## Statistical limits of high-dimensional inference problems

Présentée le 5 mars 2021
Faculté informatique et communications
Laboratoire de théorie des communications
Programme doctoral en informatique et communications
pour l'obtention du grade de Docteur ès Sciences
par

## Clément Dominique LUNEAU

Acceptée sur proposition du jury
Dr O. Lévêque, président du jury
Dr N. Macris, directeur de thèse
Prof. G. Reeves, rapporteur
Prof. A. S. Bandeira, rapporteur
Prof. E. Telatar, rapporteur

## Résumé

Cette thèse s'intéresse à deux types de problèmes d'inférence statistique en traitement du signal et en science des données. Le premier problème est l'estimation d'un tenseur informatif et structuré à partir de l'observation d'une version bruitée de celui-ci. La structure se manifeste par la possibilité de décomposer le tenseur informatif en une somme d'un petit nombre (relativement à sa dimension) de tenseurs de rang 1. Une telle structure trouve des applications en science des données où les données, organisées dans des tableaux (multidimensionnels), peuvent souvent s'expliquer par l'interaction de quelques variables explicatives qui sont caractéristiques du problème étudié. Le second problème est l'estimation d'un signal à l'entrée d'un réseau de neurones à propagation avant et dont la sortie est observée. Etudier un tel problème est important pour de nombreuses applications (phase retrieval, signaux discrétisés) où la relation entre les mesures collectées et les quantités que l'on souhaite connaître n'est pas linéaire.

Nous analysons ces deux modèles statistiques dans des limites de grandes dimensions qui correspondent à des situations pour lesquelles la quantité des observations et la taille du signal deviennent infiniment large. L'intérêt de tels régimes asymptotiques s'explique par les puissances de calculs et les capacités de stockage en constante augmentation, rendant possible le traitement et la manipulation de larges ensembles de données. Nous prenons une approche de théorie de l'information dans le but d'établir des limites statistiques fondamentales pour l'estimation dans des régimes de grandes dimensions. En particulier, les contributions principales de cette thèse sont les preuves de formules exactes pour les valeurs asymptotiques des information mutuelles normalisées associées aux problèmes d'inférence que nous étudions. Ce sont des formules variationnelles de petites dimensions qui peuvent néanmoins rendre compte du comportement de larges systèmes dont chaque composante interagit avec toutes les autres. Grâce à la relation entre l'information mutuelle et l'inférence bayésienne, nous utilisons les solutions de ces problèmes variationnels pour prédire de manière rigoureuse la limite de l'erreur quadratique moyenne minimale (MMSE), c'est-à-dire l'erreur du meilleur estimateur (l'estimateur bayésien). Ainsi nous pouvons comparer les performances d'algorithmes d'inférence à la limite statistique donnée par le MMSE.

Ces formules variationnelles pour l'information mutuelle sont dénommées ansätze symétriques des répliques et doivent leur nom à la méthode des répliques, une heuristique issue de la physique statistique grâce à laquelle elles peuvent être
prédites. Des preuves de la validité de ces prédictions ont commencé à apparaître au cours de la dernière décennie. En général, la stratégie de ces preuves est de montrer que l'ansatz symétrique des répliques est à la fois une borne supérieure et inférieure pour la limite de l'information mutuelle normalisée. Le présent travail exploite la méthode d'interpolation adaptative qui propose une manière unifiée de prouver les deux bornes. Nous étendons l'interpolation adaptative à des situations où le paramètre d'ordre du problème n'est pas scalaire mais matriciel, et à des régimes de grandes dimensions qui diffèrent de ceux pour lesquels la formule "réplique-symétrique" est habituellement conjecturée. Nos preuves démontrent aussi la modularité de la méthode d'interpolation adaptative. En effet, en utilisant des modèles statistiques précédemment étudiés dans la littérature comme briques pour construire des modèles plus complexes (e.g., le signal à estimer à un certain modèle de structure), nous déterminons des formules "réplique-symétrique" pour les informations mutuelles normalisées qui sont associées à des problèmes d'estimation pertinents pour des applications modernes.

Mots-clés : inférence bayésienne, statistique de grande dimension, estimation de tenseur, modèles linéaires généralisés, information mutuelle, formules "répliquesymétrique"

## Abstract

This thesis focuses on two kinds of statistical inference problems in signal processing and data science. The first problem is the estimation of a structured informative tensor from the observation of a noisy tensor in which it is buried. The structure comes from the possibility to decompose the informative tensor as the sum of a small number of rank-one tensors (small compared to its size). Such structure has applications in data science where data, organized into arrays, can often be explained by the interaction between a few features characteristic of the problem. The second problem is the estimation of a signal input to a feedforward neural network whose output is observed. It is relevant for many applications (phase retrieval, quantized signals) where the relation between the measurements and the quantities of interest is not linear.

We look at these two statistical models in different high-dimensional limits corresponding to situations where the amount of observations and size of the signal become infinitely large. These asymptotic regimes are motivated by the ever-increasing computational power and storage capacity that make possible the processing and handling of large data sets. We take an information-theoretic approach in order to establish fundamental statistical limits of estimation in high-dimensional regimes. In particular, the main contributions of this thesis are the proofs of exact formulas for the asymptotic normalized mutual information associated with these inference problems. These are low-dimensional variational formulas that can nonetheless capture the behavior of a large system where each component interacts with all the others. Owing to the relationship between mutual information and Bayesian inference, we use the solutions to these variational problems to rigorously predict the asymptotic minimum mean-square error (MMSE), the error achieved by the (Bayes) optimal estimator. We can thus compare algorithmic performances to the statistical limit given by the MMSE.

Variational formulas for the mutual information are referred to as replica symmetric (RS) ansätze due to the predictions of the heuristic replica method from statistical physics. In the past decade proofs of the validity of these predictions started to emerge. The general strategy is to show that the RS ansatz is both an upper and lower bound on the asymptotic normalized mutual information. The present work leverages on the adaptive interpolation method that proposes a unified way to prove the two bounds. We extend the adaptive interpolation to situations where the order parameter of the problem is not a scalar but a matrix, and to high-dimensional regimes that differ from the one for which the RS formula
is usually conjectured. Our proofs also demonstrate the modularity of the method. Indeed, using statistical models previously studied in the literature as building blocks of more complex ones (e.g., estimated signal with a model of structure), we derive RS formulas for the normalized mutual information associated with estimation problems that are relevant to modern applications.

Keywords: Bayesian inference, high-dimensional statistics, tensor estimation, generalized linear models, mutual information, replica symmetric formulas

## Acknowledgements

First and foremost, I want to express my sincere gratitude to Nicolas Macris for being my supervisor. He guided me all along my PhD journey, taught me many tricks of the trade, and it was an honor to work and discuss challenging questions with him. Most importantly he always showed his support when I was filled with doubts about where my research was going.

I am very grateful to Jean Barbier who was a postdoc in our lab during the first half of my PhD. Jean always has a lot of ideas that he is willing to share. For many problems studied in this thesis he was the impulse that started it all. He also gave me confidence, welcoming me to ENS for a couple of weeks to join forces on different problems and inviting me as a speaker for a workshop at ICTP.

It was an honor to have Afonso Bandeira, Olivier Lévêque, Galen Reeves and Emre Telatar on my thesis committee. I am thankful for the time they spent reading this work and their insightful comments. Knowing that they appreciated this manuscript made the efforts put in writing it all worth it.

I would like to thank Rüdiger Urbanke for trusting me and giving me the opportunity to join his lab. I spent most of my PhD studies in the INR building, sheltered by the Information Processing Group (a big family of four labs). Kudos to Muriel Bardet, Françoise Behn, France Faille and Damid Laurenzi for keeping IPG running smoothly and making it such a nice place to do research. I especially thank Muriel for her help and always answering so promptly to all of my administrative questions, and Damir for his support on everything touching to IT. Many thanks should also go to all the current and past members of the IPG microcosm, in particular Amadeo, Daria, Elie, Eric, Erixhen, Emanuele, Karol, Kirill, Mohamad, Raffaele, Rajai, Reka, Su and Wei.

I am happy to have had Mohamad and Eric as academic older brothers. They guided me through many steps of my PhD life and are still there when I need their help. Eric and Ying Mao were close neighbors; I owe them many fun weekends playing board/video games and sharing delicious food.

I was extremely lucky to share my office with Elie. I could not have hope a better friend through the hard(-working) days. He was always there to cheer me up, even after he graduated. He has a great spirit, and a great laugh too.

A big hug to my friend Adrien. We shared many coffee breaks at EPFL that I am thankful for. Talking with him is never boring and it is the best way to unwind when stuck on something. It is also important to have a fellow French to complain with like French people do.

Adrien, Elie, Eric and Mohamad are very humble and one can count oneself lucky when surrounded by such positive forces during a PhD life that can be too competitive.

There was not a day where I didn't talk with my friends in the "mifa": Dimitri, Guillaume, Maxime, Nicolas, Thibaud, Xavier and Yves. Even now, where we can only see each other in Grenoble over the holidays, it feels like you are never far. Thank you for having my back since primary school. I also want to thank my friends Alan, Christopher, Etienne and Paul. We don't see each other as much as I'd like but when we do these are always great moments.

None of this work would have been possible without the support and affection of my mother, Isabelle, and my father, Dominique. As a son of two chemists and crystallographers, I grew up seeing myself doing science in a lab full of colorful potions and weird compounds. I took a path safer than experimental research, doing science with pen and paper. Thank you to all my family and family-in-law, in particular Marion and my two brothers Josselin and Aymeric.

There are no words strong enough to thank my wife, María Fernanda, for her endless support, all the joy she brings me and her love.

Lausanne, February 8 ${ }^{\text {th }}$, 2021.
C.L.

## Contents

Résumé ..... i
Abstract ..... iii
Acknowledgements ..... v
Contents ..... vii
1 Introduction ..... 1
1.1 Tensor estimation ..... 2
1.2 Generalized linear models ..... 5
1.3 Optimal Bayesian estimation in high dimensions ..... 9
1.4 Mutual information and estimation ..... 12
1.5 A detour to statistical physics ..... 15
1.6 Replica symmetric formula for the mutual information ..... 21
1.7 Proving the replica predictions ..... 25
1.8 Organization and main contributions ..... 27
2 Toolbox for proofs of replica symmetric formulas ..... 31
2.1 Nishimori identity ..... 31
2.2 Gaussian integration by parts ..... 32
2.3 Mutual information of simple Gaussian channels ..... 33
2.4 Concentration inequalities ..... 36
I Tensor estimation ..... 39
3 High-dimensional rank-one nonsymmetric matrix estimation: the spherical case ..... 41
3.1 Introduction ..... 41
3.2 Problem setting and main results ..... 42
3.3 Proof of Theorem 13.2 ..... 44
3.4 Conclusion ..... 52
Appendices ..... 55
3.A Proof of Lemma|3.1 ..... 55
3.B Sum-rule of Proposition 3.4 ..... 57
3.C Properties of the interpolation paths ..... 62
3.D Concentration of the overlap ..... 64
3.E Concentration of the free entropy ..... 70
3.F Formula for the asymptotic MMSE ..... 72
4 Mutual information for low-rank even-order symmetric tensor estimation ..... 77
4.1 Introduction ..... 77
4.2 Low-rank symmetric tensor estimation ..... 78
4.3 Adaptive path interpolation ..... 81
4.4 Matching bounds ..... 82
4.5 Derivative of the average interpolating free entropy ..... 87
4.6 Concentration of the overlap matrix ..... 90
4.7 Conclusion and discussion for odd-order tensors ..... 104
Appendices ..... 107
4.A Divergence of the function $G_{n}$ ..... 107
4.B Concentration of the free entropy ..... 111
5 Tensor estimation with structured priors ..... 117
5.1 Introduction ..... 117
5.2 Asymptotic mutual information and MMSE for tensor estimation ..... 119
5.3 Examples of phase transitions and their properties ..... 123
5.4 Proof of the variational formula for the mutual information ..... 126
5.5 Derivation of the asymptotic Tensor-MMSE ..... 135
5.6 Limit of vanishing $\alpha$ ..... 136
Appendices ..... 139
5.A Auxiliary lemmas ..... 139
5.B Establishing the sum-rule ..... 141
5.C Concentration of the overlap ..... 145
5.D Concentration of the free entropy ..... 152
5.E Proof of Proposition 5.7 ..... 157
II Generalized linear models ..... 161
6 Entropy and mutual information in feedforward neural networks163
6.1 Introduction ..... 163
6.2 Two-layer generalized linear model ..... 164
6.3 Proof of the replica symmetric formula ..... 169
Appendices ..... 185
6.A Miscellaneous useful results ..... 185
6.B Interpolating average free entropy at $t=0$ ..... 187
6.C Derivative of the averaged interpolating free entropy ..... 190
6.D Concentration of the interpolating free entropy ..... 198
6.E Concentration of the overlap ..... 207
7 Information theoretic limits of learning a sparse rule ..... 213
7.1 Introduction ..... 213
7.2 Problem setting and main results ..... 216
7.3 The all-or-nothing phenomenon ..... 220
7.4 Proof of Theorem 7.2 |for a Bernoulli prior ..... 225
Appendices ..... 233
7.A Proof of Theorem 7.1 with the adaptive interpolation method ..... 233
7.B Proof of Theorem 7.2 for a general discrete prior ..... 240
7.C Asymptotic minimum mean-square error: proof of Theorem 7.3 ..... 252
7.D Properties of the mutual information $I_{P_{\text {out }}}$ ..... 262
7.E Properties of the interpolating mutual information ..... 266
7.F Concentration of the free entropy ..... 276
7.G Concentration of the overlap ..... 284
7.H Proof of Proposition 7.12 ..... 290
8 Conclusion and possible directions ..... 295
Bibliography ..... 299

## Introduction

## 1

When Claude Shannon laid the mathematical foundations of communication and established information theory in 1948, he demonstrated the existence and achievability of a fundamental limit to communication, namely, the channel capacity, the maximum rate at which information can be reliably transmitted over a communication channel. This upper limit on the information rate is given by an optimization problem where we look for the input distribution maximizing the mutual information between the input and output of the channel. While Shannon's proof does not say how to build a code that approaches the channel capacity, it highlighted the gap between the information rates of the error-correcting codes available at that time and optimal codes achieving capacity. It set a goal and gave impetus to the design of practical capacity-approaching codes.

Since then, information theory concepts and tools have been applied well beyond traditional communication channels for which they were initially developed. In this thesis, we take an information-theoretic approach to establish fundamental limits for estimation and learning problems found in signal processing, machine learning or data science. We analyze statistical models for tensor decomposition and feedforward neural networks where the goal is to estimate, with as little error as possible, a quantity of interest from noisy observations. In that regard, we study the mutual information and leverage on its connection with Bayes optimal inference to determine the minimum error that is statistically achievable.

As computational power and storage capacity steadily increased over the past decades, there is a growing interest in tackling problems involving large amounts of data. The main contribution of the present work is the proof of exact formulas for the mutual information associated with such high-dimensional problems. These formulas are low-dimensional optimization problems that can be solved numerically. Crucially, the solutions to the latter are linked to the error of the Bayes optimal estimator, the minimum error statistically achievable. Although the models that we examine are simplifications and idealizations of the situations encountered in practice, they offer a precise framework in which we can compare the error achieved by diverse reconstruction techniques to the minimum
error. In order to prove these asymptotic formulas for the mutual information, we draw on the expertise of statistical physics regarding the analysis of large systems.

In the next two sections we introduce and motivate the two kinds of highdimensional statistical models that are the focus of this thesis. In Sections 1.3 and 1.4 we explain how to use the mutual information to study the fundamental limits of estimation in these models. In Section 1.5 we take a detour to statistical physics, a field that aims to describe how macroscopic properties of large physical systems (magnetization, density, index of refraction, etc.) arise from the interactions of its microscopic constituents (particles). An important quantity in statistical physics is the free energy whose form is similar to a mutual information. In the thermodynamic limit where the physical system becomes infinitely large, a formula for the free energy can be predicted with the heuristic replica method. In Section 1.6 we show how the replica ansatz for the mutual information yields a precise characterization of the statistical limits of estimation problems. However, the replica method is not mathematically justified. The important results of this thesis are the proofs of the correctness of the replica predictions for the different problems that we study. In Section 1.7 we give some background on the techniques developed to prove these formulas. In particular, we present the Guerra-Toninelli interpolation method on which the present work builds upon. We conclude this chapter with a section outlining the organization of the thesis and contribution of each chapter.

### 1.1 Tensor estimation

Data is commonly stored and organized into arrays, either matrices or tensors. The latter are multidimensional arrays; a natural generalization of matrices to dimensions greater than two. Matrix factorizations like singular value decomposition (SVD) and principal component analysis (PCA) are widely used to discover structure in a 2D dataset [1]. Likewise, tensor decompositions are techniques to factorize tensors in order to extract information, e.g., what are the latent factors whose interactions give rise to the observed dataset. In essence, they aim to generalize different desirable properties of matrix SVD and PCA to higher-order tensors. Research on tensor decompositions started in the late 1920s [2], and until the 1990s their theoretical and algorithmic developments were essentially investigated by the psychometric [3], [4] and chemometric [5] communities. In the last twenty years the use of tensor decomposition techniques expanded outside these communities and found its way to computer science, in particular in machine learning and data science [6], [7]. Applications include parameter estimation in latent variable models [8], [9], community detection [10], collaborative filtering [11] and graph matching for computer vision [12]. The review articles [6], [13], 14] are good introductions to tensors and the two most prominent decompositions the canonical polyadic (CP) and Tucker decompositions.

CP decomposition We now introduce some notations and definitions on tensors, and then present the CP decomposition. We say that $\mathbf{T}$ is a real
order- $p$ tensor if its elements are real numbers indexed by $p$ indices, that is, $\mathbf{T}:=\left\{T_{i_{1} i_{2} \ldots i_{p}}\right\}_{\forall \ell: 1 \leq i_{\ell} \leq n_{\ell}} \in \mathbb{R}^{n_{1} \times n_{2} \times \cdots \times n_{p}}$ where each dimension $n_{\ell}$ is a positive integer. The outer product $\mathbf{v}^{(1)} \otimes \mathbf{v}^{(2)} \otimes \cdots \otimes \mathbf{v}^{(p)}$ of $p$ vectors $\mathbf{v}^{(\ell)}:=\left\{v_{i}^{(\ell)}\right\}_{i=1}^{n_{\ell}} \in \mathbb{R}^{n_{\ell}}$, $\ell=1 \ldots p$, is the real order $-p$ tensor whose elements are

$$
\left(\mathbf{v}^{(1)} \otimes \mathbf{v}^{(2)} \otimes \cdots \otimes \mathbf{v}^{(p)}\right)_{i_{1} i_{2} \ldots i_{p}}:=v_{i_{1}}^{(1)} v_{i_{2}}^{(2)} \cdots v_{i_{p}}^{(p)}
$$

If all of the vectors $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(p)}$ are nonzero then $\mathbf{v}^{(1)} \otimes \mathbf{v}^{(2)} \otimes \cdots \otimes \mathbf{v}^{(p)}$ is called a rank-one tensor. The rank of a tensor $\mathbf{T}$ is the minimum number of rank-one tensors whose sum yields exactly $\mathbf{T}$; it is denoted $\operatorname{rank}(T)^{17}$. For a fixed positive integer $K$, the CP decomposition of a tensor $\mathbf{T}$ is the best approximation of $\mathbf{T}$ as the sum of $K$ rank-one tensors. More precisely, this is the solution to

$$
\begin{equation*}
\min _{\widehat{\mathbf{T}} \in \mathcal{T}_{K}}\|\mathbf{T}-\widehat{\mathbf{T}}\| \tag{1.1}
\end{equation*}
$$

where

$$
\begin{equation*}
\mathcal{T}_{K}:=\left\{\sum_{k=1}^{K} X_{\cdot, k}^{(1)} \otimes X_{\cdot, k}^{(2)} \otimes \cdots \otimes X_{\cdot, k}^{(p)}: \mathbf{X}^{(\ell)} \in \mathbb{R}^{n_{\ell} \times K}, \ell=1 \ldots p\right\} \tag{1.2}
\end{equation*}
$$

In the above, $X_{\cdot, k}$ denotes the $k^{\text {th }}$ column of a matrix $\mathbf{X}$ and $\|\cdot\|$ is the usual Euclidean norm (also called Frobenius norm), i.e., $\|\mathbf{T}\|^{2}:=\sum_{i_{1}, i_{2}, \ldots, i_{p}}\left(T_{i_{1} i_{2} \ldots i_{p}}\right)^{2}$. For matrices, the solution to (1.1) is given by the Eckart-Young-Mirsky theorem [15]. For higher-order tensors, (1.1) is an ill-posed problem because the set $\mathcal{T}_{K}$ on which we minimize is not closed [16]. A natural way to fix the ill-posedness of (1.1) is to look for weak solutions, that is, to minimize the objective function $\|\mathbf{T}-\mathbf{T}\|$ on the closure of $\mathcal{T}_{K}$ instead. For example, 16 gives a characterization of the closure when $K=2, p=3$ that can be used to define an objective function amenable to minimization.

Statistical models for CP decomposition Solving (1.1) for a general tensor $\mathbf{T}$ is computationally hard [17]. Instead, we adopt the point of view of tensor estimation which is the task of extracting meaningful information from a noisy data tensor. A prototypical example is community detection from observed interactions between individuals (18).

Example 1.1 (Symmetric two-group stochastic block model). Consider $n$ individuals divided into two communities $\mathcal{C}_{-1}$ and $\mathcal{C}_{+1}$. For the individual $i \in\{1, \ldots, n\}$ we define $X_{i}=-1$ if they belong to the community $\mathcal{C}_{-1}$ and $X_{i}=1$ if they belong to $\mathcal{C}_{+1}$. Each individual has a fifty-fifty chance to be part of either $\mathcal{C}_{-1}$ or $\mathcal{C}_{+1}$, i.e., $\mathbb{P}\left(X_{i}=1\right)=\mathbb{P}\left(X_{i}=-1\right)=0.5$, and this independently of the others. Define the column vector $\mathbf{X}=\left[X_{1}, X_{2}, \ldots, X_{n}\right]^{\top}$. Individuals interact with each other. Let $Y_{i j}=1$ if $i$ and $j$ know each other, and $Y_{i j}=-1$ otherwise. We assume that the $Y_{i j}$ 's are distributed according to the model

$$
\begin{equation*}
\mathbb{P}\left(Y_{i j}=1 \mid X_{i} X_{j}=1\right)=p_{n}, \quad \mathbb{P}\left(Y_{i j}=1 \mid X_{i} X_{j}=-1\right)=q_{n}, \tag{1.3}
\end{equation*}
$$

[^0]where $0 \leq q_{n}<p_{n} \leq 1$. Under this model, two individuals are more likely to know each other if they belong to the same community. The matrix $\mathbf{Y}=\left\{Y_{i j}\right\}$ can be seen as a noisy version of the rank-1 matrix $\mathbf{X} \mathbf{X}^{\top}$; each entry of $\mathbf{X} \mathbf{X}^{\top}$ is observed through the channel (1.3). Suppose that we do not know the composition of each community; however, we know if two indivuals are familiar with each other. In a nutshell, we know $\mathbf{Y}$ but not $\mathbf{X}$. One of the central question in community detection is whether we can recover $\mathbf{X X} \mathbf{X}^{\top}$ from $\mathbf{Y}$. Note that we can easily determine $\mathbf{X}$ up to a plus or minus sign from $\mathbf{X X} \mathbf{X}^{\top}$.

In this thesis, we study different statistical models for tensor decomposition as proposed in [19]. The observed data is a noisy version of a finite-rank tensor of interest:

Model 1.1 (Low-rank asymmetric tensor estimation). Let $\mathbf{X}^{(\ell)} \in \mathbb{R}^{\left\lfloor\alpha_{\ell} n\right\rfloor \times K}$, $\ell=1, \ldots, p$, be independent random matrices where the factors $\alpha_{1}, \ldots, \alpha_{p}$ are positive real numbers and the dimensions $n, K$ are positive integers. Define the order- $p$ tensor

$$
\mathbf{T}:=\sum_{k=1}^{K} X_{\cdot, k}^{(1)} \otimes X_{\cdot, k}^{(2)} \otimes \cdots \otimes X_{\cdot, k}^{(p)}
$$

We observe $\mathbf{T}$ under an additive white Gaussian noise (AWGN), i.e., we observe

$$
\mathbf{Y}:=n^{(1-p) / 2} \mathbf{T}+\sqrt{\Delta} \mathbf{Z}
$$

where the elements of $\mathbf{Z} \in \mathbb{R}^{\left\lfloor\alpha_{1} n\right\rfloor \times \cdots \times\left\lfloor\alpha_{p} n\right\rfloor}$ are independent standard Gaussians and $\Delta>0$ is the noise variance.

Model 1.2 (Low-rank symmetric tensor estimation). Let $\mathbf{X} \in \mathbb{R}^{n \times K}$ be a random matrix where the dimensions $n, K$ are positive integers. Define the order- $p$ tensor

$$
\mathbf{T}:=\sum_{k=1}^{K} X_{\cdot, k} \otimes X_{\cdot, k} \otimes \cdots \otimes X_{\cdot, k}
$$

The tensor $\mathbf{T}$ is symmetric, meaning that $T_{i_{\pi(1)} i_{\pi(2)} \ldots i_{\pi(p)}}=T_{i_{1} i_{2} \ldots i_{p}}$ for any permutation $\pi:\{1, \ldots, p\} \rightarrow\{1, \ldots, p\}$ and $p$-tuple $\left(i_{1}, i_{2}, \ldots, i_{p}\right)$. We observe $\mathbf{T}$ under a symmetric AWGN, i.e., we observe

$$
\mathbf{Y}:=n^{(1-p) / 2} \mathbf{T}+\sqrt{\Delta} \mathbf{Z}
$$

where $\mathbf{Z} \in \mathbb{R}^{n \times n \times \cdots \times n}$ is a symmetric tensor whose elements $Z_{i_{1} i_{2} \ldots i_{p}}$ are independent standard Gaussians for $1 \leq i_{1} \leq \cdots \leq i_{p} \leq n$ and $\Delta>0$ is the noise variance.

In both models the definition of the underlying tensor $\mathbf{T}$ is reminiscent of a CP decomposition and entails that the rank of $\mathbf{T}$ is at most $K$. We study these models in a high-dimensional regime where $K$ is negligible compared to the dimensions of the tensor; hence the adjective "low-rank" in their names. Specifically, we fix $K$ (and each factor $\alpha_{\ell}$ in the case of Model 1.1) and take $n$ infinitely large.

In the descriptions that we give, the distributions of the matrices $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(p)}$ or $\mathbf{X}$ are voluntarily kept general. In Chapter 3 we analyze Model 1.1 in the case $p=2, K=1$ and $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}$ are uniformly distributed on spheres of radii proportional to $\sqrt{n}$. In Chapter 4 we consider Model 1.2 for $\mathbf{X}$ having independent and identically distributed (i.i.d.) rows. In this case, unless the support of the distribution of the rows spans a proper subspace of $\mathbb{R}^{K}, \mathbf{X}$ has asymptotically almost surely full rank $]^{2}$ and, by Kruskal's theorem $[20$, the rank of $\mathbf{T}$ is exactly $K$.

We are usually interested in problems where each entry of the tensor $\mathbf{T}$ is observed at the output of a noisy channel described by a conditional probability distribution $P_{\text {out }}(y \mid x)$, that is, each element of the data $\mathbf{Y}$ satisfies

$$
Y_{i_{1} i_{2} \ldots i_{p}} \sim P_{\text {out }}\left(\cdot \mid n^{(1-p) / 2} T_{i_{1} i_{2} \ldots i_{p}}\right),
$$

independently of the others. In Models 1.1 and 1.2. $P_{\text {out }}(y \mid x)=e^{\frac{-(y-x)^{2}}{2 \Delta}} / \sqrt{2 \pi \Delta}$; far from a channel like (1.3) in Example 1.1. However, the AWGN is not as restrictive as it might seem at first sight. Indeed, in the high-dimensional regime that we consider, many channels turn out to be equivalent to an AWGN channel with appropriate noise variance $\Delta$. This equivalence, called channel universality, is proved in [21] for a class of channels satisfying some regularity assumptions. For the channel of Example 1.1, the equivalence is proved in [18] (for an appropriate scaling of $p_{n}$ and $q_{n}$ as $n$ diverges) and allows the authors to study the symmetric two-group SBM through the analysis of Model 1.2 (for $K=1, p=2$ ).

One achievement of this thesis is the determination of fundamental statistical limits on how well we can recover the tensor $\mathbf{T}$ of interest from the noisy data $\mathbf{Y}$. More precisely, we establish exact formulas for the performance of the optimal estimator of $\mathbf{T}$, based on the observation $\mathbf{Y}$, in the high-dimensional regime. Of course, the optimal estimator depends on how we measure performances. In this work, we assess the performance of an estimator $\widehat{\mathbf{T}}(\mathbf{Y})$ in terms of its mean-square error (MSE) $n^{-p} \mathbb{E}\|\mathbf{T}-\widehat{\mathbf{T}}(\mathbf{Y})\|^{2}$, in which case the Bayes estimator $\mathbb{E}[\mathbf{T} \mid \mathbf{Y}]$ is optimal and its error is referred to as the minimum mean-square error (MMSE). As the MMSE is a lower bar on the error of any estimator, it constitutes a limit to approach as closely as possible by an algorithm (subject or not to constraints on its computational complexity); the MMSE acts as a reference to which we can compare existing algorithms. Before diving further into Bayesian inference and the MMSE in Section 1.3, we first present the other estimation problem studied in this thesis.

### 1.2 Generalized linear models

Linear models of the form

$$
\begin{equation*}
\mathbf{Y}:=\mathbf{W X}+\mathbf{Z}, \tag{1.4}
\end{equation*}
$$

[^1]where $\mathbf{X}$ is a signal of interest, $\mathbf{W}$ a known measurement matrix and $\mathbf{Z}$ an AWGN with noise variance $\sigma^{2}$, are ubiquitous in statistical inference and least-squares methods. Despite their usefulness, there are many applications for which the observations are rather nonlinear transformations of linear measurements. Classical examples include phase retrieval problems in optics and X-ray crystallography [22], [23] where only the amplitude of the linear measurements are observed, and dequantization [24], [25] when the measurements are discretized due to digital recording and further quantized for compression purposes. In linear models (1.4), the observations $Y_{1}, Y_{2}, \ldots, Y_{m}$ are conditional on $\mathbf{X}$ independent normal random variables of same variance $\sigma^{2}$ and means $(\mathbf{W X})_{1},(\mathbf{W X})_{2}, \ldots,(\mathbf{W X})_{m}$. Generalized linear models [26] have an increased expressivity by extending the linear ones in two ways: (i) the conditional distribution of the observations belongs to an exponential family, and (ii) the conditional means are $g^{-1}\left((\mathbf{W X})_{i}\right)$ for $i=1 \ldots m$ where $g$ is refered as the link function in statistics. In this thesis, we consider a more general form of generalized linear models (GLMs), perhaps more familiar to readers with a machine learning background.

Model 1.3 (1-layer GLM). Let $\mathbf{X}$ be a $n$-dimensional random vector of interest, $k_{A}$ a natural number, $P_{A}$ a probability distribution on $\mathbb{R}^{k_{A}}$, and $\varphi: \mathbb{R} \times \mathbb{R}^{k_{A}} \rightarrow \mathbb{R}$ a function (known as an activation function in machine learning). In the one-layer GLM, we are given a $m \times n$ matrix $\mathbf{W}$ and $m$ observations of the form

$$
\begin{equation*}
Y_{i}:=\varphi\left(\frac{(\mathbf{W X})_{i}}{\sqrt{n}}, \mathbf{A}_{i}\right)+\sqrt{\Delta} Z_{i} \tag{1.5}
\end{equation*}
$$

where $\mathbf{A}:=\left\{\mathbf{A}_{i}\right\}_{i=1}^{m} \stackrel{\text { i.i.d. }}{\sim} P_{A}, \mathbf{Z}:=\left\{Z_{i}\right\}_{i=1}^{m} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ is an AWGN and $\Delta \geq 0$ is a noise variance parameter.

Remark. We can recover the original GLM of statistics by choosing $\Delta:=0$, $k_{A}:=1, P_{A}:=\mathcal{U}([0,1])$ the uniform distribution on $[0,1]$, and

$$
\varphi(x, a):=F\left(g^{-1}(x), a\right)
$$

where the function $F$ is such that the distribution of $F(\mu, A), A \sim \mathcal{U}([0,1])$, is the desired exponential family distribution with mean $\mu \in \mathbb{R}$.

Going back to the examples given earlier, the phase retrieval problem corresponds to $\varphi(x, A)=|x+A|$ with $A \sim \mathcal{N}(0,1)$, and 1-bit compressed sensing (quantization with a 1 -bit quantizer) to $\varphi(x, A)=\operatorname{sign}(x+A)$ with $A \sim \mathcal{N}(0,1)$ and sign the sign function.

We recognize in $\varphi(\mathbf{W} \mathbf{x} / \sqrt{n}, \mathbf{A}):=\left\{\varphi\left((\mathbf{W} \mathbf{X})_{i} / \sqrt{n}, \mathbf{A}_{i}\right)\right\}_{i=1}^{n}$ the layer of a feedforward neural network where $\mathbf{X}$ is the input, each row of $\mathbf{w} / \sqrt{n}$ the weights of one of the $m$ nodes in the layer, and $\varphi\left((\mathbf{W} \mathbf{X})_{i} / \sqrt{n}, \mathbf{A}_{i}\right)$ the output of the $i^{\text {th }}$ node. The observations $\mathbf{Y}:=\left\{Y_{i}\right\}_{i=1}^{m}$ in Model 1.3 then correspond to noisy versions of these outputs. In Figure 1.1 we draw a graphical representation of Model 1.3 that suits well the inference point of view that we have just described. From this point of view, $\mathbf{X}$ is the input of a multi-output one-layer neural network and we want to estimate $\mathbf{X}$ from the noisy outputs. Although we mostly focus on the latter


Figure 1.1: Graphical representation of Model 1.3 with an inference point of view. The matrix $\mathbf{W} / \sqrt{n}$ are the weights of a one-layer feedforward neural network with $n$ input nodes and $m$ output nodes, and $\mathbf{Y}$ is a noisy output produced by feeding $\mathbf{X}$ to the neural network.
interpretation, the "learning" point of view is worth mentioning. We now see $\mathrm{X} / \sqrt{n}$ as the weights of a one-layer neural network with $n$ inputs nodes and a unique output node. When a row of $\mathbf{W}$ is fed to the neural network, the preactivation $\sum_{j=1}^{n} \mathbf{W}_{i j} X_{j} / \sqrt{n}$ passes through the stochastic activation function ${ }^{3} \varphi\left(\cdot, \mathbf{A}_{i}\right), \mathbf{A}_{i} \sim P_{A}$, thus producing an output that we observe after it has been corrupted by the noise $\sqrt{ } \Delta Z_{i}$. We draw in Figure 1.2 a graphical representation of Model 1.3 that is more suited to this learning point of view. In this interpration, the task is to use the $m$ input-output pairs $\left(\left\{W_{1 j}\right\}_{j=1}^{n}, Y_{1}\right),\left(\left\{W_{2 j}\right\}_{j=1}^{n}, Y_{2}\right), \ldots,\left(\left\{W_{m j}\right\}_{j=1}^{n}, Y_{m}\right)$ in order to learn the weights $\mathbf{X}$. Here, more than the mean-square error of the estimate of $\mathbf{X}$, the important metric is the generalization error. We want to learn the weights to accurately predict the label $Y_{\text {new }}$ of an input $\left\{W_{\text {new }, i}\right\}_{i=1}^{n}$ that has never been seen before.

In recent years, deep neural networks built by stacking numerous layers on top of each other have been used as probabilistic generative models of complex data. We refer to these feedforward neural networks as multilayer GLMs and we give a more formal definition below [27], [28].

Model 1.4 ( $L$-layer GLM). Let X be a $n$-dimensional random vector of interest and $L$ a natural number. For $\ell \in\{1, \ldots, L\}$, let $k_{\ell}$ be a natural number, $P_{A}^{(\ell)}$ a probability distribution on $\mathbb{R}^{k_{\ell}}$, and $\varphi_{\ell}: \mathbb{R} \times \mathbb{R}^{k_{\ell}} \rightarrow \mathbb{R}$ an activation function. For $\ell \in\{1, \ldots, L\}$, let $n_{\ell}$ be a positive integer and $\mathbf{W}^{(\ell)}$ a $n_{\ell} \times n_{\ell-1}$ matrix, with $n_{0}:=n$. Starting from $\mathbf{X}^{(0)}:=\mathbf{X}$, define recursively $\forall \ell \in\{1, \ldots, L\}:$

$$
\mathbf{X}^{(\ell)}:=\varphi_{\ell}\left(\frac{\mathbf{W}^{(\ell)} \mathbf{X}^{(\ell-1)}}{\sqrt{n_{\ell-1}}}, \mathbf{A}^{(\ell)}\right)
$$

[^2]

Figure 1.2: Graphical representation of Model 1.3 with a learning point of view. The vector $\mathbf{X} / \sqrt{n}$ are the weights of a one-layer feedforward neural network with $n$ input nodes and 1 output node. The neural network is fed $m$ different inputs $\left\{W_{1 j}\right\}_{j=1}^{n},\left\{W_{2 j}\right\}_{j=1}^{n}, \ldots,\left\{W_{m j}\right\}_{j=1}^{n}$ and produces $m$ outputs $Y_{1}, Y_{2}, \ldots, Y_{m}$.
where $\mathbf{A}^{(\ell)}:=\left\{\mathbf{A}_{i}^{(\ell)}\right\}_{i=1}^{n_{\ell}} \stackrel{\text { i.i.d. }}{\sim} P_{A}^{(\ell)}$ and $\varphi_{\ell}$ is applied componentwise. In the $L$-layer GLM, we are given $\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(L)}$ and $n_{L}$ observations of the form

$$
Y_{i}:=X_{i}^{(L)}+\sqrt{\Delta} Z_{i} ;
$$

where $\mathbf{Z}:=\left\{Z_{i}\right\}_{i=1}^{n_{L}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ is an AWGN and $\Delta \geq 0$ is a noise variance parameter.

Note that Model 1.3 is a special case of Model 1.4 for $L=1$. In machine learning, we have access to a set of $M$ training pairs $\left\{\left(\mathbf{X}_{j}, \mathbf{Y}_{j}\right)\right\}_{j=1}^{M}$ and we optimize the weight matrices $\mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \ldots, \mathbf{W}^{(L)}$ in order to minimize the error of predicting $\left\{\mathbf{Y}_{j}\right\}_{j=1}^{M}$ using the outputs obtained by feeding each input $\mathbf{X}_{j}$ to the neural network. The end goal is that the predictive power of the neural networks generalizes to previously unseen inputs; an important consideration to take into account when optimizing the weights. In Model 1.4, we know the weights and we observe the noisy output $\mathbf{Y}:=\left\{Y_{i}\right\}_{i=1}^{n_{L}}$. Our goal is to estimate the input $\mathbf{X}$ that produced the observations. We thus study multilayer GLMs from an inference point of view.

We analyze the inference problem in two distinct high-dimensional regimes. In Chapter 6, we look at Model 1.4 in the high-dimensional regime where all of the dimensions $n_{0}, n_{1}, \ldots, n_{L}$ diverge to infinity while the ratios $n_{\ell} / n_{\ell-1}$ for $\ell \in\{1, \ldots, L\}$ are kept fixed. This regime was first rigorously studied in [29] for the 1-layer GLM and we show how to extend the analysis to the 2-layer GLM. In Chapter 7, we consider Model 1.3 in the high-dimensional regime where the input dimension $n$ diverges to infinity but the number $m$ of observations as well as the sparsity $\|\mathbf{X}\|_{0}$ of $\mathbf{X}$ (the amount of nonzero components in $\mathbf{X}$ ) are sublinear in $n$, that is, both $m / n$ and $\|\mathbf{X}\|_{0} / n$ vanish as $n$ diverges. This regime is motivated by the observation, popularized by compressed sensing [30], [31], that the sparsity of a signal $\mathbf{X}$ can be exploited to recover $\mathbf{X}$ with fewer observations than what would
be usually required. In the high-dimensional regime of sublinear sparsity and number of measurements, the MMSE differs both qualitatively and quantitatively from the one in the classical high-dimensional regime of Chapter 6. This was first evidenced in $[32]$ and shown rigorously in $[33]$ for the linear model with binary signal X.

### 1.3 Optimal Bayesian estimation in high dimensions

We have now introduced the two classes of problems for which we study estimation. In this section, we present Bayes optimal inference in a discussion tailored to the different models of Sections 1.1 and 1.3

Bayesian inference Let $\mathbf{X}$ be a set of real variables whose values we do not know. Instead, we have some observations $\mathbf{Y}$ that bear a relation to $\mathbf{X}^{4}$. For example, $\mathbf{Y}$ can be a noisy measurement of $\mathbf{X}$ or a function of $\mathbf{X}$, see Models 1.1t to 1.4. On the basis of these observations, we want to deduce (statistical) properties of $\mathbf{X}$. This process, called statistical inference, requires making assumptions on the generation of the data $\mathbf{Y}$ and on the distribution of $\mathbf{X}$ when devoid of any observation. In Bayesian inference, these assumptions take the forms of "beliefs"; a prior belief $B_{\mathbf{X}}$ on the distribution of $\mathbf{X}$ and a likelihood belief $B_{\mathbf{Y} \mid \mathbf{X}}$ on the distribution of $\mathbf{Y}$ given a specific realization of $\mathbf{X}$. Bayesian inference then computes a posterior belief following Bayes' rul $\square^{5}$

$$
\begin{equation*}
B_{\mathbf{X} \mid \mathbf{Y}}(\mathbf{x} \mid \mathbf{y}):=\frac{B_{\mathbf{X}}(\mathbf{x}) B_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{y} \mid \mathbf{x})}{\int B_{\mathbf{X}}\left(\mathbf{x}^{\prime}\right) B_{\mathbf{Y} \mid \mathbf{X}}\left(\mathbf{y} \mid \mathbf{x}^{\prime}\right) d \mathbf{x}^{\prime}} . \tag{1.6}
\end{equation*}
$$

We can use the posterior belief to estimate $\mathbf{X}$, e.g., with the maximum a posteriori (MAP) estimator $\arg \max _{\mathbf{x}} B_{\mathbf{X} \mid \mathbf{Y}}(\mathbf{x} \mid \mathbf{Y})$ or the Bayes estimator $\mathbb{E}_{\mathbf{x} \sim B_{\mathbf{X} \mid \mathbf{Y}}(\cdot \mid \mathbf{Y})}[\mathbf{x}]$. Of course, wrong assumptions lead to faulty conclusions.

Bayes optimal inference refers to the ideal situation where the beliefs exactly match the true distributions, that is, we know both the prior distribution $P_{\mathbf{X}}$ and the likelihood distribution $P_{\mathbf{Y} \mid \mathbf{X}}$. In this case, Bayes' rule is an identity for the posterior distribution of $\mathbf{X}$ given a particular realization of the observations $\mathbf{Y}$,

$$
\begin{equation*}
P_{\mathbf{X} \mid \mathbf{Y}}(\mathbf{x} \mid \mathbf{y})=\frac{P_{\mathbf{X}}(\mathbf{x}) P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{y} \mid \mathbf{x})}{\int P_{\mathbf{X}}\left(\mathbf{x}^{\prime}\right) P_{\mathbf{Y} \mid \mathbf{X}}\left(\mathbf{y} \mid \mathbf{x}^{\prime}\right) d \mathbf{x}^{\prime}} . \tag{1.7}
\end{equation*}
$$

It is fair to say that such a situation is never achieved in practice. However, in all of Models 1.1 to 1.4 we fully describe the process generating $\mathbf{Y}$ and, although these models do not make any assumption on $\mathbf{X}$, we analyze them for a well-specified

[^3]prior (e.g., Model 1.2 with $\mathbf{X}$ having rows i.i.d. with respect to a given distribution on $\mathbb{R}^{K}$ ). Thus, we have all the information we need to study these problems in the Bayes optimal setting. While unrealistic in practice, the advantage of analyzing such fully-specified models is that we have access to a baseline (optimal Bayesian inference) to which we can compare other inference techniques.

Bayes optimal estimation In order to make this last point clearer, let us discuss the estimation of $\mathbf{X}$. Let $\widehat{\mathbf{X}}(\mathbf{Y})$ be an estimator of $\mathbf{X}$ and let $n$ be the size of $\mathbf{X}$. We evaluate the performance of this estimator using its average loss $\mathbb{E}[\ell(\mathbf{X}, \widehat{\mathbf{X}}(\mathbf{Y}))]$ where $\ell(\cdot, \cdot) \geq 0$ is designed to measure the similarity of its two arguments and is known as a loss function. The smaller the average loss, the better the estimator. What is the best estimator of course depends on the selected loss. In this thesis, we exclusively work with the quadratic loss $\ell(\mathbf{x}, \widehat{\mathbf{x}}):=\|\mathbf{x}-\widehat{\mathbf{x}}\|^{2} / n$, where $\|\cdot\|$ is the Euclidean norm. This loss is well-suited to both continuous and discrete real variables $s^{6}$, and yields the mean-square error $\mathbb{E}\|\mathbf{X}-\widehat{\mathbf{X}}(\mathbf{Y})\|^{2} / n$. For our choice of loss function, the Bayes optimal estimator

$$
\begin{equation*}
\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]:=\int \mathbf{x} P_{\mathbf{X} \mid \mathbf{Y}}(\mathbf{x} \mid \mathbf{Y}) d \mathbf{x}=\frac{\int \mathbf{x} P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{x}) P_{\mathbf{X}}(\mathbf{x}) d \mathbf{x}}{\int P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{x}) P_{\mathbf{X}}(\mathbf{x}) d \mathbf{x}} \tag{1.8}
\end{equation*}
$$

is the best estimator; the one achieving the minimum mean-square error (MMSE),

$$
\begin{equation*}
\operatorname{MMSE}(\mathbf{X} \mid \mathbf{Y}):=\min _{\widehat{\mathbf{X}}(\cdot)} \frac{\mathbb{E}\|\mathbf{X}-\widehat{\mathbf{X}}(\mathbf{Y})\|^{2}}{n}=\frac{\mathbb{E}\|\mathbf{X}-\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]\|^{2}}{n} \tag{1.9}
\end{equation*}
$$

Indeed, a classic computation shows that

$$
\begin{aligned}
\mathbb{E}\|\mathbf{X}-\widehat{\mathbf{X}}(\mathbf{Y})\|^{2} & =\mathbb{E}\|\mathbf{X}-\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]+\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]-\widehat{\mathbf{X}}(\mathbf{Y})\|^{2} \\
& =\mathbb{E}\|\mathbf{X}-\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]\|^{2}+\mathbb{E}\|\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]-\widehat{\mathbf{X}}(\mathbf{Y})\|^{2} \geq \mathbb{E}\|\mathbf{X}-\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]\|^{2}
\end{aligned}
$$

where the second equality follows from first expanding the squared norm and then realizing that $\mathbb{E}[\langle\mathbf{X}-\mathbb{E}[\mathbf{X} \mid \mathbf{Y}], \mathbb{E}[\mathbf{X} \mid \mathbf{Y}]-\widehat{\mathbf{X}}(\mathbf{Y})\rangle]$, where $\langle\cdot, \cdot\rangle$ is the Euclidean inner product, is zero because

$$
\begin{aligned}
\mathbb{E}[\langle\mathbf{X}, \mathbb{E}[\mathbf{X} \mid \mathbf{Y}]-\widehat{\mathbf{X}}(\mathbf{Y})\rangle] & =\mathbb{E}\left[\int\langle\mathbf{x}, \mathbb{E}[\mathbf{X} \mid \mathbf{Y}]-\widehat{\mathbf{X}}(\mathbf{Y})\rangle P_{\mathbf{X} \mid \mathbf{Y}}(\mathbf{x} \mid \mathbf{Y}) d \mathbf{x}\right] \\
& =\mathbb{E}[\langle\mathbb{E}[\mathbf{X} \mid \mathbf{Y}], \mathbb{E}[\mathbf{X} \mid \mathbf{Y}]-\widehat{\mathbf{X}}(\mathbf{Y})\rangle]
\end{aligned}
$$

We see that the Bayes optimal estimator not only offers a baseline but also a lower bound on the mean-square error of any estimator and, in particular, any estimator output by an algorithm; the mean-square error of $\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]$ constitutes a limit to approach as closely as possible. In practice, we cannot compute the MMSE as we do not have access to the true prior and likelihood distributions to begin with. On the contrary, fully-specified models like the ones analyzed in this thesis make it possible to determine this universal lower bound and how far from it are the performances of existing algorithms.

[^4]MMSE in high-dimensional regimes In order to evaluate the MMSE, we have to compute first the normalization factor of the posterior distribution in (1.7) and then the posterior mean in (1.8). Both steps involve computing a $n$-dimensional integral. While we might consider evaluating such integrals on low-dimensional problems where $n$ is just a handful, the computation of these integrals suffers from the curse of dimensionality. For example, we can try to approximate these integrals by selecting evenly spaced points at which we evaluate the integrand, but the number of points needed for a good approximation grows exponentially with $n$. The difficulty of computing the Bayes estimator appears clearly in the next example that corresponds to Model 1.2 when $p=2, K=1$ and the components of $\mathbf{X}$ are independent and uniformly distributed on $\{-1,1\}$. We use it as a running example in the remaining parts of the introduction due to its simplicity as well as being a prototypical problem for which proof techniques have first been developed [18], [21], [34]-[37].

Example 1.2 (Running example). Let $\mathrm{X} \in \mathbb{R}^{n}$ be a random vector whose components are independent and uniformly distributed on $\{-1,1\}$. We observe the $n \times n$ real symmetric matrix

$$
\begin{equation*}
\mathbf{Y}:=\frac{\mathbf{X} \mathbf{X}^{\top}}{\sqrt{n}}+\sqrt{\Delta} \mathbf{Z} \tag{1.10}
\end{equation*}
$$

where $\mathbf{Z}$ is a $n \times n$ symmetric matrix whose elements $Z_{i j}$ are independent standard Gaussians for $1 \leq i \leq j \leq K$ and $\Delta>0$ is the noise variance.

The prior distribution of $\mathbf{X}$ is $P_{\mathbf{X}}(\mathbf{x})=1 / 2^{n}$ for every $\mathbf{x} \in\{-1,1\}^{n}$. The distribution of $\mathbf{Y}$ conditional on $\mathbf{X}$ is supported on the space of $n \times n$ symmetric matrices, and the corresponding conditional density function is

$$
\begin{equation*}
P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{y} \mid \mathbf{x})=\prod_{1 \leq i \leq j \leq n} \frac{1}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}\left(y_{i j}-x_{i} x_{j} / \sqrt{n}\right)^{2}} \tag{1.11}
\end{equation*}
$$

where $\mathbf{y}$ is a real symmetric matrix and $\mathbf{x} \in\{-1,1\}^{n}$. The Bayes optimal estimator reads (after simplifying the factors common to the numerator and denominator)

$$
\begin{equation*}
\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]=\frac{\sum_{\mathbf{x} \in\{-1,1\}^{n}} \mathbf{x} \prod_{1 \leq i \leq j \leq n} e^{\frac{x_{i} x_{j} Y_{i j}}{\Delta \sqrt{n}}}}{\sum_{\mathbf{x} \in\{-1,1\}^{n}} \prod_{1 \leq i \leq j \leq n} e^{\frac{x_{i} x_{j} j_{i j}}{\Delta \sqrt{n}}}} . \tag{1.12}
\end{equation*}
$$

We see that evaluating the Bayes estimator requires us to compute two sums over a set of size $2^{n}$ growing exponentially with $n$.

Let us remark that the sums in 1.12 are hard to compute because the components of $\mathbf{X}$ - despite being a priori independent - are coupled by the observations. The same remark applies to any model of Section 1.1 and 1.2. The coupling stems from the tensor product defining $\mathbf{T}$ in Models 1.1 and 1.2, and from the linear transformation ${ }^{7}$ preceding an activation function in Models 1.3

[^5]and 1.4. On the contrary, if we consider Model 1.3 with a diagonal weight matrix W then there is no coupling and we have $n$ decoupled estimation problems $Y_{i}:=\varphi\left(W_{i i} X_{i}, \mathbf{A}_{i}\right)+\sqrt{\Delta} Z_{i}$; the complexity of computing all of the Bayes optimal estimators $\mathbb{E}\left[X_{i} \mid Y_{i}\right]=\mathbb{E}\left[X_{i} \mid \mathbf{Y}\right]$ scales linearly with $n$, instead of exponentially. Also note that computing the Bayes estimator is related to sampling from the posterior distribution (after all, it is the mean of the posterior). In fact, Markov chain Monte Carlo (MCMC) methods for sampling from a distribution [38] were originally developed and used for computing multi-dimensional integrals similar to (1.8) [39], [40]. MCMC sampling is more efficient than the evenly-spaced sampling previously mentioned but is still subject to the curse of dimensionality.

Seen under this prism, evaluating the MMSE for $n$ large seems a hopeless task. Yet, over the last decade, rigorous proofs have emerged of computable exact formulas for the MMSE of different estimation problems in the high-dimensional limit $n \rightarrow+\infty$. These formulas were first conjectured using heuristics from statistical mechanics dating back to the 1970s and 80s [41], [42], see Section 1.5. Demonstrating the exactness of these predictions is a challenge that has driven - and continues to drive - a long line of research in statistical mechanics and information theory [34], [36], [43]-[47] to which this thesis belongs. Indeed, the main contribution of the present work is in proving such computable formulas for the MMSE in models of tensor decomposition and GLMs. Remarkably enough, one method of choice to prove the exactness of these formulas in the high-dimensional limit is to interpolate from the original problem - whose coupling in the observations is responsible for the curse of dimensionality - to a simpler problem where the variables to estimate are decoupled in the observations. This blessing of dimensionality [48] is due to the concentration of measure phenomenon, the mathematical background to statistical mechanics 49].

In the next section we introduce the mutual information, the central quantity in information theory, and show how we obtain the MMSE as a byproduct from computing the former. Because the mutual information is closely related to the free energy in statistical mechanics, we can borrow tools developed over decades of research on large disordered systems in order to analyze (and even compute) the mutual information in high-dimensional regimes, see Section 1.6.

### 1.4 Mutual information and estimation

The mutual information of two random variables was introduced by Claude Shannon in the paper [50] that gave birth to the field of information theory. It is understood as a measure of the "amount of shared information" between the two random variables or, equivalently, of how much information we gain about one random variable from observing the other. The mutual information was initially used to study the fundamental limits of communicating through a channel, with the two random variables being the corresponding input and output. In this thesis, making use of the notations of the previous sections, the signal $\mathbf{X}$ to estimate is the input, the observations $\mathbf{Y}$ is the output, and Models 1.1 to 1.4 are the different channels under consideration.

Let $\mathbf{X}$ and $\mathbf{Y}$ be two random variables with joint distribution $P_{\mathbf{X}, \mathbf{Y}}$ and marginal distributions $P_{\mathbf{X}}, P_{\mathbf{Y}}$. The mutual information between $\mathbf{X}$ and $\mathbf{Y}$ is

$$
\begin{equation*}
I(\mathbf{X} ; \mathbf{Y}):=\mathbb{E}\left[\ln \left(\frac{P_{\mathbf{X}, \mathbf{Y}}(\mathbf{X}, \mathbf{Y})}{P_{\mathbf{X}}(\mathbf{X}) P_{\mathbf{Y}}(\mathbf{Y})}\right)\right] \tag{1.13}
\end{equation*}
$$

The mutual information has intuitively pleasing properties: it is symmetric (shared information), nonnegative and zero if, and only if, $\mathbf{X}$ and $\mathbf{Y}$ are independent.

From the mutual information to the MMSE We can mathematically define a communication channel by a conditional distribution $P_{\mathbf{Y} \mid \mathbf{X}}$ of the ouput $\mathbf{Y}$ given the input $\mathbf{X}$. Then, the channel capacity of a communication channel $P_{\mathbf{Y} \mid \mathbf{X}}$ is the maximum of the mutual information between the input and the output when optimizing over the input distribution. The key result from Shannon's landmark paper [50] is that the channel capacity is a tight upper bound on the rate at which we can reliably transmit information through the channel. This result gives an appealing physical interpretation to the mutual information in communication problems.

By looking purely at the definition of the mutual information, it is not clear how to interpret it from an estimation point of view. The answer is given by a formula called the I-MMSE relationship [51]. It connects the derivative of the mutual information to the MMSE, where the derivative is with respect to a parameter of the model akin to a signal-to-noise ratio (SNR). Let $\mathbf{X}$ be a set of $n$ real random variables arranged in a vector and satisfying $\mathbb{E}\|\mathbf{X}\|^{2}<+\infty$. We observe $\widetilde{\mathbf{Y}}{ }^{(\tau)}:=\sqrt{\tau} \mathbf{X}+\widetilde{\mathbf{Z}}$, where $\widetilde{\mathbf{Z}}$ is a vector with independent standard Gaussian components and $\tau \geq 0$ is a known SNR. By [51, Theorem 2], we have the identity

$$
\begin{equation*}
\frac{\partial}{\partial \tau}\left(\frac{I\left(\mathbf{X} ; \widetilde{\mathbf{Y}}^{(\tau)}\right)}{n}\right)=\frac{\operatorname{MMSE}\left(\mathbf{X} \mid \tilde{\mathbf{Y}}^{(\tau)}\right)}{2} \tag{1.14}
\end{equation*}
$$

where $\operatorname{MMSE}\left(\mathbf{X} \mid \mathbf{Y}^{(\tau)}\right)$ - the MMSE for estimating $\mathbf{X}$ from $\widetilde{\mathbf{Y}}^{(\tau)}$ - is defined in (1.9). In the models of Sections 1.1 and 1.2 , the inverse $\Delta^{-1}$ of the noise variance is the SNR. For example, applied to Models 1.2 and 1.3 the I-MMSE relationship yields

$$
\begin{equation*}
\frac{\partial}{\partial \Delta^{-1}}\left(\frac{I(\mathbf{X} ; \mathbf{Y})}{n}\right)=\frac{\operatorname{MMSE}(\mathbf{T} \mid \mathbf{Y})}{2 p!} \tag{1.15}
\end{equation*}
$$

and

$$
\begin{equation*}
\frac{\partial}{\partial \Delta^{-1}}\left(\frac{I(\mathbf{X} ; \mathbf{Y})}{n}\right)=\frac{\operatorname{MMSE}(\varphi(\mathbf{w} \mathbf{x} / n) \mid \mathbf{Y})}{2} \tag{1.16}
\end{equation*}
$$

respectively. In these models, if the high-dimensional behavior of $\operatorname{MMSE}(\mathbf{X} \mid \mathbf{Y})$ is what we are really interested in then we have to consider an inference problem $\underset{\sim}{w}$ here, in addition to the original observations $\mathbf{Y}$, we have the side information $\widetilde{\mathbf{Y}}^{(\tau)}:=\sqrt{\tau} \mathbf{X}+\widetilde{\mathbf{Z}}$. At $\tau=0$, the latter observations are pure noise and the derivative of the normalized mutual information with respect to $\tau$ is directly
related to the MMSE of the original problem,

$$
\begin{equation*}
\left.\frac{\partial}{\partial \tau}\left(\frac{I\left(\mathbf{X} ; \mathbf{Y}, \tilde{\mathbf{Y}}^{(\tau)}\right)}{n}\right)\right|_{\tau=0}=\frac{\operatorname{MMSE}\left(\mathbf{X} \mid \mathbf{Y}, \tilde{\mathbf{Y}}^{(0)}\right)}{2}=\frac{\operatorname{MMSE}(\mathbf{X} \mid \mathbf{Y})}{2} \tag{1.17}
\end{equation*}
$$

All in all, thanks to the I-MMSE relationship, we can study the fundamental limits of an estimation problem by looking at the mutual information between the observations and the quantity to estimate. In particular, we can analyze the asymptotic behavior of the mutual information when $n$ diverges in order to determine the MMSE in the high-dimensional regime. It amounts to first computing the limit of the normalized mutual information, and then the derivative of the limit, to finally obtain the asymptotic MMSE thanks to (1.17). Note that, strictly speaking, (1.17) says that we should first compute the derivative of the normalized mutual information, and then the limit of the derivative as $n$ diverges. However, concavity and differentiability of the mutual information with respect to $\tau$ imply that if the normalized mutual information converges pointwise to the function $i(\tau)$ then its derivative converges to the derivative $i^{\prime}(\tau)$ at every point where the derivative exists [52, Appendix A] (see also [53, Griffiths' lemma]).

Mutual information in high-dimensional regimes Hence, we now shift our interest to the computation of the normalized mutual information. Playing with the definition of the mutual information, we find that

$$
\begin{equation*}
\frac{I(\mathbf{X} ; \mathbf{Y})}{n}=-\frac{1}{n} \mathbb{E}\left[\ln \int P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{x}) P_{\mathbf{X}}(\mathbf{x}) d \mathbf{x}\right]+\frac{1}{n} \mathbb{E}\left[\ln P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{X})\right] \tag{1.18}
\end{equation*}
$$

where we use that $P_{\mathbf{Y}}(\mathbf{y})=\int P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{y} \mid \mathbf{x}) P_{\mathbf{X}}(\mathbf{x}) d \mathbf{x}$ and $P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{y} \mid \mathbf{x})=P_{\mathbf{X}, \mathbf{Y}(\mathbf{x}, \mathbf{y}) / P_{\mathbf{X}}(\mathbf{x})}$. We can usually reduce the second expectation on the right-hand side of (1.18) to a simple form, in particular for the models we work with because the observations $\mathbf{Y}$ are - possibly stochastic - functions of $\mathbf{X}$ under an AWGN channel. More often than not, the simplification is similar to the one demonstrated below.

Example 1.2 continuing from p. 11. Plugging $\mathbf{X}$ and $\mathbf{Y}:=\frac{\mathbf{X X}^{\top}}{\sqrt{n}}+\sqrt{\Delta} \mathbf{Z}$ in the conditional density function (1.11), we obtain

$$
P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{X})=\prod_{1 \leq i \leq j \leq n} \frac{e^{-\frac{\left(Y_{i j}-x_{i} X_{j / \sqrt{n}}\right)^{2}}{2 \Delta}}}{\sqrt{2 \pi \Delta}}=\frac{1}{(2 \pi \Delta)^{\frac{n(n+1)}{4}}} \prod_{1 \leq i \leq j \leq n} e^{-\frac{z_{i j}^{2}}{2}}
$$

It directly follows that

$$
\frac{1}{n} \mathbb{E}\left[\ln P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{X})\right]=-\frac{n+1}{4} \ln (2 \pi \Delta)-\sum_{1 \leq i \leq j \leq n} \frac{\mathbb{E} Z_{i j}^{2}}{2 n}=-\frac{n+1}{4} \ln (2 \pi \Delta e)
$$

Although this quantity diverges in the high-dimensional limit, it cancels out with its opposite coming from the first expectation on the right-hand side of (1.18;

$$
\begin{aligned}
-\frac{1}{n} \mathbb{E} & {\left[\ln \int P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{x}) P_{\mathbf{X}}(\mathbf{x}) d \mathbf{x}\right]=-\frac{1}{n} \mathbb{E}\left[\ln \sum_{\mathbf{x} \in\{-1,1\}^{n}} \frac{1}{2^{n}} \prod_{i \leq j} \frac{e^{-\frac{\left(Y_{i j}-x_{i} x_{j} / \sqrt{n}\right)^{2}}{2 \Delta}}}{\sqrt{2 \pi \Delta}}\right] } \\
& =\ln (2)+\frac{n+1}{4} \ln (2 \pi \Delta)+\sum_{i \leq j} \frac{\mathbb{E} Y_{i j}^{2}+n^{-1}}{2 \Delta n}-\frac{1}{n} \mathbb{E}\left[\ln \sum_{\mathbf{x} \in\{-1,1\}^{n}} \prod_{i \leq j} e^{\frac{Y_{i j} x_{i j} x_{j}}{\Delta \sqrt{n}}}\right] \\
& =\ln (2)+\frac{n+1}{4} \ln (2 \pi e \Delta)+\frac{n+1}{2 \Delta n}-\frac{1}{n} \mathbb{E}\left[\ln \sum_{\mathbf{x} \in\{-1,1\}^{n}} e^{\sum_{i \leq j}}\right]
\end{aligned}
$$

where we use that $\mathbb{E} Z_{i j}^{2}=n^{-1}+\Delta$ in the last equality. The normalized mutual information simplifies to

$$
\begin{equation*}
\frac{I(\mathbf{X} ; \mathbf{Y})}{n}=\ln (2)+\frac{1+n^{-1}}{2 \Delta}-\frac{1}{n} \mathbb{E}\left[\ln \sum_{\mathbf{x} \in\{-1,1\}^{n}} e e^{\sum_{i \leq j \leq n} \frac{Y_{i j} x_{i j} x_{j}}{\Delta \sqrt{n}}}\right] . \tag{1.19}
\end{equation*}
$$

Therefore, computing the normalized mutual information reduces to computing the normalized expected logarithm of $\int P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{x}) P_{\mathbf{X}}(\mathbf{x}) d \mathbf{x}$. The latter integral is the normalization factor of the posterior (1.7) whose difficulty to calculate was already discussed in Section 1.3. It begs the question of whether the asymptotic of

$$
\begin{equation*}
-\frac{1}{n} \mathbb{E}\left[\ln \int P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{x}) P_{\mathbf{X}}(\mathbf{x}) d \mathbf{x}\right] \tag{1.20}
\end{equation*}
$$

is easier to analyze than for the MMSE. It turns out that 1.20 has the same form than a fundamental quantity called the average free energy in statistical mechanics, where we postulate a distribution on the possible states of a physical system and are also confronted to the task of normalizing it. Hence, it is worth looking at the rigorous and heuristic tools developed by statistical physicists to analyze such normalization factors.

### 1.5 A detour to statistical physics

Statistical physics was created at the end of the $19^{\text {th }}$ century for the purpose of providing a framework capable to explain the macroscopic properties of a physical material (properties "observable" at our scale, e.g., pressure, magnetization) from the properties and interactions of its constituents (e.g., atoms, molecules, etc.). It describes a system as an ensemble of possible states; each state corresponding to a particular microscopic configuration and being associated with a probability that the system will actually be in that state.

Spin models We focus our discussion on spin systems because the methods developed in statistical mechanics to analyze them are particurlarly relevant for
the estimation problems studied in this thesis. A spin system is described by $n$ degrees of freedom called spins and denoted by $x_{i}$ for $i=1 \ldots n$. A (Ising) spin $x_{i}$ is either 1 or -1 ; it represents the magnetic dipole moment associated to one particle in the system. The state of the system is the vector $\mathbf{x}=\left[x_{1}, x_{2}, \ldots, x_{n}\right]^{\top}$ collecting all the spins. The energy (or cost) of a specific configuration $\mathbf{x}$ is given by $\mathcal{H}(\mathbf{x})$ where the function $\mathcal{H}$ is known as the Hamiltonian. We consider spin systems that are closed, i.e., $n$ is fixed, and in equilibrium at the temperature $T$. Then, the probability to find the system in a state $\mathbf{x}$ follows the Boltzmann distribution $8^{8}$

$$
\begin{equation*}
P_{\beta}(\mathbf{x}):=\frac{e^{-\beta \mathcal{H}(\mathbf{x})}}{\mathcal{Z}(\beta)} \tag{1.21}
\end{equation*}
$$

where $\beta:=1 / T$ is an inverse temperature and $\mathcal{Z}(\beta)$ is the normalization factor of the distribution. The latter is also known as the partition function. Note that the sum in $\mathcal{Z}(\beta):=\sum_{\mathbf{x}} e^{-\beta \mathcal{H}(\mathbf{x})}$ runs over the $2^{n}$ possible states, an exponential growth alike to the one encountered in Bayes optimal estimation. The distribution (1.21) assigns higher probabilities to states of lower energies, the assignment being more or less extreme depending on the temperature. At high temperature $(\beta \rightarrow 0)$, all the states become equiprobable no matter their energies while, at low temperature $(\beta \rightarrow+\infty)$, the system is found in the states of lowest energy.

Free energy In statistical mechanics, a lot of interest is concentrated on the free energy,

$$
\begin{equation*}
f_{n}(\beta):=-\frac{1}{n \beta} \ln \mathcal{Z}(\beta) \tag{1.22}
\end{equation*}
$$

The opposite of the free energy (1.22) is sometimes dubbed free entropy in the information theory literature. This quantity is fundamental not only for its relation to the partition function, but also because macroscopic properties of the system can be determined from its knowledge. The latter is better seen on an example.

Example 1.3 (Ising model). In the canonical Ising model, the $n$ spins are located on a $d$-dimensional square grid. The notation $i \sim j$ means that the spins $x_{i}$ and $x_{j}$ are neighbors. The Hamiltonian reads

$$
\begin{equation*}
\mathcal{H}(\mathbf{x}):=-\sum_{i \sim j} J x_{i} x_{j}-\sum_{i=1}^{n} h x_{i} \tag{1.23}
\end{equation*}
$$

where the first sum runs over all pairs of neighboring spins on the grid, the variable $J$ controls the strength of the interaction between two neighbors, and $h$ is an external field. For $J>0$, states whose neighboring spins are aligned (i.e., $x_{i} x_{j}=1$ ) have lower energies. It is the opposite for $J<0$. The external

[^6]field favors configurations whose macroscopic magnetizations $M(\mathbf{x}):=\sum_{i=1}^{n} x_{i}$ have the same sign than $h$. We see that the free energy (1.22) of the Ising model satisfies
\[

$$
\begin{equation*}
\frac{\partial \beta f_{n}(\beta)}{\partial \beta}=\sum_{\mathbf{x}} \frac{\mathcal{H}(\mathbf{x})}{n} P_{\beta}(\mathbf{x}) \quad \text { and } \quad-\frac{\partial f_{n}(\beta)}{\partial h}=\sum_{\mathbf{x}} \frac{M(\mathbf{x})}{n} P_{\beta}(\mathbf{x}) . \tag{1.24}
\end{equation*}
$$

\]

Hence, we can obtain the average energy and magnetization per spin of the system from its free energy.

In Section 1.4 we have already pointed out that the mutual information is akin to a free energy. We now realize that the I-MMSE relationship (1.17) is much like the link between free energy and macroscopic properties of a system (see 1.24).

Variational formula for the free energy Statistical physics is naturally interested in systems for which the number of interacting particles is large. We are thus faced with the challenge of analyzing the free energy when $n$ diverges; a challenge similar to the one encountered in high-dimensional Bayesian estimation. Ising models are notoriously difficult to analyze (in fact, there is no analytical expression of the free energy for $d \geq 3$ ). Instead, we turn to a mean-field approximation of the Ising mode ${ }^{9}$, whose free energy in the high-dimensional limit is given by a closed-form expression and already exhibits an interesting behavior.

Example 1.4 (Curie-Weiss model). In the Curie-Weiss model, the Hamiltonian is

$$
\begin{equation*}
\mathcal{H}(\mathbf{x}):=-\sum_{i<j} \frac{J}{n} x_{i} x_{j}-\sum_{i=1}^{n} h x_{i}, \tag{1.25}
\end{equation*}
$$

where $J / n$ controls the strength of the interaction between two spins and the variable $h$ is an external field. Therefore, all the spins interact with each other and with the same strength. For this reason, the Hamiltonian can be conveniently rewritten as a function of the global magnetization per $\operatorname{spin} m(\mathbf{x}):=\frac{1}{n} \sum_{i=1}^{n} x_{i}$,

$$
\begin{equation*}
\mathcal{H}(\mathbf{x}):=-n\left(\frac{J}{2} m(\mathbf{x})^{2}+h m(\mathbf{x})\right)+\frac{J}{2} . \tag{1.26}
\end{equation*}
$$

Thus, the sum over $2^{n}$ states in the partition function can be crucially replaced by a sum over the $n+1$ possible values of the global magnetization per spin,

$$
\begin{equation*}
\mathcal{Z}(\beta)=\sum_{m \in\left\{1-2 \frac{k}{n}: 0 \leq k \leq n\right\}}\binom{n}{k} e^{n \beta\left(\frac{J m^{2}}{2}+h m\right)+\frac{\beta J}{2}}, \tag{1.27}
\end{equation*}
$$

[^7]where the binomial coefficient $\binom{n}{k}$ counts the number of states with a global magnetization $1-\frac{2 k}{n}$. By Stirling's approximation ${ }^{10}$, we have the bounds (valid for all $n \geq 2$ and $0 \leq k \leq n$ )
\[

$$
\begin{equation*}
\frac{1}{\sqrt{n}} e^{n h_{b}\left(\frac{k}{n}\right)} \leq\binom{ n}{k} \leq e^{n h_{b}\left(\frac{k}{n}\right)} \tag{1.28}
\end{equation*}
$$

\]

where we introduced the binary entropy function $h_{b}: p \mapsto-p \ln (p)-(1-p) \ln (1-p)$ for $p \in[0,1]$. We define the function

$$
\begin{equation*}
\phi(m ; \beta, J, H):=\frac{J m^{2}}{2}+h m+\beta^{-1} h_{b}\left(\frac{1-m}{2}\right) . \tag{1.29}
\end{equation*}
$$

Making use of (1.28), we find that

$$
\frac{1}{\sqrt{n}} e^{\frac{\beta J}{2}+n \beta \max _{0 \leq k \leq n} \phi\left(1-\frac{2 k}{n} ; \beta, J, h\right)} \leq \mathcal{Z}(\beta) \leq(n+1) e^{\frac{\beta J}{2}+n \beta \max _{0 \leq k \leq n} \phi\left(1-\frac{2 k}{n} ; \beta, J, h\right)}
$$

Hence, we see that the free energy $f_{n}(\beta):=-(n \beta)^{-1} \ln \mathcal{Z}(\beta)$ satisfies

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} f_{n}(\beta)=-\max _{m \in[-1,1]} \phi(m ; \beta, J, H)=\min _{m \in[-1,1]}-\phi(m ; \beta, J, H) \tag{1.30}
\end{equation*}
$$

Despite the coupling of the spins in the Hamiltonian (1.25), we have just shown that the free energy of the Curie Weiss model is given by a simple variational formula in the high-dimensional limit. Computing the limit of the free energy amounts to minimizing a scalar function $-\phi(\cdot ; \beta, J, H)$ known as a potential function. The straightforward identification of the limit and its proof were made possible by the interaction that is deterministic and identical for all pairs of spins in the system. It would be desirable to have a step-by-step recipe that we can follow in order to prove, or at least predict, variational formulas similar to 1.30 for the free energy of more complex systems. For example, the normalized mutual information (1.19) is equal - up to a simple additive term - to the average free energy at inverse temperature $\beta=1$ of a spin system whose Hamiltonian reads

$$
\begin{equation*}
\mathcal{H}(\mathbf{x}):=-\sum_{i<j} \frac{Y_{i j}}{\Delta \sqrt{n}} x_{i} x_{j} \tag{1.31}
\end{equation*}
$$

Thus, the interaction $Y_{i j} / \Delta \sqrt{n}$ between two spins is random and depends on the specific pair of spins. We need to first compute the free energy for a realization of the random interactions and then average over all possible realizations. Spin models with random interactions have also been studied in statistical mechanics and are called spin glasses ${ }^{11}$. The replica method [41], at which we look next, is a

[^8]heuristic tool developed by statistical physicists to predict the average free energy of such disordered systems in the high-dimensional limit.

Replica method Consider a system of $n$ spins with an Hamiltonian $\mathcal{H}(\mathbf{x} ; \mathbf{J})$ that depends explicitly on random variables $\mathbf{J}$ controlling the interactions between spins. We denote by $f_{n}(\beta)$ the average free energy, that is, the free energy of the system averaged over all possible realizations of $\mathbf{J}$;

$$
\begin{equation*}
f_{n}(\beta):=-\frac{1}{n \beta} \mathbb{E}[\ln \mathcal{Z}(\beta, \mathbf{J})], \tag{1.32}
\end{equation*}
$$

where $\mathcal{Z}(\beta, \mathbf{J}):=\sum_{\mathbf{x}} e^{-\beta \mathcal{H}(\mathbf{x} ; \mathbf{J})}$. The replica method starts with the following observation,

$$
\begin{aligned}
\lim _{a \rightarrow 0^{+}} \frac{\ln \left(\mathbb{E}\left[\mathcal{Z}(\beta, \mathbf{J})^{a}\right]\right)}{a} & =\lim _{a \rightarrow 0^{+}} \frac{\partial \ln \left(\mathbb{E}\left[\mathcal{Z}(\beta, \mathbf{J})^{a}\right]\right)}{\partial a} \\
& =\lim _{a \rightarrow 0^{+}} \frac{\mathbb{E}\left[\mathcal{Z}(\beta, \mathbf{J})^{a} \ln (\mathcal{Z}(\beta, \mathbf{J}))\right]}{\mathbb{E}\left[\mathcal{Z}(\beta, \mathbf{J})^{a}\right]}=\mathbb{E}[\ln \mathcal{Z}(\beta, \mathbf{J})]
\end{aligned}
$$

Hence, the average free energy (1.32) satisfies

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} f_{n}(\beta)=\lim _{n \rightarrow+\infty} \lim _{a \rightarrow 0^{+}}-\frac{1}{a n \beta} \ln \left(\mathbb{E}\left[\mathcal{Z}(\beta, \mathbf{J})^{a}\right]\right) \tag{1.33}
\end{equation*}
$$

Then, the replica method takes a jump away from mathematical rigor by inverting the two limits in (1.33),

$$
\lim _{n \rightarrow+\infty} f_{n}(\beta) \stackrel{!?}{=} \lim _{a \rightarrow 0^{+}} \lim _{n \rightarrow+\infty}-\frac{1}{a n \beta} \ln \left(\mathbb{E}\left[\mathcal{Z}(\beta, \mathbf{J})^{a}\right]\right)
$$

Next, we assume that $a$ is a positive integer. The power $\mathcal{Z}(\beta, \mathbf{J})^{a}$ is the partition function of $a$ independent replicated systems,

$$
\begin{equation*}
\mathcal{Z}(\beta, \mathbf{J})^{a}=\sum_{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(a)}} e^{-\beta \sum_{r=1}^{a} \mathcal{H}\left(\mathbf{x}^{(r)} ; \mathbf{J}\right)}, \tag{1.34}
\end{equation*}
$$

where $\mathbf{x}^{(r)}$ is the state of the $r^{\text {th }}$ replica of the system. The previous manipulations of the replica method help us to pass the logarithm outside of the expectation. Instead, we have to compute the high-dimensional limit of the annealed free energy $-\ln \mathbb{E}\left[\mathcal{Z}(\beta, \mathbf{J})^{a}\right] / a n \beta$ of the replicated system. Similarly to what is done in Example 1.4 we have to determine which states $\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(a)}\right)$ of the replicated system contribute the most to the partition function (1.34). Doing so is difficult and, at this point, some assumptions have to be made to simplify the analysis. A common assumption is the replica symmetric (RS) assumption ${ }^{122}$. Under these

[^9]assumptions, we obtain a formula for $\lim _{n \rightarrow+\infty}-\ln \mathbb{E}\left[\mathcal{Z}(\beta, \mathbf{J})^{a}\right] /$ an $\beta$ when $a$ is a positive integer. Finally, and without any form of mathematical justification, the replica method extends the formula obtained for positive integers to the positive real numbers and then computes the limit when $a$ vanishes. In all generality, we end up with a variational formula of the form
$$
\lim _{n \rightarrow+\infty} f_{n}(\beta) \stackrel{!?}{=} \underset{\mathbf{m} \in \mathcal{M}}{\operatorname{extr}} f_{\mathrm{RS}}(\mathbf{m})
$$
where extr is a combination of supremums and infimums, and $\mathbf{m}$ are parameters constrained to the set $\mathcal{M}$ over which the function $f_{\mathrm{RS}}(\mathbf{m})$ is extremized. The function $f_{\mathrm{RS}}$ is called the replica symmetric potential.

No one denies the divinatory nature of the replica method ${ }^{13}$. It is nevertheless a powerful tool to guess a formula that can later be proved or disproved by other means. In this regard, the Sherrington-Kirkpatrick (SK) model is an excellent example.

Example 1.5 (SK model). The Hamiltonian of the SK model in the absence of external field is given by

$$
\begin{equation*}
\mathcal{H}(\mathbf{x} ; \mathbf{J}):=-\sum_{1 \leq i<j \leq n} \frac{J_{i j}}{\sqrt{n}} x_{i} x_{j} \tag{1.35}
\end{equation*}
$$

where $\mathbf{J}:=\left\{J_{i j}\right\}_{1 \leq i<j \leq n}$ are independent centered Gaussian random variables with variance $J^{2}$. This spin glass model was introduced by D. Sherrington and S. Kirkpatrick in 1975 [55]. In this latter reference, they applied the replica method under the RS assumption and proposed the replica symmetric ansatz

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \beta f_{n}(\beta) \stackrel{!?}{=}-\min _{q \in[0,1]} \frac{(\beta J)^{2}}{4}(1-q)^{2}+\ln 2+\mathbb{E}[\ln (\cosh (\beta J \sqrt{q} Z))] \tag{1.36}
\end{equation*}
$$

where $Z$ is a standard Gaussian variable. They observed that (1.36) yields an "unphysical behavior" [55] at low temperature ${ }^{14}(\beta \rightarrow+\infty)$. It turns out that the replica symmetric ansatz is a correct prediction of the average free energy only when $\beta \leq \frac{1}{J}$, in which case the minimum in (1.36) is always achieved at $q=0$. Five years later, G. Parisi proposed another ansatz [56], [57] based on the replica method but under a replica symmetry breaking (RSB) assumption less stringent than the RS one. The formula is given in [58, Chapter 3] and was proved to be correct by M. Talagrand (59.

In the next section, we look at replica symmetric ansätze for the mutual information. Before concluding our detour, let us stress that the contribution of statistical physics to Bayesian estimation is not limited to the mathematically dubious replica trick. In order to prove the exactness of the RS ansätze for the

[^10]mutual information, we will have to analyze the concentration of complicated random quantities. Therefore, the expertise of statistical mechanics on the concentration of measure phenomenon will come in handy. Finally, a key ingredient in our proofs of the RS formulas is a refined version of the interpolation method [60], [61] created by F. Guerra and F. Toninelli to study the SK model (in fact, it plays a crucial part in the proof of the RSB formula).

### 1.6 Replica symmetric formula for the mutual information

Since the relationship between statistical physics of spins and coding theory was first pointed out [62], the replica formula has gained popularity within the information theory community. It has been applied to coding problems [63][65] and more recently to estimation ones. Given that the normalized mutual information (1.18) is similar to a free energy (up to a simple additive term), the replica trick is a method of choice to predict the high-dimensional behavior of the mutual information and the MMSE. In order to illustrate the latter, let us come back to Example 1.2.

Example 1.2 continuing from p. 14). The replica symmetric ansatz predicts

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \mathbf{Y})}{n}=\min _{q \in[0,1]} f_{\mathrm{RS}}(q, \Delta) \tag{1.37}
\end{equation*}
$$

where $f_{\mathrm{RS}}(q, \Delta)$ is the RS potential function (in what follows $Z$ is a standard Gaussian random variable),

$$
\begin{equation*}
f_{\mathrm{RS}}(q, \Delta):=\frac{(1+q)^{2}}{4 \Delta}-\mathbb{E}\left[\ln \left(\cosh \left(\frac{q}{\Delta}+\sqrt{\frac{q}{\Delta}} Z\right)\right)\right] . \tag{1.38}
\end{equation*}
$$

The RS ansatz (1.37) was proved to be exact in (66]. Remember that, in this example, the components of $\mathbf{X}$ are independent and uniformly distributed on $\{-1,1\}$. The RS prediction for a more general random vector $\mathbf{X}$ with i.i.d. components was derived later in [67, Equation (30) with $M=Q]$ and proved correct in [34], [36], [37]. We review the tools developed to prove replica formulas in the next section. For now, we show how to finally compute the asymptotic MMSE after we have ascertained the RS ansatz. The global minimum in (1.37) is achieved by a solution to the stationary point equation

$$
\begin{equation*}
\left.\frac{\partial f_{\mathrm{RS}}}{\partial q}\right|_{q, \Delta}=0 \Leftrightarrow q=1-\mathbb{E}\left[\left(1-\tanh \left(\frac{q}{\Delta}+\sqrt{\frac{q}{\Delta}} Z\right)\right)^{2}\right] . \tag{1.39}
\end{equation*}
$$

Of course, for a general differentiable function on a closed interval, the global minimum is not necessarily a stationary point if it is one of the endpoint. This is not a problem here because the endpoint $q=0$ is always a solution to (1.39) and the opposite endpoint $q=1$ cannot be the global minimum (the right-derivative at $q=1$


Figure 1.3: Left: Asymptotic normalized mutual information and $\operatorname{MMSE}\left(\mathbf{X X}^{\top} \mid \mathbf{Y}\right)$ as a function of the inverse noise variance $\Delta^{-1}$ for the estimation problem of Example 1.2 The dotted vertical line marks the phase transition at $\Delta_{c}=1$; once passed this threshold the MMSE starts decreasing from its maximum value. Right: Offset RS potential function $q \in[0,1] \mapsto f_{\mathrm{RS}}(q, \Delta)-f_{\mathrm{RS}}(0, \Delta)$ for different values of the noise variance before and after the phase transition. The global minimum of a potential is indicated with a circle.
is positive). If $f_{\mathrm{RS}}(\cdot, \widetilde{\Delta})$ has a unique global minimum $q^{*}(\widetilde{\Delta})$ for all $\widetilde{\Delta}$ in a segment surrounding $\Delta$ then we can show ${ }^{15}$ that the derivative of $\widetilde{\Delta}^{-1} \mapsto \min _{q \in[0,1]} f_{\mathrm{RS}}(q, \widetilde{\Delta})$ at $\widetilde{\Delta}^{-1}=\Delta^{-1}$ exists and is equal to ${ }^{\left(1-q^{*}(\Delta)^{2}\right) / 4 \text {. Then, the I-MMSE relationship }}$ (1.15) directly implies

$$
\begin{align*}
\lim _{n \rightarrow+\infty} \operatorname{MMSE}\left(\mathbf{X X}^{\top} \mid \mathbf{Y}\right) & =\lim _{n \rightarrow+\infty} 4 \frac{\partial}{\partial \Delta^{-1}}\left(\frac{I(\mathbf{X} ; \mathbf{Y})}{n}\right) \\
& =4 \frac{\partial}{\partial \Delta^{-1}}\left(\lim _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \mathbf{Y})}{n}\right)=1-q^{*}(\Delta)^{2} \tag{1.40}
\end{align*}
$$

where $\operatorname{MMSE}\left(\mathbf{X X}^{\top} \mid \mathbf{Y}\right):=\mathbb{E}\left\|\mathbf{X} \mathbf{X}^{\top}-\mathbb{E}\left[\mathbf{X X}^{\top} \mid \mathbf{Y}\right]\right\|^{2} / n^{2}$. The inversion of the derivative and the limit in the second equality is mathematically justified as long as the pointwise limit 1.37 ) is differentiable at $\Delta^{-1}$ (see the concluding remark of From the mutual information to the MMSE in Section (1.4).

We can thus determine the asymptotic MMSE by solving the minimization problem of the RS formula. Practically, we numerically solve (1.39) with a fixed point iterative scheme started from different initializations. We keep the fixed point $q^{*}$ minimizing the RS potential to then compute the asymptotic MMSE thanks to (1.40). The resulting asymptotic normalized mutual information and MMSE are drawn in Figure 1.3. Interestingly, the estimation problem exhibits a phase transition at $\Delta_{c}=1$ in the high-dimensional limit. When the noise variance is above this threshold $\Delta_{c}$, the asymptotic MMSE is equal to $\lim _{n \rightarrow+\infty} \mathbb{E}\left\|\mathbf{X} \mathbf{X}^{\top}-\mathbb{E}\left[\mathbf{X X}^{\top}\right]\right\|^{2} / n^{2}=1$; the observations $\mathbf{Y}$ are too noisy to help recovering $\mathbf{X}$ and the best we can do is a random guess of $\mathbf{X}$ based on the prior $P_{\mathbf{X}}$. When the noise variance passes below $\Delta_{c}$, the MMSE starts decreasing continuously towards zero; it becomes statistically possible ${ }^{16}$ to do better than a random guess thanks to the observations.

[^11]Therefore, we can derive without much effort the exact performance of the Bayes optimal estimator from the replica symmetric ansatz assuming it is correct. Since the proofs of (1.37) in [34], [46], RS formulas for more general models of matrix and tensor estimation have been proved [45], [46], [69]-[72]. The validity of the RS predictions is not limited to tensor estimation; exact formulas have also been proved for linear estimation [44], [73], [74] and generalized linear estimation [29], [75]. These examples suggest that the replica symmetric assumption is generally valid in the setting of Bayes optimal inference, see for example the discussion in [76, 2.7. No RSB in Bayes-optimal inference]. Before presenting the main ideas behind the proofs of RS formulas, and to conclude this section, let us discuss how the performances of algorithms compare with the statistical limit that is the MMSE.

Algorithmic limits The MMSE is by definition a lower bound on the error of any estimator. It is the best performance that we can achieve in an ideal situation devoid of any time and space constraints where the data generating process is known exactly. A RS formula for the mutual information thus offers a great insight into the intrinsic difficulty of the estimation problem that we study, e.g., it reveals the existence of phase transitions that separate regimes in which it is possible or not to extract information on the estimated signal from the observations. Algorithms designed for estimation have to compose with both time and space constraints (e.g., their time complexity should scale polynomially with the size of the problem) as well as an imperfectly known data generating process. Determining how far algorithms stand from the MMSE in term of performance is crucial to improve upon them. In this regard, there is a class of iterative algorithms called approximate message passing (AMP) whose performances can be analyzed thanks to the RS formulas.

AMP was originally proposed for compressed sensing [77]. It is an approximation of loopy belief propagation (BP) [78] for estimation problems whose graphical representations are densily connected ${ }^{17}$. The time complexity of AMP scales polynomially with the size of the problem - the dimension $n$ in all of Models 1.1 to $1.4-$, instead of exponentially for loopy BP. AMP algorithms try to iteratively approach the MAP estimator $\arg \max _{\mathbf{x}} B_{\mathbf{X} \mid \mathbf{Y}}(\mathbf{x} \mid \mathbf{Y})$ or the Bayes estimator $\mathbb{E}_{\mathbf{x} \sim B_{\mathbf{X} \mid \mathbf{Y}}(\cdot \mid \mathbf{Y})}[\mathbf{x}]$, where $B_{\mathbf{X} \mid \mathbf{Y}}$ is the posterior belief defined in (1.6) and based on postulated prior and likelihood distributions. In the high-dimensional limit, the performance of AMP is precisely tracked by a set of equations called state evolution (SE) [79].

Let us focus on the situation where the postulated distributions match their true counterparts and our AMP algorithm tries to approach the Bayes optimal estimator $\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]$. In this case, SE corresponds to the fixed point iterative scheme derived from the stationary point equation of the RS potential function. Remember that in Example 1.2 we used the same fixed point iteration to look for

[^12]the global minimum of the RS potential function and computed the asymptotic MMSE thanks to this global minimum. Similarly, the mean-square error of the AMP estimate depends on the fixed point reached by SE. If the fixed point is the global minimum then our estimate achieves the MMSE, if not then there is a discrepancy between what we algorithmically achieve with AMP and what is statistically achievable. We again illustrate the latter on Example 1.2.

Example 1.2 continuing from p.21. The AMP algorithm of this estimation problem was obtained in [80]. We initialize the algorithm with $\widehat{\mathbf{x}}_{0}:=\mathbb{E}[\mathbf{X}]$ and we denote $\widehat{\mathbf{x}}_{k} \in \mathbb{R}^{n}$ the estimate of $\mathbf{X}$ after $k$ iterations of AMP. Remember that the entries of $\mathbf{X}$ are independent uniformly distributed on $\{-1,1\}$ and $P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid \mathbf{x})=P_{\mathbf{Y} \mid \mathbf{X}}(\mathbf{Y} \mid-\mathbf{x})$. Hence, the best we can do in this problem is to recover $\mathbf{X}$ up to a plus or minus sign. We thus study the evolution of the mean-square error of $\widehat{\mathbf{x}}_{k} \widehat{\mathbf{x}}_{k}^{\top}$,

$$
\begin{equation*}
\frac{\left\|\mathbf{X} \mathbf{X}^{\top}-\widehat{\mathbf{x}}_{k} \widehat{\mathbf{x}}_{k}^{\top}\right\|^{2}}{n^{2}}=1+\frac{\left\|\widehat{\mathbf{x}}_{k}\right\|^{4}}{n^{2}}-2 \frac{\left|\widehat{\mathbf{x}}_{k}^{\top} \mathbf{X}\right|^{2}}{n^{2}} \tag{1.41}
\end{equation*}
$$

As stated above, we suppose that our AMP algorithm tries to approach the Bayes estimator where the true prior and likelihood distributions are known. In this situation, both $\left\|\widehat{\mathbf{x}}_{k}\right\|^{2} / n$ and $\left|\widehat{\mathbf{x}}_{k}^{\top} \mathbf{x}\right| / n$ converge in the high-dimensional limit to the same scalar $q_{k} \in[0,1]$ that follows the state evolution

$$
\begin{equation*}
q_{k+1}=\operatorname{SE}\left(q_{k}\right):=\left.\frac{\partial f_{\mathrm{RS}}}{\partial q}\right|_{q_{k}, \Delta} \tag{1.42}
\end{equation*}
$$

where $f_{\mathrm{RS}}$ is the RS potential function (1.38). Therefore, the mean-square error (1.41) of the $k^{\text {th }}$ iterate satisfies $\lim _{k \rightarrow+\infty} \lim _{n \rightarrow+\infty} \frac{\left\|\mathbf{X X}^{\top}-\widehat{\mathbf{x}}_{2} \widehat{\mathbf{x}}_{k}^{\top}\right\|^{2}}{n^{2}}=1-q^{2}$ where $q$ is the fixed point reached by the state evolution (1.42).

When the prior is symmetric, like in the present example, $q_{0}=0$ and SE predicts that AMP is stuck in this uninformative fixed point; the mean-square error is as bad as it can get. However, in practice, we draw a nonzero random vector such as $\widehat{\mathbf{x}}_{0} \sim \mathcal{N}\left(0, \sigma_{0}^{2} I\right)$ and $n$ is finite. In that case, $\left|\widehat{\mathbf{x}}_{0}^{\top} \mathbf{x}\right| / n=O\left(\sigma_{0} / \sqrt{n}\right)$ and $\left\|\widehat{\mathrm{x}}_{0}\right\|^{2} / n=O\left(\sigma_{0}^{2} / \sqrt{n}\right)$ are not exactly zero but close to it as long as $\sigma_{0}$ is negligible compared to $\sqrt{n}$ (we run into convergence issues otherwise). Then, despite $n$ being finite, empirical findings demonstrate that we can still rely on SE initialized with a very small $q_{0}$ in order to predict the performance of AMP. Recent theoretical results also go in that direction [81], [82]. We saw in Figure 1.3 that $q=0$ is the unique global minimum of the RS potential when $\Delta^{-1}<1$ and that the asymptotic MMSE starts decreasing when $\Delta^{-1}>1$ as another solution to 1.39 with a lower potential appears. It turns out that $\left|\partial^{2} f_{\mathrm{RS}} / \partial q^{2}\right|_{q=0, \Delta} \mid$ is equal to $\Delta^{-1}$, that is, $q=0$ is a stable fixed point for $\Delta^{-1}<1$ and is unstable otherwise. Hence, in this particular example, state evolution initialized arbitrarily close to zero always converges to the global minimum of the potential; the mean-square error of the AMP estimate matches the MMSE curve of Figure 1.3 .

The performance of AMP does not always match the MMSE. In Figure 1.4 we draw the asymptotic MMSE and MSE of AMP for the same estimation problem


Figure 1.4: Model 1.2 for $p=2, K=1$ and $\mathbf{X} \in\{-1,0,1\}^{n}$ such that $X_{1}, X_{2} \ldots, X_{n} \stackrel{\text { i.i.d. }}{\sim}$ $(1-\rho) \delta_{0}+\frac{\rho}{2}\left(\delta_{1}+\delta_{-1}\right)$ where $\rho=0.05$. Left: Asymptotic MMSE (solid blue line) and MSE of AMP in the high-dimensional limit (dashed orange line, obtained with SE initialized at $q / \rho=10^{-6}$ ) as functions of $\rho^{2} \Delta^{-1}$. Right: Offset RS potential function $q \in[0, \rho] \mapsto f_{\mathrm{RS}}(q, \Delta)-f_{\mathrm{RS}}(0, \Delta)$ for $\Delta^{-1}$ below the information theoretic threshold (top curve), above it but below the algorithmic threshold (middle curve), and above the algorithmic threshold (bottom curve). Circles indicate global minima and triangles fixed points reached by SE. Note that circle and triangle differ in location on the middle curve; hence the computational-to-statistical gap in this region of $\Delta^{-1}$.
than Example 1.2 except that the entries of $\mathbf{X} \in\{-1,0,1\}^{n}$ are i.i.d. with respect to $(1-\rho) \delta_{0}+\frac{\rho}{2}\left(\delta_{1}+\delta_{-1}\right)$ where $\rho=0.05$. We notice a computational-to-statistical gap; there is an information theoretic threshold for $\Delta^{-1}$ above which the MMSE is nontrivia ${ }^{18}$, and an higher algorithmic thereshold above which the MSE of AMP becomes nontrivial too. Between the two thresholds, AMP does not output an estimate better than a random guess even if it is statistically possible.

To conclude, let us mention that there exist AMP algorithms for Models 1.1 and 1.2 in tensor estimation [67], [80], [83], [84] as well as for Models 1.3 and 1.4 in generalized linear estimation [28], [85]. The RS formulas that we prove for the asymptotic normalized mutual information of these different models shed light on the performance of these algorithms.

### 1.7 Proving the replica predictions

Thanks to the replica method, we can predict formulas for the normalized mutual information and MMSE associated with high-dimensional estimation problems. We are left with proving that these formulas are indeed correct. There is little hope to obtain a proof by making every step of the replica method mathematically rigorous. The most famous replica formula is the one conjectured by Parisi for the SK model of Example 1.5. The two important tools that led to a proof of the Parisi formula [58], [59] are the Guerra-Toninelli interpolation method [60], [61] and the Aizenman-Sims-Starr scheme [86].

These tools are not specific to the SK model and they have now been applied to prove RS formulas for a variety of estimation problems. Proofs go by first demonstrating that the RS formula is an upper bound on the asymptotic normalized mutual information, and then that it is a lower bound as well. The

[^13]interpolation method is the way to go in order to establish the upper bound. Proving the converse bound is in general more difficult. With the exception of the proof of 1.37] in 66] that is based on a complex interpolation scheme specific to the case $\mathbf{X} \in\{-1,1\}^{n}$, it has originally been dealt with thanks to the Aizenman-Sims-Starr scheme [36], [45]-[47] or spatial coupling [34], [44]. The latter was primarily developed to construct capacity-achieving error-correcting codes and later turned into a proof technique 43]. Recently, the interpolation method has been improved upon [37], [87] so that it can handle the lower bound, too. This adaptive interpolation method also allows the analysis of broader models, such as the generalized linear ones [29], that are out of reach of the canonical interpolation method. Altogether, these improvements offer a unified strategy to prove RS formulas.

Interpolation method We want to determine what is the normalized mutual information between a signal $\mathbf{X}$ of interest and some observations $\mathbf{Y}$ in the highdimensional limit. The difficulty of computing the mutual information stems from the coupling of the signal entries in the observations. The interpolation method circumvents this problem by analyzing not the original mutual information, but one whose observations $\mathbf{Y}^{(t)}$ depend continuously on a parameter $t$. This parameter varies from zero to one in such a way that at $t=0$ we recover the original observations and mutual information, while at $t=1$ the signal entries are not coupled in the observations. Hence, the mutual information at $t=1$ has a simple form that is easy to compute even in the high-dimensional limit. The choice of the observations at $t=1$ is not random but is guided by the RS formula itself. Indeed, part of the RS potential function is equal to the normalized mutual information between $\mathbf{X}$ and observations in which the entries of $\mathbf{X}$ are not coupled. We thus design the interpolation to mimick these observations when $t$ equals one.

In order to compare the original mutual information to the simple one, we have to understand how the mutual information evolves when $t$ increases from zero to one. To do so, we compute its derivative with respect to $t$ and write it as a sum of two terms. One term matches the part of the RS potential function that is not explained by the mutual information at $t=1$, the other is a remainder that hopefully has nice properties that we can exploit. The way that the observations $\mathbf{Y}^{(t)}$ depend on $t$ is through functions of $t$ akin to signal-to-noise ratios and called the interpolation path. In the canonical interpolation method, these are linear functions of $t$. Under this choice, the remainder is nonnegative for all $t$ and we obtain an upper bound on the original mutual information. Instead, in the adaptive interpolation method, the interpolation path is the solution to an ordinary differential equation (ODE) in $t$. The specific form of the ODE is determined a posteriori. Based on the derivative of the mutual information for a general interpolation path, we identify what should be the ODE in order to have a "well-behaved" remainder. Typically, the ODE is such that the remainder has a constant sign or vanishes completely in the high-dimensional limit. All in all, the adaptive interpolation method gives us greater leeway in the choice of the interpolation path. Let us show what the interpolation method looks like on our
running example.
Example 1.2 (continuing from p.24). The RS potential function (1.38) can be rewritten as

$$
f_{\mathrm{RS}}(q, \Delta):=\frac{(1-q)^{2}}{4 \Delta}+I(X ; \sqrt{q} X+\sqrt{\Delta} \widetilde{Z})
$$

where $X$ is uniformly distributed on $\{-1,1\}$ and independent of the standard Gaussian random variable $\widetilde{Z}$. Note that $I(X ; \sqrt{q} X+\sqrt{\Delta} \widetilde{Z})$ is equal to the normalized mutual information between $\mathbf{X}$ and $\widetilde{\mathbf{Y}}:=\sqrt{q} \mathbf{X}+\sqrt{\Delta} \widetilde{\mathbf{Z}}$ where the standard Gaussian random vector $\widetilde{\mathbf{Z}}$ is independent of $\mathbf{X}$. For this reason, we can prove the upper bound $\lim _{n \rightarrow+\infty} I(\mathbf{X} ; \mathbf{Y}) / n \leq \min _{q \in[0,1]} f_{\mathrm{RS}}(q, \Delta)$ by considering the observations

$$
\left\{\begin{array}{l}
\mathbf{Y}^{(t)}:=\sqrt{1-t} \mathbf{X X}^{\boldsymbol{\top}}+\sqrt{\Delta} \mathbf{Z} \\
\widetilde{\mathbf{Y}}^{(t)}:=\sqrt{q t} \mathbf{X}+\sqrt{\Delta} \widetilde{\mathbf{Z}}
\end{array}\right.
$$

where the interpolation path $t \in[0,1] \mapsto(1-t, q t)$ is linear. We see that the normalized mutual information $I\left(\mathbf{X} ;\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right)\right) / n$ equals $I(\mathbf{X} ; \mathbf{Y}) / n$ at $t=0$ and $I(X ; \sqrt{q} X+\sqrt{\Delta} \widetilde{Z})$ at $t=1$. Hence,

$$
\frac{I(\mathbf{X} ; \mathbf{Y})}{n}=I(X ; \sqrt{q} X+\sqrt{\Delta} \widetilde{Z})-\int_{0}^{1} \frac{d I\left(\mathbf{X} ;\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right)\right)}{n d t} d t
$$

The upper bound follows from an analysis of the derivative of the normalized mutual information. Finally, one way to prove the converse lower bound is via an adaptive interpolation [37].

### 1.8 Organization and main contributions

This thesis focuses on proving replica symmetric formulas for the normalized mutual information associated with high-dimensional statistical models. One of the main by-product of these formulas is the minimum mean-square error of the estimation task for each of these models. In this thesis, all the proofs of RS formulas are based on the adaptive interpolation method. They demonstrate its versatility by extending its applicability to more general settings. The thesis is broadly divided into two parts; one is dedicated to statistical models for tensor estimation, the other to generalized linear models. Ahead of these two parts, we present in Chapter 2 technical tools that are essential to our proofs. We refer to and heavily use them in all the remaining chapters.

In Chapter 3 we look at what is maybe the simplest instance of Model 1.1 , the estimation of a rank-one matrix $\mathbf{U V}^{\top} / \sqrt{n}$ whose entries are observed under an additive white Gaussian noise channel. There exist multiple proofs of the RS formula when each of $\mathbf{U}$ and $\mathbf{V}$ has independent and identically distributed entries [46], [69]. Here we consider a situation where the entries of these vectors are correlated as both $\mathbf{U}$ and $\mathbf{V}$ are constrained on spheres of radii $O(\sqrt{n})$ and we carry out the proof of a RS formula. The formula turns out to be identical to the one we obtain when the entries of $\mathbf{U}$ and $\mathbf{V}$ are independent standard Gaussian
random variables. The result is perhaps unsurprising given that a high-dimensional standard Gaussian random vector is thinly concentrated around the sphere of radius $\sqrt{n}$, though it should be noted that models of spherical and Gaussian spin glasses (where the system is described by a spherical or standard Gaussian random vector instead of $\pm 1$ spins) are not equivalent [88|. The important take out of this chapter is that the adaptive interpolation can handle non-i.i.d. prior on the estimated signals. It also serves as an introduction to the adaptive interpolation method; it lays down on a simple example the main steps of any proof based on it.

In Chapter 4 we go beyond rank-one tensor estimation and study Model 1.2 when the rank $K$ of the tensor buried in noise remains fixed as $n$ diverges. We prove that the RS formula conjectured in [45 is exact for even-order tensors. Prior to this work, the formula had been proved in the matrix case (order-2 tensors) using Guerra's interpolation technique and the Aizenman-Sims-Starr scheme [36]. Our proof leverages solely on the adaptive interpolation method and is not a trivial extension of the rank-one case. Ultimately, the proof shows that the method can handle estimation problems whose overlap order parameter is not a scalar but a $K \times K$ matrix.

In Chapter 5 we study the impact of a structured spike $\mathbf{X} \in \mathbb{R}^{n}$ on symmetric rank-one tensor estimation. The structure does not come from restricting $\mathbf{X}$ to a sphere like in Chapter 3. Instead, $\mathbf{X}$ is generated by a latent vector input to a generalized linear model (think of Model 1.3 with $\Delta=0$ ). We obtain a RS formula that allows us to analyze how the structure of data can be exploited in tensor estimation. For example, if the size of the latent vector is small compared to the size of $\mathbf{X}$ then the high-dimensional spike lies on a lower-dimensional manifold, a kind of structure often exhibited by natural data ${ }^{19}$, and the tensor can be estimated in higher noise regimes than when the spike has i.i.d. components. The proof demonstrates the modularity of the adaptive interpolation. Indeed, it combines an interpolation scheme similar to the one used in the i.i.d. case [37] with the RS formula associated with Model 1.3 that was proved by adaptive interpolation in 87. This chapter constitutes a natural bridge between this part on tensor estimation and the next one on generalized linear models.

In Chapter 6 we prove the RS formula associated with a two-layer generalized linear model (see Model (1.4) and conjectured in [27]. Once again, the proof shows the modularity of the adaptive interpolation since it interpolates from the two-layer GLM to two independent channels, a one-layer GLM and a random vector with i.i.d. entries observed under AWGN. We can thus rely on [29] to compute the asymptotic normalized mutual information at $t=1$. In fact, the result suggests a proof by induction of the RS formula associated with a $L$-layer GLM, where we interpolate from the $L$-layer GLM to two independent channels, one of them being a GLM with $L-1$ layers.

In Chapter 7 we study Model 1.3 in a high-dimensional regime that is not the one usually looked at in statistical physics or high-dimensional statistics. Barbier

[^14]et al. analyzed this model in the classical thermodynamic regime where the size $n$ of the signal $\mathbf{X}$ diverges while its entries are independent and identically distributed with respect to a fixed prior distribution [29]. In particular, each entry of the signal can be nonzero with a fixed probability $\rho$, yielding a sparse signal $\mathbf{X}$ that has $\rho \cdot n$ nonzero entries in expectation. In that case, the normalized mutual information and MMSE associated with the model are nontrivial, with the appearance of phase transitions, when the number $m$ of measurements scales linearly with $n$, that is, the sampling rate $m / n$ converges to a positive value $\alpha$. In this high-dimensional regime of linear sparsity and sampling rate, [29] proved a RS formula for the normalized mutual information. Instead, in Chapter 7, we consider a signal whose entries are still i.i.d. but they are nonzero with a probability $\rho_{n}$ that vanishes in the high-dimensional limit. The sparsity $\rho_{n} \cdot n$ of the signal is thus sublinear in $n$. We show that the normalized mutual information and MMSE associated with the model are nontrivial if the number $m$ of measurements also scales sublinearly with $n$. More precisely, we establish that $m / n$ should scale like $\rho_{n}\left|\ln \rho_{n}\right|$ and prove a RS formula in this high-dimensional regime of sublinear sparsity and samping rate. Our proof extends the adaptive interpolation outside of the regime for which it was initially developed. We then use the RS formula to show that the MMSE, as a function of the sampling rate, differs both qualitatively and quantitatively from the one in the classical high-dimensional regime of [29|. We demonstrate that the MMSE exhibits sharp phase transitions separating regions where it is constant. Our results thus generalize the all-or-nothing phenomenon that was evidenced in [32] and proved in [33] for the linear model with sparse binary signal $\mathbf{X}$.

This thesis is concluded by Chapter 8 where we summarize its findings and present possible research directions.

Reading guide Chapter 2 lists different tools used in the proofs of all the following chapters. The reader can skim through or skip this chapter, and come back to it when needed.

All the Chapters 3 to 7 can be read independently. Chapter 3 is a good introduction to the adaptive interpolation as we study a model that is relatively simpler compared to the ones in subsequent chapters. The main contribution of Chapter 4 is the extension of the adaptive interpolation to estimation problems where the overlap order parameter is not a scalar but a matrix. Even so the chapters are independent, it is probably a good advice to the reader not to make Chapter 4 the first one they read because of its technicality.

Both Chapters 5 and 6 demonstrate the modularity of the adaptive interpolation, but the proof of the RS formula in Chapter 5 is simpler. Besides, Chapter 5 has a discussion of the phenomenology of phase transitions accompanied by figures, making it more appealing than Chapter 6 if we have to pick between the two.

Finally, in Chapter 7, we study a high-dimensional regime of a different nature compared to the previous chapters. The adaptive interpolation is only the first step and the chapter presents a novel technique, specific to this regime, in order to simplify the RS formula into a discrete minimization problem. The phenomenology of phase transitions changes, which we discuss and illustrate with figures.

Bibliographic notes The content of Chapter 3 was presented in 71. A summary of Chapter 4 was presented in 89 and an extended version was published in 70 . The content of Chapter 5 was published in 90 . The results in Chapter 6 are part of the work presented in [75]. Finally, the research of Chapter 7 was presented in [91].

# Toolbox for proofs of replica symmetric formulas 

In this chapter we briefly review some of the basic tools that are used repeatedly and repeatedly in all our proofs of RS formulas. The reader can quickly browse through the main statements and come back to this chapter when needed.

### 2.1 Nishimori identity

The Nishimori identity is a simple application of Bayes' rule. We abundantly use it in our computations; it is a great tool to simplify expressions and realize that two - apparently different - quantities are equal.

Lemma 2.1 (Nishimori identity). Let $\mathbf{X} \in \mathbb{R}^{n}$ and $\mathbf{Y} \in \mathbb{R}^{m}$ be jointly distributed random vectors. Let $k$ be a positive integer and $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)} k$ independent samples drawn from $P_{\mathbf{X} \mid \mathbf{Y}}(\cdot \mid \mathbf{Y})$, where $P_{\mathbf{X} \mid \mathbf{Y}}$ is the conditional probability distribution of $\mathbf{X}$ given $\mathbf{Y}$. We denote by angular brackets $\langle-\rangle$ the expectation with respect to these samples and by $\mathbb{E}$ the expectation with respect to $(\mathbf{X}, \mathbf{Y})$. Then, for every continuous function $g$ such that $\mathbb{E}\langle | g\left(\mathbf{Y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)}\right)\rangle<+\infty$, we have

$$
\mathbb{E}\left\langle g\left(\mathbf{Y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)}\right)\right\rangle=\mathbb{E}\left\langle g\left(\mathbf{Y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k-1)}, \mathbf{X}\right)\right\rangle .
$$

Proof. Let $P_{\mathbf{X}, \mathbf{Y}}$ be the probability distribution of $(\mathbf{X}, \mathbf{Y})$ and $P_{\mathbf{X}}, P_{\mathbf{Y}}$ the corresponding marginal distribution. A formal computation using Bayes' rule gives

$$
\begin{aligned}
\mathbb{E}\left\langle g\left(\mathbf{Y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)}\right)\right\rangle & =\int g\left(\mathbf{y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)}\right) d P_{\mathbf{Y}}(\mathbf{y}) \prod_{i=1}^{k} d P_{\mathbf{X} \mid \mathbf{Y}}\left(\mathbf{x}^{(k)} \mid \mathbf{y}\right) \\
& =\int g\left(\mathbf{y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k-1)}, \mathbf{x}\right) d P_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, \mathbf{y}) \prod_{i=1}^{k-1} d P_{\mathbf{X} \mid \mathbf{Y}}\left(\mathbf{x}^{(k)} \mid \mathbf{y}\right) \\
& =\int\left\langle g\left(\mathbf{y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k-1)}, \mathbf{x}\right)\right\rangle d P_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, \mathbf{y}) \\
& =\mathbb{E}\left\langle g\left(\mathbf{Y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k-1)}, \mathbf{X}\right)\right\rangle
\end{aligned}
$$

In words, it is equivalent to draw $(\mathbf{X}, \mathbf{Y})$ from its joint distribution, or to first draw $\mathbf{Y}$ from its marginal distribution and then sample $\mathbf{X}$ from its probability distribution given $\mathbf{Y}$. Hence, $\left(\mathbf{Y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)}\right)$ and $\left(\mathbf{Y}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k-1)}, \mathbf{X}\right)$ are equal in law.

Note that the Nishimori identity holds only for Bayes optimal inference. It is not valid when the samples are drawn from a posterior belief (1.6) that does not agree with the true posterior (1.7). In most cases we use two simple instances of the Nishimori identity, namely

$$
\mathbb{E}\left\langle\left\|\mathbf{x}^{(1)}\right\|^{2}\right\rangle=\mathbb{E}\|\mathbf{X}\|^{2} \quad \text { and } \quad \mathbb{E}\left\|\left\langle\mathbf{x}^{(1)}\right\rangle\right\|^{2}=\mathbb{E}\left\langle\mathbf{x}^{(1)} \cdot \mathbf{x}^{(2)}\right\rangle=\mathbb{E}\left\langle\mathbf{x}^{(1)} \cdot \mathbf{X}\right\rangle,
$$

where $\cdot$ denotes the inner product inducing the norm $\|\cdot\|$.

### 2.2 Gaussian integration by parts

Lemma 2.2 (Gaussian integration by parts). Let $Z$ be a standard Gaussian random variable and $f: \mathbb{R} \mapsto \mathbb{R}$ an absolutely continuous function such that $\mathbb{E}\left|f^{\prime}(Z)\right|<+\infty$. Then,

$$
\mathbb{E}[Z f(Z)]=\mathbb{E}\left[f^{\prime}(Z)\right]
$$

Proof. The proof given here follows the one given in [92]. If $f$ has compact support in ( $a, b$ ) then the result follows from an integration by parts,

$$
\mathbb{E}[Z f(Z)]=\int_{a}^{b} z f(z) \frac{e^{-\frac{z^{2}}{2}} d z}{\sqrt{2 \pi}}=\left[-f(z) \frac{e^{-\frac{z^{2}}{2}}}{\sqrt{2 \pi}}\right]_{a}^{b}+\int_{a}^{b} f^{\prime}(z) \frac{e^{-\frac{z^{2}}{2}} d z}{\sqrt{2 \pi}}=\mathbb{E}\left[f^{\prime}(Z)\right]
$$

Otherwise, suppose that all of $\mathbb{E}|f(Z)|, \mathbb{E}|Z f(Z)|, \mathbb{E}\left|f^{\prime}(Z)\right|$ are finite. We consider the sequence of absolutely continuous functions $f_{n}: z \mapsto f(z) g(z / n)$ where $g$ is continuously differentiable, takes value 1 on $[-1,1]$, and is 0 outside $[-2,2]$. Hence, $f_{n}$ is compactly supported on $[-2,2]$ and $\mathbb{E}\left[Z f_{n}(Z)\right]=\mathbb{E}\left[f_{n}^{\prime}(Z)\right]$. Let $C$ and $C^{\prime}$ be positive real numbers that upper bound $|g|$ and $\left|g^{\prime}\right|$, respectively. The sequence $\left(f_{n}\right)_{n \geq 1}$ converges pointwise to $f$ and $\left|z f_{n}(z)\right| \leq C|z f(z)|$ for all $z$ so, by the dominated convergence theorem, $\lim _{n \rightarrow+\infty} \mathbb{E}\left[Z f_{n}(Z)\right]=\mathbb{E}[Z f(Z)]$. Similarly, $\left(f_{n}^{\prime}\right)_{n \geq 1}$ converges pointwise to $f^{\prime}$ and $\left|f_{n}^{\prime}(z)\right| \leq C^{\prime}\left|f^{\prime}(z)\right|+C|f(z)|$ for all $z$ so $\lim _{n \rightarrow+\infty} \mathbb{E}\left[f_{n}^{\prime}(Z)\right]=\mathbb{E}\left[f^{\prime}(Z)\right]$. It follows that $\mathbb{E}[Z f(Z)]=\mathbb{E}\left[f^{\prime}(Z)\right]$.

Finally, if $\mathbb{E}\left|f^{\prime}(Z)\right|<+\infty$ then $\mathbb{E}|Z f(Z)|<+\infty$ because

$$
\begin{aligned}
\int_{0}^{+\infty}|z f(z)| \frac{e^{-\frac{z^{2}}{2}}}{\sqrt{2 \pi}} d z & =\int_{0}^{+\infty} z\left|f(0)+\int_{0}^{z} f^{\prime}(x) d x\right| \frac{e^{-\frac{z^{2}}{2}}}{\sqrt{2 \pi}} d z \\
& \leq \frac{|f(0)|}{\sqrt{2 \pi}}+\int_{0}^{+\infty} z \frac{e^{-\frac{z^{2}}{2}}}{\sqrt{2 \pi}} d z \int_{0}^{z}\left|f^{\prime}(x)\right| d x \\
& \leq \frac{|f(0)|}{\sqrt{2 \pi}}+\int_{0}^{+\infty}\left|f^{\prime}(x)\right| d x \int_{x}^{+\infty} z \frac{e^{-\frac{z^{2}}{2}}}{\sqrt{2 \pi}} d z \leq \frac{|f(0)|}{\sqrt{2 \pi}}+\int_{0}^{+\infty}\left|f^{\prime}(x)\right| \frac{e^{-\frac{x^{2}}{2}}}{\sqrt{2 \pi}} d x
\end{aligned}
$$

and $\mathbb{E}|f(Z)| \leq \sup _{x \in[-1,1]}|f(x)|+\mathbb{E}|Z f(Z)|<+\infty$.

In general, we have to deal with expectations of the form $\mathbb{E}[Z f(Z, \mathbf{U})]$ where $\mathbf{U}$ are all the other sources of randomness in our problem. In this case, if $Z \sim \mathcal{N}(0,1)$ is independent of $\mathbf{U}$, the Gaussian integration by parts reads

$$
\mathbb{E}[Z f(Z, \mathbf{U})]=\mathbb{E}\left[\frac{\partial f(Z, \mathbf{U})}{\partial Z}\right]
$$

### 2.3 Mutual information of simple Gaussian channels

The RS potential functions to optimize in the RS formulas involve mutual informations associated with low-dimensional - most often scalar - Gaussian channels. We rely on properties of these mutual informations to prove the RS formulas.

Lemma 2.3. Let $P_{X}$ be a probability distribution on $\mathbb{R}$ with finite second moment, i.e., $\mathbb{E} X^{2}<+\infty$ where $X \sim P_{X}$. Let $Z$ be a standard Gaussian random variable independent of $X$. For all $R \in \mathbb{R}^{+}$, we define the mutual information

$$
I_{P_{X}}(R):=I(X ; \sqrt{R} X+Z),
$$

and

$$
\psi_{P_{X}}(R):=\mathbb{E}\left[\ln \int d P_{X}(x) \exp \left((R X+\sqrt{R} Z) x-\frac{R x^{2}}{2}\right)\right] .
$$

Then,

$$
I_{P_{X}}(R)=\frac{R \mathbb{E} X^{2}}{2}-\psi_{P_{X}}(R)
$$

The function $\psi_{P_{X}}: R \in \mathbb{R}^{+} \rightarrow \mathbb{R}$ is nondecreasing, convex and Lipschitz continuous with Lipschitz constant ( $\mathbb{E X}^{2} / 2$ ). Equivalently, $I_{P_{X}}: R \in \mathbb{R}^{+} \rightarrow \mathbb{R}^{+}$is nondecreasing, concave and $\left(\mathbb{E}^{2} / 2\right)$-Lipschitz continuous. Besides, if $P_{X}$ is not a deterministic distribution then $\psi_{P_{X}}$ is strictly convex, and $I_{P_{X}}$ is strictly concave.

Lemma 2.4. Let $K$ be a positive integer and $P_{X}$ a probability distribution on $\mathbb{R}^{K}$ with a finite matrix of second moments denoted $\Sigma_{X}:=\mathbb{E} X X^{\top}, X \sim P_{X}$. Let $Z \in \mathbb{R}^{K}$ be a standard Gaussian random vector independent of $X$. Let $S_{K}^{+}$be the cone of $K \times K$ symmetric positive semidefinite matrices. For all $R \in S_{K}^{+}$, we define the mutual information

$$
I_{P_{X}}(R):=I(X ; \sqrt{R} X+Z),
$$

where $\sqrt{R}$ is the positive semidefinite square root of $R$, and

$$
\psi_{P_{X}}(R):=\mathbb{E}\left[\ln \int d P_{X}(x) \exp \left((R X+\sqrt{R} Z)^{\top} x-\frac{x^{\top} R x}{2}\right)\right] .
$$

Then,

$$
I_{P_{X}}(R)=\frac{\operatorname{Tr}\left(R \Sigma_{X}\right)}{2}-\psi_{P_{X}}(R)
$$

The function $\psi_{P_{X}}: R \in \mathbb{R}^{+} \rightarrow \mathbb{R}$ is convex and Lipschitz continuous with Lipschitz constant $\operatorname{Tr}\left(\Sigma_{X}\right) / 2$, that is, $\forall(R, Q) \in\left(S_{K}^{+}\right)^{2}$ :

$$
\left|\psi_{P_{X}}(R)-\psi_{P_{X}}(Q)\right| \leq \frac{\operatorname{Tr}\left(\Sigma_{X}\right)}{2}\|R-Q\|_{2}
$$

where $\|\cdot\|_{2}$ is the matrix norm induced by the Euclidean norm on vectors. Equivalently, $I_{P_{X}}: R \in S_{K}^{+} \rightarrow \mathbb{R}^{+}$is concave and $\operatorname{Tr}\left(\Sigma_{X}\right) / 2$-Lipschitz continuous.

Lemma 2.3 is just a specialization of Lemma 2.4 for $K=1$ and we prove the lemmas for a general $K \geq 1$ directly. It is a good exercise to read through the proof; it shows on a simple example how we use Gaussian integration by parts and the Nishimori identity to compute and simplify the derivative of a mutual information in the rest of this thesis.

Proof of Lemma 2.3 and 2.4. Fix $R \in S_{K}^{+}$and define $Y_{R}:=\sqrt{R} X+Z$. The posterior distribution of X given $Y_{R}$ is simply

$$
\begin{equation*}
d P\left(x \mid Y_{R}\right):=\frac{1}{\mathcal{Z}_{R}\left(Y_{R}\right)} d P_{X}(x) \exp \left(Y_{R}^{\top} \sqrt{R} x-\frac{x^{\top} R x}{2}\right) \tag{2.1}
\end{equation*}
$$

where $\mathcal{Z}_{R}\left(Y_{R}\right)=\int d P_{X}(x) \exp \left(Y_{R}^{\top} \sqrt{R} x-\frac{x^{\top} R x}{2}\right)$. The angular brackets $\langle-\rangle_{R}$ denote an expectation with respect to $x$ drawn from the posterior (2.1). Note that $\psi_{P_{X}}(R)=\mathbb{E} \ln \mathcal{Z}_{R}\left(Y_{R}\right)$ and

$$
I_{P_{X}}(R)=-\mathbb{E}\left[\ln \left(\mathcal{Z}_{R}\left(Y_{R}\right) \frac{e^{-\frac{\|Y\|^{2}}{2}}}{(2 \pi)^{K / 2}}\right)\right]+\mathbb{E}\left[\ln \frac{e^{-\frac{\|Z\|^{2}}{2}}}{(2 \pi)^{K / 2}}\right]=\frac{\operatorname{Tr}\left(R \Sigma_{X}\right)}{2}-\psi_{P_{X}}(R)
$$

Thanks to this identity, the properties of $I_{P_{X}}$ follow directly from those of $\psi_{P_{X}}$.
We now fix $S_{0}, S_{1}$ in the cone of positive definite matrices $S_{K}^{++}$and define $R: t \in[0,1] \mapsto(1-t) S_{0}+t S_{1}$. We prove that $\psi: t \in[0,1] \mapsto \psi_{P_{X}}(R(t))$ is convex and Lipschitz continuous with Lipschitz constant $\operatorname{Tr}\left(\Sigma_{X}\right)\left\|S_{1}-S_{0}\right\|_{2} / 2$. It directly implies the properties of $\psi_{P_{X}}$ on $S_{K}^{++}$and, by continuity, on $S_{K}^{+}$, the closure of $S_{K}^{++}$. The derivative of $\psi: t \mapsto \mathbb{E} \ln \mathcal{Z}_{R(t)}\left(Y_{R(t)}\right)$ reads

$$
\begin{align*}
\psi^{\prime}(t) & =\mathbb{E}\left[\frac{\partial \ln \mathcal{Z}_{R(t)}\left(Y_{R(t)}\right)}{\partial t}\right]=\mathbb{E}\left[\frac{1}{\mathcal{Z}_{R(t)}\left(Y_{R(t)}\right)} \frac{\partial \mathcal{Z}_{R(t)}\left(Y_{R(t)}\right)}{\partial t}\right] \\
& =\mathbb{E}\left[\left\langle X^{\top}\left(S_{1}-S_{0}\right) x-\frac{x^{\top}\left(S_{1}-S_{0}\right) x}{2}+Z^{\top} \frac{d \sqrt{R(t)}}{d t} x\right\rangle_{R(t)}\right] . \tag{2.2}
\end{align*}
$$

To simplify $\psi^{\prime}(t)$, we perform a Gaussian integration by parts with respect to each entry of $Z$,

$$
\begin{align*}
\mathbb{E} & {\left[\left\langle Z^{\top} \frac{d \sqrt{R(t)}}{d t} x\right\rangle_{R(t)}\right]=\mathbb{E}\left[Z^{\top} \int \frac{d \sqrt{R(t)}}{d t} x \frac{e^{(R(t) X+\sqrt{R(t)} Z)^{\top} x-\frac{x^{\top} R(t) x}{2}}}{\mathcal{Z}_{R(t)}\left(Y_{R(t)}\right)} d P_{X}(x)\right] } \\
& =\mathbb{E}\left[\left\langle x^{\top} \sqrt{R(t)} \frac{d \sqrt{R(t)}}{d t} x\right\rangle_{R(t)}\right]-\mathbb{E}\left[\left\langle x^{\top}\right\rangle_{R(t)} \sqrt{R(t)} \frac{d \sqrt{R(t)}}{d t}\langle x\rangle_{R(t)}\right] \tag{2.3}
\end{align*}
$$

By symmetry of $\sqrt{R(t)}$, we have $\forall v \in \mathbb{R}^{K}$ :

$$
\begin{align*}
v^{\top} \sqrt{R(t)} \frac{d \sqrt{R(t)}}{d t} v & =\frac{1}{2} v^{\top} \sqrt{R(t)} \frac{d \sqrt{R(t)}}{d t} v+\frac{1}{2} v^{\top}\left(\sqrt{R(t)} \frac{d \sqrt{R(t)}}{d t}\right)^{\top} v \\
& =v^{\top}\left(\frac{\sqrt{R(t)}}{2} \frac{d \sqrt{R(t)}}{d t}+\frac{d \sqrt{R(t)}}{d t} \frac{\sqrt{R(t)}}{2}\right) v \\
& =v^{\top} \frac{d R(t)}{2 d t} v=v^{\top} \frac{S_{1}-S_{0}}{2} v . \tag{2.4}
\end{align*}
$$

Thanks to (2.4), (2.3) further simplifies to

$$
\mathbb{E}\left[\left\langle Z^{\top} \frac{d \sqrt{R(t)}}{d t} x\right\rangle_{R(t)}\right]=\mathbb{E}\left[\left\langle x^{\top} \frac{S_{1}-S_{0}}{2} x\right\rangle_{R(t)}\right]-\mathbb{E}\left[\left\langle x^{\top}\right\rangle_{R(t)} \frac{S_{1}-S_{0}}{2}\langle x\rangle_{R(t)}\right] .
$$

We plug this last identity back in (2.2), and use the Nishimori identity

$$
\mathbb{E}\left[X^{\top}\left(S_{1}-S_{0}\right)\langle x\rangle_{R(t)}\right]=\mathbb{E}\left[\left\langle x^{\top}\right\rangle_{R(t)}\left(S_{1}-S_{0}\right)\langle x\rangle_{R(t)}\right]
$$

to finally obtain

$$
\psi^{\prime}(t)=\mathbb{E}\left[\left\langle x^{\boldsymbol{\top}}\right\rangle_{R(t)} \frac{S_{1}-S_{0}}{2}\langle x\rangle_{R(t)}\right] .
$$

When $K=1$ and $S_{0}<S_{1}$, we see that $\psi^{\prime}(t)=\left(S_{1}-S_{0}\right) / 2 \mathbb{E}\left[\langle x\rangle_{R(t)}^{2}\right] \geq 0$ so $\psi_{P_{X}}$ is nondecreasing on $[0,+\infty)$. Whatever $K$, we have for all $t \in[0,1]$ :

$$
\begin{equation*}
\left|\psi^{\prime}(t)\right| \leq \mathbb{E}\left|\left\langle x^{\top}\right\rangle_{R(t)} \frac{S_{1}-S_{0}}{2}\langle x\rangle_{R(t)}\right| \leq \mathbb{E}\left[\left\|\langle x\rangle_{R(t)}\right\|^{2}\right] \frac{\left\|S_{1}-S_{0}\right\|_{2}}{2}, \tag{2.5}
\end{equation*}
$$

where the first inequality follows from Jensen's inequality and the second from the definition of the matrix norm. Applying first Jensen's inequality, and then the Nishimori identity, the expectation on the right-hand side of (2.5) satisfies

$$
\mathbb{E}\left[\left\|\langle x\rangle_{R(t)}\right\|^{2}\right] \leq \mathbb{E}\left\langle\|x\|^{2}\right\rangle_{R(t)}=\mathbb{E}\|X\|^{2}=\operatorname{Tr}\left(\Sigma_{X}\right)
$$

Hence, $\left|\psi^{\prime}(t)\right| \leq \operatorname{Tr}\left(\Sigma_{X}\right)\left\|S_{1}-S_{0}\right\|_{2} / 2$ and $\psi$ is Lipschitz continuous with Lipschitz constant $\operatorname{Tr}\left(\Sigma_{X}\right)\left\|S_{1}-S_{0}\right\|_{2} / 2$. We now differentiate $\psi^{\prime}(t)=\mathbb{E}\left[X^{\top} \frac{S_{1}-S_{0}}{2}\langle x\rangle_{R(t)}\right]$,

$$
\begin{align*}
& \psi^{\prime \prime}(t)= \mathbb{E}\left[X^{\top} \frac{S_{1}-S_{0}}{2}\left\langle x\left(X^{\top}\left(S_{1}-S_{0}\right) x-\frac{x^{\top}\left(S_{1}-S_{0}\right) x}{2}+Z^{\top} \frac{d \sqrt{R(t)}}{d t} x\right)\right\rangle_{R(t)}\right] \\
&-\mathbb{E}\left[X^{\top} \frac{S_{1}-S_{0}}{2}\langle x\rangle_{R(t)}\left\langle X^{\top}\left(S_{1}-S_{0}\right) x-\frac{x^{\top}\left(S_{1}-S_{0}\right) x}{2}+Z^{\top} \frac{d \sqrt{R(t)}}{d t} x\right\rangle_{R(t)}\right] \\
&= \frac{1}{2} \mathbb{E}\left[\left\langle\left(X^{\top}\left(S_{1}-S_{0}\right) x\right)^{2}\right\rangle_{R(t)}\right] \\
&-\mathbb{E}\left[\left(X^{\top}\left(S_{1}-S_{0}\right)\langle x\rangle_{R(t)}\right)^{2}\right]  \tag{2.6}\\
&+\frac{1}{2} \mathbb{E}\left[\left(\langle x\rangle_{R(t)}^{\top}\left(S_{1}-S_{0}\right)\langle x\rangle_{R(t)}\right)^{2}\right]
\end{align*}
$$

To get the second equality, we first perform a Gaussian integration by parts with respect to $Z$, and then use the Nishimori identity and (2.4) to further simplify the
expression. We want to show that $\psi^{\prime \prime}(t) \geq 0$ to prove that $\psi$ is convex. We have

$$
\begin{aligned}
& \left.\mathbb{E}\left[\left\langle\left(X^{\top}\left(S_{1}-S_{0}\right) x\right)^{2}\right\rangle_{R(t)}\right]-\mathbb{E}\left[\left\langle X^{\top}\left(S_{1}-S_{0}\right)\right) x\right\rangle_{R(t)}^{2}\right] \\
& \quad=\mathbb{E}\left\langle\left(X^{\top}\left(S_{1}-S_{0}\right)\left(x-\langle x\rangle_{R(t)}\right)\right)^{2}\right\rangle_{R(t)} \\
& \begin{aligned}
\mathbb{E}\left[\left(X^{\top}\left(S_{1}-S_{0}\right)\langle x\rangle_{R(t)}\right)^{2}\right]-\mathbb{E}\left[\left(\langle x\rangle_{R(t)}^{\top}\right.\right. & \left.\left.\left(S_{1}-S_{0}\right)\langle x\rangle_{R(t)}\right)^{2}\right] \\
& =\mathbb{E}\left[\left(\langle x\rangle_{R(t)}^{\top}\left(S_{1}-S_{0}\right)\left(X-\langle x\rangle_{R(t)}\right)\right)^{2}\right]
\end{aligned}
\end{aligned}
$$

Both identities can be checked by extending the square on the right-hand side and using the Nishimori identity $\mathbb{E}\left[\left(X^{\top}\left(S_{1}-S_{0}\right)\langle x\rangle_{R(t)}\right)^{2}\right]=\mathbb{E}\left[\left(\langle x\rangle_{R(t)}^{\top}\left(S_{1}-S_{0}\right)\langle x\rangle_{R(t)}\right)^{2}\right]$. We use these identities to simplify (2.6) further,

$$
\psi^{\prime \prime}(t)=\frac{\mathbb{E}\left\langle\left(X^{\top}\left(S_{1}-S_{0}\right)\left(x-\langle x\rangle_{R(t)}\right)\right)^{2}\right\rangle_{R(t)}-\mathbb{E}\left[\left(\langle x\rangle_{R(t)}^{\top}\left(S_{1}-S_{0}\right)\left(X-\langle x\rangle_{R(t)}\right)\right)^{2}\right]}{2} .
$$

The numerator in this last expression is nonnegative because

$$
\begin{align*}
\mathbb{E}\left[\left(\langle x\rangle_{R(t)}^{\top}\left(S_{1}-S_{0}\right)\left(X-\langle x\rangle_{R(t)}\right)\right)^{2}\right] & =\mathbb{E}\left\langle x^{\top}\left(S_{1}-S_{0}\right)\left(X-\langle x\rangle_{R(t)}\right)\right\rangle_{R(t)}^{2} \\
& \leq \mathbb{E}\left\langle\left(x^{\top}\left(S_{1}-S_{0}\right)\left(X-\langle x\rangle_{R(t)}\right)\right)^{2}\right\rangle_{R(t)} \\
& =\mathbb{E}\left\langle\left(X^{\top}\left(S_{1}-S_{0}\right)\left(x-\langle x\rangle_{R(t)}\right)\right)^{2}\right\rangle_{R(t)} \tag{2.7}
\end{align*}
$$

where the inequality follows from Jensen's inequality and the subsequent equality from the Nishimori identity. All in all, $\psi^{\prime \prime}(t) \geq 0$ for all $t \in[0,1]$ and $\psi$ is convex. When $K=1$, we can directly factor (2.6) into

$$
\psi^{\prime \prime}(t)=\frac{\left(S_{1}-S_{0}\right)^{2}}{2} \mathbb{E}\left[\left(\left\langle x^{2}\right\rangle_{R(t)}-\langle x\rangle_{R(t)}^{2}\right)^{2}\right]=\frac{\left(S_{1}-S_{0}\right)^{2}}{2} \mathbb{E}\left[\left\langle\left(x-\langle x\rangle_{R(t)}\right)^{2}\right\rangle_{R(t)}^{2}\right]
$$

If $S_{0}<S_{1}$, we see that $\psi^{\prime \prime}(t)=0$ if, and only if, $x=\langle x\rangle_{R(t)}$ almost surely (where $\left.x \sim d P\left(\cdot \mid Y_{R(t)}\right)\right)$. The latter is equivalent to $P_{X}$ being deterministic. Hence, $\psi$ is strictly convex if $P_{X}$ is not deterministic.

### 2.4 Concentration inequalities

We explain in Section 1.7 how we prove a RS formula for the asymptotic normalized mutual information by interpolating from the original channel at $t=0$ to a simpler channel at $t=1$. To understand how the mutual informations associated with the two extremes relate to each other, we study the derivative of the normalized mutual information with respect to $t \in(0,1)$.

As discussed in Section 1.4, the normalized mutual information at $t$ is directly related to the average free entropy $\mathbb{E} \ln \mathcal{Z}_{t} / n$, where $\mathcal{Z}_{t}$ is the normalization factor of the posterior distribution associated with the interpolating estimation problem. Being a function of the noisy observations $\mathbf{Y}_{t}, \mathcal{Z}_{t}$ is itself a random variable. In order to control some terms in the derivative of the normalized mutual information, we have to show that the free entropy $\ln \mathcal{Z}_{t} / n$ is tightly concentrated around its mean $\mathbb{E} \ln \mathcal{Z}_{t} / n$ in the high-dimensional limit $n \rightarrow+\infty$. We prove the latter thanks to the classical variance bounds stated below.

Proposition 2.5 (Efron-Stein inequality). Let $U_{1}, \ldots, U_{N}, U_{1}^{\prime}, \ldots, U_{N}^{\prime}$ be $2 N$ independent random variables with $U_{i}$ and $U_{i}^{\prime}$ having the same distribution for all $i \in\{1, \ldots, N\}$. Let $\mathcal{U} \subseteq \mathbb{R}^{N}$ be the support of $\mathbf{U}:=\left(U_{1}, U_{2}, \ldots, U_{N}\right)$ and $g: \mathcal{U} \rightarrow \mathbb{R}$ a function such that $g(\mathbf{U})$ is square integrable. For all $i \in\{1, \ldots, N\}$, $\mathbf{U}^{(i)}$ is the random vector obtained from $\mathbf{U}$ by replacing $U_{i}$ by $U_{i}^{\prime}$, that is,

$$
\mathbf{U}^{(i)}:=\left(U_{1}, \ldots, U_{i-1}, U_{i}^{\prime}, U_{i+1}, \ldots, U_{N}\right)
$$

Then,

$$
\operatorname{Var} g(\mathbf{U}) \leq \frac{1}{2} \sum_{i=1}^{N} \mathbb{E}\left[\left(g(\mathbf{U})-g\left(\mathbf{U}^{(i)}\right)\right)^{2}\right] .
$$

Proposition 2.6 (McDiarmid's inequality). Let $\mathcal{U}$ be a subset of $\mathbb{R}$. Let $g: \mathcal{U}^{N} \rightarrow \mathbb{R}$ be a function that satisfies the bounded difference property, i.e., there exist constants $c_{1}, c_{2}, \ldots, c_{N}$ such that $\forall i \in\{1, \ldots, N\}$ :

$$
\sup _{\substack{\left(u_{1}, \ldots, u_{N}\right) \in \mathcal{U}^{N} \\ u_{i}^{\prime} \in \mathcal{U}}}\left|g\left(u_{1}, \ldots, u_{i}, \ldots, u_{N}\right)-g\left(u_{1}, \ldots, u_{i-1}, u_{i}^{\prime}, u_{i+1}, \ldots, u_{N}\right)\right| \leq c_{i}
$$

Let $\mathbf{U}:=\left(U_{1}, U_{2}, \ldots, U_{N}\right)$ where $U_{1}, U_{2} \ldots, U_{N}$ are independent random variables supported on $\mathcal{U}$. Then,

$$
\operatorname{Var} g(\mathbf{U}) \leq \frac{1}{4} \sum_{i=1}^{N} c_{i}^{2}
$$

Proposition 2.7 (Gaussian Poincaré inequality). Let $\mathbf{U}=\left(U_{1}, U_{2}, \ldots, U_{N}\right)$ be a vector of $N$ independent standard Gaussian random variables. Let $g: \mathbb{R}^{N} \rightarrow \mathbb{R}$ be a continuously differentiable function and $\nabla g: \mathbb{R}^{N} \rightarrow \mathbb{R}^{N}$ its gradient. Then,

$$
\operatorname{Var} g(\mathbf{U}) \leq \mathbb{E}\|\nabla g(\mathbf{U})\|^{2}
$$

McDiarmid's inequality and the Gaussian Poincaré inequality are proved thanks to the more general Efron-Stein inequality. We refer to [93, Chapter 3] for proofs of all these propositions.

We stated earlier that we use the concentration of the free entropy to control terms in the derivative of the normalized mutual information. However, the average free entropy is equal (up to some additive term) to the normalized mutual information, not to its derivative. Hence, what we ultimately need is a result on the concentration of the partial derivative of the free entropy with respect to an SNR. Thanks to the next lemma, we can use upper bounds on the variance of the free entropy $\ln \mathcal{Z}_{t} / n$ in order to bound the variations of the partial derivative.

Lemma 2.8 (Bound on difference of derivatives of convex functions). Let $G$ and $g$ be two differentiable convex functions defined on a real interval $I \subseteq \mathbb{R}$. Let $r \in I$ and $\delta>0$ be such that $r \pm \delta \in I$. Then, $C_{\delta}(r):=g^{\prime}(r+\delta)-g^{\prime}(r-\delta) \geq 0$ and

$$
\left|G^{\prime}(r)-g^{\prime}(r)\right| \leq C_{\delta}(r)+\frac{1}{\delta} \sum_{u \in\{-\delta, 0, \delta\}}|G(r+u)-g(r+u)| .
$$

Proof. Fix $r \in I$ and $\delta>0$ such that $r \pm \delta \in I$. The functions $G$ and $g$ are convex so $G^{\prime}(r) \leq \frac{G(r+\delta)-G(r)}{\delta}, g^{\prime}(r-\delta) \leq g^{\prime}(r)$ and $\frac{g(r+\delta)-g(r)}{\delta} \leq g^{\prime}(r+\delta)$. Hence,

$$
\begin{align*}
G^{\prime}(r)-g^{\prime}(r) & \leq \frac{G(r+\delta)-G(r)}{\delta}-g^{\prime}(r-\delta) \\
& \leq \frac{G(r+\delta)-G(r)}{\delta}-\frac{g(r+\delta)-g(r)}{\delta}+g^{\prime}(r+\delta)-g^{\prime}(r-\delta) \\
& =\frac{G(r+\delta)-g(r+\delta)}{\delta}-\frac{G(r)-g(r)}{\delta}+C_{\delta}(r) \\
& =\frac{|G(r+\delta)-g(r+\delta)|}{\delta}+\frac{|G(r)-g(r)|}{\delta}+C_{\delta}(r) \tag{2.8}
\end{align*}
$$

Similarly, using $g^{\prime}(r) \leq g^{\prime}(r+\delta), \frac{G(r)-G(r-\delta)}{\delta} \leq G^{\prime}(r)$ and $g^{\prime}(r-\delta) \leq \frac{g(r)-g(r-\delta)}{\delta}$, we have

$$
\begin{align*}
g^{\prime}(r)-G^{\prime}(r) & \leq g^{\prime}(r+\delta)-\frac{G(r)-G(r-\delta)}{\delta} \\
& \leq g^{\prime}(r+\delta)-g^{\prime}(r-\delta)+\frac{g(r)-g(r-\delta)}{\delta}-\frac{G(r)-G(r-\delta)}{\delta} \\
& =C_{\delta}(r)+\frac{G(r-\delta)-g(r-\delta)}{\delta}+\frac{g(r)-G(r)}{\delta} \\
& =\frac{|G(r-\delta)-g(r-\delta)|}{\delta}+\frac{|G(r)-g(r)|}{\delta}+C_{\delta}(r) . \tag{2.9}
\end{align*}
$$

The desired upper bound on $\left|G^{\prime}(r)-g^{\prime}(r)\right|$ directly follows from (2.8) and (2.9).
Roughly speaking, we apply Lemma 2.8 to $g(r)=-\ln \mathcal{Z}_{t} / n+C r^{2}, G(r)=\mathbb{E}[g(r)]$, where $r$ is an SNR of the problem and the quadratic perturbation $C r^{2}$ with $C$ a positive random variable makes $g$ convex (then $G$ is convex as well).

## Part I

## Tensor estimation

# High-dimensional rank-one nonsymmetric matrix estimation: the spherical case 

## 3

### 3.1 Introduction

Tensor decomposition, which originated with Hitchcock in 1927 [2], has found many applications in signal processing, graph analysis, data mining and machine learning in the past two decades [6], [13], [94]. While tensor decomposition was originally developed in a deterministic and algebraic context, it is of interest for these applications to develop a statistical approach [19]. Some important questions in this setting are, for example, under which conditions and how can we recover a low-rank tensor - the signal of interest - from noisy observations of it? This work focuses on answering - at least in part - these questions in the most elementary, but yet rich, setting of a nonsymmetric rank-one matrix signal buried within noise. Namely, we observe under additive white Gaussian noise (AWGN) a $n_{u} \times n_{v}$ rank-one matrix $\mathbf{U V}^{\top}$ where $\mathbf{U}$ and $\mathbf{V}$ are random vectors that we wish to recover as well as possible. This problem, and its symmetric version, have generated important results in the past ten years [95], [96].

Our approach is in the continuity of a line of research establishing lowdimensional variational formulas for the normalized mutual information between a signal of interest and noisy observations in the high-dimensional regime 29, [36], [46], [97. Such formulas are valuable because they link the mutual information of a high-dimensional channel whose outputs are coupled to those of simple decoupled scalar channels. We can then determine, by solving a low-dimensional variational problem, phase transitions as well as performance measures related to the minimum mean-square error (MMSE). We also gain important insight on the performance of (message passing) algorithms designed to estimate input signals. In fact, the fixed points of the state evolution equations tracking the performance of approximate message passing in the high-dimensional regime can be identified among the critical points of the variational expression for the mutual information.

For the problem at hand, the variational formula - that was predicted using the replica trick from statistical physics - has already been proven rigorously when $\mathbf{U}$ and $\mathbf{V}$ have independent and identically distributed (i.i.d.) entries 36], 46].

These results were extended beyond the matrix case to rank-one nonsymmetric tensor decomposition [69], [84]. The replica prediction has also been shown to be true for low-rank symmetric tensor decomposition [36], [89].

A natural follow-up interrogation is what happens when either $\mathbf{U}$ or $\mathbf{V}$ doesn't have independent entries anymore. Can the average mutual information in the high-dimensional regime still be given by a simple, low-dimensional, variational formula? In this work, we study the simple case in which both $\mathbf{U}$ and $\mathbf{V}$ are uniformly distributed on spheres (whose radii scale like $\sqrt{n_{u}}$ and $\sqrt{n_{v}}$, respectively) and give a rigorous and positive answer to the question above. To the best of our knowledge fully rigorous results on this issue are scarce. Recently, [98] analyzed (under natural assumptions) another situation in which $\mathbf{U}$ and $\mathbf{V}$ are generated by a generalized linear model.

In Section 3.2 we present the problem and our main results. In Section 3.3 we give the reader an outline of the proof of the variational formula for the average mutual information. We conclude in Section 3.4 with a discussion of the relation between the present problem and the classical spherical spin-glass model of statistical mechanics.

### 3.2 Problem setting and main results

Let $n_{u}, n_{v}$ be positive integers and $\rho_{u}, \rho_{v}$ positive real numbers. Let $\mathbf{U} \in \mathbb{R}^{n_{u}}$ and $\mathbf{V} \in \mathbb{R}^{n_{v}}$ be independent random vectors uniformly distributed on the spheres of radii $\sqrt{\rho_{u} n_{u}}$ and $\sqrt{\rho_{v} n_{v}}$, respectively. We denote $P_{u}$ and $P_{v}$ their respective probability distributions. We consider the task of estimating both vectors $\mathbf{U}$ and $\mathbf{V}$ from a noisy version of the scaled rank-one matrix $\mathbf{U V}^{\top}$. More precisely, we observe the matrix $\mathbf{Y} \in \mathbb{R}^{n_{u} \times n_{v}}$ whose entries are $\forall(i, j) \in\left\{1, \ldots, n_{u}\right\} \times\left\{1, \ldots, n_{v}\right\}$ :

$$
\begin{equation*}
Y_{i j}:=\sqrt{\frac{\lambda}{n}} U_{i} V_{j}+Z_{i j} . \tag{3.1}
\end{equation*}
$$

Here, the elements of the noise $\mathbf{Z}:=\left\{Z_{i j}\right\}_{i, j} \in \mathbb{R}^{n_{u} \times n_{v}}$ are standard Gaussian random variables, the positive real number $\lambda$ plays the role of a signal-to-noise ratio (SNR), and the positive integer $n$ scales like $n_{u}$ and $n_{v}$, i.e., there exist positive real numbers $\alpha_{u}$ and $\alpha_{v}$ such that

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{n_{u}}{n}=\alpha_{u}, \lim _{n \rightarrow+\infty} \frac{n_{v}}{n}=\alpha_{v} \tag{3.2}
\end{equation*}
$$

The normalization $1 / \sqrt{n}$ in (3.1) with the scaling (3.2) makes the estimation problem nontrivial. Finally, we define the vector of hyperparameters for this problem, $\Theta:=\left[\begin{array}{lllll}\lambda & \alpha_{u} & \alpha_{v} & \rho_{u} & \rho_{v}\end{array}\right]$.

### 3.2.1 Variational formula for the normalized mutual information

A central role is played by the normalized mutual information associated with a simple linear AWGN channel.

Lemma 3.1. Let $\mathbf{X}$ be a n-dimensional random vector uniformly distributed on the sphere of radius $\sqrt{n}$. The vector $\mathbf{X}$ is observed at the output of the noisy linear channel

$$
\begin{equation*}
\widetilde{\mathbf{Y}}:=\sqrt{m} \mathbf{X}+\widetilde{\mathbf{Z}}, \tag{3.3}
\end{equation*}
$$

where $\widetilde{Z}_{i} \stackrel{i . i . d .}{\sim} \mathcal{N}(0,1)$ for $i=1, \ldots, n$ and $m>0$ plays the role of a SNR. The normalized mutual information between $\mathbf{X}$ and $\widetilde{\mathbf{Y}}$ converges in the high-dimensional limit and

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \tilde{\mathbf{Y}})}{n}=\frac{\ln (1+m)}{2} \tag{3.4}
\end{equation*}
$$

Note that the limit is equal to the normalized mutual information between $\mathbf{X}$ and $\widetilde{\mathbf{Y}}$ where this time the entries of the signal $\mathbf{X}$ are i.i.d. with respect to $\mathcal{N}(0,1)$. It is well-known that such a vector $\mathbf{X}$ is approximately uniformly distributed on the sphere of radius $\sqrt{n}$ in high-dimension [99, Section 3.3.3]. We now use (3.4) to describe the normalized mutual information between $(\mathbf{U}, \mathbf{V})$ and $\mathbf{Y}$.

Theorem 3.2. Define the replica symmetric ( $R S$ ) potential function

$$
\begin{align*}
i_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right):=\frac{\lambda \alpha_{u} \alpha_{v}}{2} & \left(\rho_{u}-m_{u}\right)\left(\rho_{v}-m_{v}\right) \\
& +\alpha_{u} \frac{\ln \left(1+\lambda \alpha_{v} \rho_{u} m_{v}\right)}{2}+\alpha_{v} \frac{\ln \left(1+\lambda \alpha_{u} \rho_{v} m_{u}\right)}{2} . \tag{3.5}
\end{align*}
$$

In the high-dimensional limit $n \rightarrow+\infty$ where $n_{u} / n \rightarrow \alpha_{u}$ and $n_{v} / n \rightarrow \alpha_{v}$, the normalized mutual information between $(\mathbf{U}, \mathbf{V})$ and $\mathbf{Y}$ defined by (3.1) satisfies

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{I(\mathbf{U}, \mathbf{V} ; \mathbf{Y})}{n}=\inf _{m_{u} \in\left[0, \rho_{u}\right]} \sup _{m_{v} \in\left[0, \rho_{v}\right]} i_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right) \tag{3.6}
\end{equation*}
$$

We prove Theorem 3.2 in Section 3.3. Note that the last two summands in (3.5) are the asymptotic normalized mutual informations associated with two decoupled linear AWGN channels ( $\widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}$ are standard Gaussian random vectors),

$$
\begin{aligned}
\lim _{n \rightarrow+\infty} \frac{I\left(\mathbf{U} ; \sqrt{\lambda \alpha_{v} m_{v}} \mathbf{U}+\widetilde{\mathbf{Z}}\right)}{n} & =\frac{\alpha_{u} \ln \left(1+\lambda \alpha_{v} \rho_{u} m_{v}\right)}{2} \\
\lim _{n \rightarrow+\infty} \frac{I\left(\mathbf{V} ; \sqrt{\lambda \alpha_{u} m_{u}} \mathbf{V}+\overline{\mathbf{Z}}\right)}{n} & =\frac{\alpha_{v} \ln \left(1+\lambda \alpha_{u} \rho_{v} m_{u}\right)}{2}
\end{aligned}
$$

We remark that the limit of $I(\mathbf{U}, \mathbf{V} ; \mathbf{Y}) / n$ is the same if both $\mathbf{U}$ and $\mathbf{V}$ have i.i.d. standard Gaussian components (see [46], [69]). The equivalence of the spherical and Gaussian cases is not an obvious fact when it comes to make a precise argument. We discuss this point further in Section 3.4.

### 3.2.2 Minimum mean-square error

It is well-known that the mean-square error of an estimator of $\mathbf{U V}^{\top}$ that is only a function of $\mathbf{Y}$ is minimized by the posterior mean $\mathbb{E}\left[\mathbf{U V}^{\top} \mid \mathbf{Y}\right]$. We denote by
$\operatorname{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right)$ the minimum mean-square error,

$$
\begin{equation*}
\left.\operatorname{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right): \left.=\frac{\mathbb{E} \| \mathbf{U V}^{\top}-\mathbb{E}[\mathbf{U V}}{}{ }^{\top} \right\rvert\, \mathbf{Y}\right] \|^{2} n_{u} n_{v} \tag{3.7}
\end{equation*}
$$

It depends on $\lambda$ through the observations $\mathbf{Y}$. We combine Theorem 3.2 with the I-MMSE relationship [51]

$$
\frac{\partial}{\partial \lambda}\left(\frac{I(\mathbf{U}, \mathbf{V} ; \mathbf{Y})}{n}\right)=\frac{n_{u}}{n} \frac{n_{v}}{n} \frac{\mathrm{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right)}{2}
$$

to get the asymptotic MMSE. The proof is given in Appendix 3.F.
Theorem 3.3. Define $\lambda_{\mathrm{IT}}:=\left(\rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}\right)^{-1}$ and for all $\lambda \in(0,+\infty)$ :

$$
\left(m_{u}^{*}(\lambda), m_{v}^{*}(\lambda)\right)=\left\{\begin{array}{cl}
(0,0) & \text { if } 0<\lambda \leq \lambda_{\mathrm{IT}} \\
\left(\frac{\lambda^{2} \alpha_{u} \alpha_{v} \rho_{v} \rho_{u}^{2}-1}{\lambda \alpha_{u} \rho_{v}\left(1+\lambda \alpha_{v} \rho_{v} \rho_{u}\right)}, \frac{\lambda^{2} \alpha_{u} \alpha_{v} \rho_{v}^{2} \rho_{u}^{2}-1}{\lambda \alpha_{v} \rho_{u}\left(1+\lambda \alpha_{u} \rho_{v} \rho_{u}\right)}\right) & \text { if } \lambda>\lambda_{\mathrm{IT}}
\end{array} .\right.
$$

The pair $\left(m_{u}^{*}(\lambda), m_{v}^{*}(\lambda)\right)$ is the unique solution to the extremization over $\left(m_{u}, m_{v}\right)$ on the right-hand side of (3.6), and $\operatorname{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right)$ satisfies

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \operatorname{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right)=\rho_{u} \rho_{v}-m_{u}^{*}(\lambda) m_{v}^{*}(\lambda) \tag{3.8}
\end{equation*}
$$

Hence, the asymptotic MMSE is less than $\rho_{u} \rho_{v}$ if, and only if, $\lambda>\lambda_{\mathrm{IT}}$.
Theorems 3.2 and 3.3 provide important insight on the inference problem. Nonanalytic points of (3.6) correspond to the location of phase transitions where the MMSE changes behavior. In the present problem, we find by an explicit analysis a unique continuous phase transition point $\lambda_{\mathrm{IT}}$. The mutual information is continuously differentiable for all $\lambda>0$ and its second derivative has a jump at $\lambda_{\text {IT }}$. Correspondingly, the MMSE is continuous with a jump in its first derivative at $\lambda_{\text {IT }}$. More precisely, the MMSE is $\rho_{u} \rho_{v}$ for $\lambda \leq \lambda_{\text {IT }}$ and it continuously departs from $\rho_{u} \rho_{v}$ once $\lambda$ becomes greater than $\lambda_{\text {IT }}$. Thus, $\lambda_{\text {IT }}$ is the lowest SNR for which an estimate of the matrix $\mathbf{U V}^{\top}$ is information-theoretically possible. The general phenomenological picture has been uncovered in a number of situations (including richer ones) by direct analysis of the RS formula for the asymptotic normalized mutual information. We refer to [97] for more details.

In Figure 3.1, we use Theorem 3.3 and draw the asymptotic MMSE as a function of the signal-to-noise ratio $\lambda$ for different values of $\alpha_{v}$ and $\alpha_{u}=\rho_{u}=\rho_{v}=1$. We see that the asymptotic MMSE decreases as $\lambda$ or $\alpha_{v}$ increases. In Figure 3.2, we plot the asymptotic MMSE in the plane $(x, y)=\left(\lambda \rho_{u} \rho_{v} \alpha_{u}, \lambda \rho_{u} \rho_{v} \alpha_{v}\right)$. Note that $y=x^{-1}$ if, and only, if $\lambda=\lambda_{\text {IT }}$. Above the curve $y=x^{-1}$ it becomes statistically possible to do better than the trivial estimate $\mathbb{E}\left[\mathbf{U V}^{\top}\right]=0$.

### 3.3 Proof of Theorem 3.2

We present in this section the main ideas and steps in the proof of Theorem 3.2, We refer to the appendices, that contain all the technicalities of the proof, when


Figure 3.1: Asymptotic MMSE, normalized by $\rho_{u} \rho_{v}$, as a function of $\lambda$ for different values of $\alpha_{v}$ and $\alpha_{u}=\rho_{u}=\rho_{v}=1$. Each of the dashed vertical lines is located at the information-theoretic threshold $\lambda_{\text {IT }}$ of the corresponding $\alpha_{v}$.


Figure 3.2: Asymptotic MMSE, normalized by $\rho_{u} \rho_{v}$, plotted in the plane $(x, y)=\left(\lambda \rho_{u} \rho_{v} \alpha_{u}, \lambda \rho_{u} \rho_{v} \alpha_{v}\right)$. The dashed orange curve is the curve $y=x^{-1} \Leftrightarrow \lambda=\lambda_{\text {IT }}$. Below this curve the normalized MMSE is maximum equal to 1, i.e., it is not possible to give an estimate better than $\mathbb{E}\left[\mathbf{U V}^{\top}\right]=0$.
necessary. The proof is based on the adaptive interpolation method introduced in [37, [87]. The main difference with the canonical interpolation method developed by Guerra and Toninelli in the context of spin glasses [60], [61] is the increased flexibility in choosing the path followed by the interpolation between its two
extremes. By choosing two different interpolation paths, we bound the asymptotic normalized mutual information from above and below by the same variational formula. We can reduce the proof to the case $\lambda=1$ as we can always rescale $\rho_{u}$ to $\lambda \rho_{u}$ and $\lambda$ to 1 .

### 3.3.1 Adaptive path interpolation

We introduce a real parameter $t \in[0,1]$. The adaptive interpolation interpolates from the original channel (3.1) at $t=0$ to two independent channels similar to (3.3) at $t=1$, one for $\mathbf{U}$ and the other for $\mathbf{V}$. In between, we follow an interpolation path

$$
R(\cdot, \epsilon)=\left(R_{u}(\cdot, \epsilon), R_{v}(\cdot, \epsilon)\right),
$$

where $R_{u}(\cdot, \epsilon)$ and $R_{v}(\cdot, \epsilon)$ are continuously differentiable functions from $[0,1]$ to $[0,+\infty)$ parametrized by a "small perturbation" $\epsilon=\left(\epsilon_{u}, \epsilon_{v}\right) \in[0,+\infty)^{2}$ and such that $R(0, \epsilon)=\epsilon$. More precisely, for $t \in[0,1]$, we observe

$$
\begin{cases}\mathbf{Y}^{(t)} & :=\sqrt{\frac{1-t}{n}} \mathbf{U}^{\mathbf{\top}}+\mathbf{Z}  \tag{3.9}\\ \widetilde{\mathbf{Y}}^{(t, \epsilon)} & :=\sqrt{\alpha_{v} R_{v}(t, \epsilon)} \mathbf{U}+\widetilde{\mathbf{Z}} \\ \overline{\mathbf{Y}}^{(t, \epsilon)} & :=\sqrt{\alpha_{u} R_{u}(t, \epsilon)} \mathbf{V}+\overline{\mathbf{Z}}\end{cases}
$$

where $\mathbf{U} \sim P_{u}, \mathbf{V} \sim P_{v}$ and all the noises $\mathbf{Z} \in \mathbb{R}^{n_{u} \times n_{v}}, \widetilde{\mathbf{Z}} \in \mathbb{R}^{n_{u}}, \overline{\mathbf{Z}} \in \mathbb{R}^{n_{v}}$ have i.i.d. entries with respect to $\mathcal{N}(0,1)$. Applying Bayes' rule, the posterior distribution of $(\mathbf{U}, \mathbf{V})$ given $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)$ is

$$
\begin{equation*}
d P\left(\mathbf{u}, \mathbf{v} \mid \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right):=\frac{d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)}}{\mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)}, \tag{3.10}
\end{equation*}
$$

where we introduced the interpolating Hamiltonian

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right):= & \sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} \frac{1-t}{2 n} u_{i}^{2} v_{j}^{2}-\sqrt{\frac{1-t}{n}} u_{i} v_{j} Y_{i j}^{(t)} \\
& +\sum_{i=1}^{n_{u}} \frac{\alpha_{v} R_{v}(t, \epsilon)}{2} u_{i}^{2}-\sqrt{\alpha_{v} R_{v}(t, \epsilon)} u_{i} \widetilde{Y}_{i}^{(t, \epsilon)} \\
& +\sum_{j=1}^{n_{v}} \frac{\alpha_{u} R_{u}(t, \epsilon)}{2} v_{j}^{2}-\sqrt{\alpha_{u} R_{u}(t, \epsilon)} v_{j} \bar{Y}_{j}^{(t, \epsilon)} \tag{3.11}
\end{align*}
$$

and $\mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)$ properly normalizes the posterior. Note that 3.11) could be simplified using the spherical constraints,e.g., $\sum_{i} u_{i}^{2}=\rho_{u} n_{u}$, but this general form is convenient for the analysis. We denote by angular brackets $\langle-\rangle_{t, \epsilon}$ the expectation with respect to (w.r.t.) the posterior distribution (3.10), i.e.,

$$
\langle g(\mathbf{u}, \mathbf{v})\rangle_{t, \epsilon}=\int g(\mathbf{u}, \mathbf{v}) d P\left(\mathbf{u}, \mathbf{v} \mid \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)
$$

The interpolating average free entropy defined as

$$
\begin{equation*}
f_{n}(t, \epsilon):=\frac{1}{n} \mathbb{E} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right) \tag{3.12}
\end{equation*}
$$

is intimately linked to the normalized mutual information. In particular,

$$
f_{n}:=f_{n}(0,0)=\frac{n_{u} n_{v} \rho_{u} \rho_{v}}{2 n^{2}}-\frac{I(\mathbf{U}, \mathbf{V} ; \mathbf{Y})}{n} .
$$

Hence, Theorem 3.2 is equivalent to

$$
\begin{equation*}
\lim _{n \rightarrow \infty} f_{n}=\sup _{m_{u}} \inf _{m_{v}} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right), \tag{3.13}
\end{equation*}
$$

where the potential $\phi_{\mathrm{RS}}$ is defined, using $\varphi: m \in[0,+\infty) \mapsto(m-\ln (1+m)) / 2$, by

$$
\phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right):=\alpha_{u} \varphi\left(\alpha_{v} \rho_{u} m_{v}\right)+\alpha_{v} \varphi\left(\alpha_{u} \rho_{v} m_{u}\right)-\frac{\alpha_{u} \alpha_{v} m_{u} m_{v}}{2} .
$$

Looking at how $f_{n}(t, \epsilon)$ varies from $t=0$ to $t=1$ yields the following important sum-rule that we evaluate for different interpolation paths later.
Proposition 3.4. Define the scalar overlaps $Q_{u}:=\frac{\mathbf{u}^{\top} \mathbf{U}}{n_{u}}$ and $Q_{v}:=\frac{\mathbf{v}^{\top} \mathbf{V}}{n_{v}}$. Denote $R_{u}^{\prime}(\cdot, \epsilon), R_{v}^{\prime}(\cdot, \epsilon)$ the derivatives of $R_{u}(\cdot, \epsilon), R_{v}(\cdot, \epsilon)$, respectively. Assume that $R_{u}^{\prime}(t, \epsilon)$ and $R_{v}^{\prime}(t, \epsilon)$ are uniformly bounded for $(t, \epsilon) \in[0,1] \times[0,+\infty)^{2}$. Then,

$$
\begin{aligned}
f_{n}= & O(\|\epsilon\|)+o_{n}(1)+\alpha_{u} \varphi\left(\alpha_{v} \rho_{u} R_{v}(1, \epsilon)\right)+\alpha_{v} \varphi\left(\alpha_{u} \rho_{v} R_{u}(1, \epsilon)\right) \\
& -\frac{\alpha_{u} \alpha_{v}}{2} \int_{0}^{1} d t R_{u}^{\prime}(t, \epsilon) R_{v}^{\prime}(t, \epsilon)+\frac{\alpha_{u} \alpha_{v}}{2} \int_{0}^{1} d t \mathbb{E}\left\langle\left(Q_{u}-R_{u}^{\prime}(t, \epsilon)\right)\left(Q_{v}-R_{v}^{\prime}(t, \epsilon)\right)\right\rangle_{t, \epsilon},
\end{aligned}
$$

where $O_{n}(1)$ is a quantity that vanishes uniformly in $\epsilon$ as $n$ gets large, and $O(\|\epsilon\|)$ is a quantity whose absolute value is upper bounded by $C\|\epsilon\|$ for some constant $C$ independent of both $n$ and $\epsilon$.

Proof. If we evaluate (3.12) at both extremes of the interpolation, we get

$$
f_{n}(0, \epsilon)=f_{n}(0,0)+O(\|\epsilon\|)=f_{n}+O(\|\epsilon\|)
$$

and

$$
f_{n}(1, \epsilon)=\alpha_{u} \varphi\left(\alpha_{v} \rho_{u} R_{v}(1, \epsilon)\right)+\alpha_{v} \varphi\left(\alpha_{u} \rho_{v} R_{u}(1, \epsilon)\right)+o_{n}(1),
$$

where $O(\|\epsilon\|)$ and $o_{n}(1)$ are the quantities defined in the proposition. The first identity follows from $f_{n}(0, \cdot)$ being Lipschitz continuous and the second one from a direct application of Lemma 3.1, see Lemma 3.7 in Appendix 3.B for a detailed proof. We obtain the proposition by combining these identities at $t=0$ and $t=1$ with the fundamental theorem of calculus,

$$
f_{n}(0, \epsilon)=f_{n}(1, \epsilon)-\int_{0}^{1} f_{n}^{\prime}(t, \epsilon) d t
$$

where $f_{n}^{\prime}(\cdot, \epsilon)$ is the derivative of $f_{n}(\cdot, \epsilon)$. We compute $f_{n}^{\prime}(\cdot, \epsilon)$ using Gaussian integration by parts (Lemma 2.2) and the Nishimori identity (Lemma 2.1), see Lemma 3.8 in Appendix 3.B.

### 3.3.2 Interpolation paths as solutions to ODEs

To prove Theorem 3.2 , we lower bound $\lim _{\inf }^{n} f_{n}$ and upper bound $\limsup { }_{n} f_{n}$ by the same quantity $\sup _{m_{u}} \inf _{m_{v}} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right)$. To do so, we plug two different choices for $R(\cdot, \epsilon)$ in the sum-rule of Proposition 3.4. In both cases, $R(\cdot, \epsilon)$ is the solution to a first-order ordinary differential equation (ODE). We now describe these ODEs before diving further into the proofs of the matching bounds.

For $t \in[0,1]$ and $R=\left(R_{u}, R_{v}\right) \in[0,+\infty)^{2}$, consider the problem of estimating ( $\mathbf{U}, \mathbf{V}$ ) from the observations

$$
\left\{\begin{align*}
\mathbf{Y}^{(t)} & :=\sqrt{\frac{1-t}{n}} \mathbf{U}^{\top}+\mathbf{Z}  \tag{3.14}\\
\widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)} & :=\sqrt{\alpha_{v} R_{v}} \mathbf{U}+\widetilde{\mathbf{Z}} \\
\overline{\mathbf{Y}}^{\left(t, R_{u}\right)} & :=\sqrt{\alpha_{u} R_{u}} \mathbf{V}+\overline{\mathbf{Z}}
\end{align*}\right.
$$

where $\mathbf{U} \sim P_{u}, \mathbf{V} \sim P_{v}$ and all the noises $\mathbf{Z} \in \mathbb{R}^{n_{u} \times n_{v}}, \widetilde{\mathbf{Z}} \in \mathbb{R}^{n_{u}}, \overline{\mathbf{Z}} \in \mathbb{R}^{n_{v}}$ have i.i.d. entries with respect to $\mathcal{N}(0,1)$. The posterior distribution of $(\mathbf{U}, \mathbf{V})$ given $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)$ is, up to the normalization factor,

$$
\begin{equation*}
d P\left(\mathbf{u}, \mathbf{v} \mid \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right) \propto d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}_{t, R}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{(t, R u)}\right)} \tag{3.15}
\end{equation*}
$$

where $\mathcal{H}_{t, R}$ denotes the associated Hamiltonian,

$$
\begin{aligned}
\mathcal{H}_{t, R}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right):= & \sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} \frac{1-t}{2 n} u_{i}^{2} v_{j}^{2}-\sqrt{\frac{1-t}{n}} u_{i} v_{j} Y_{i j}^{(t)} \\
& +\sum_{i=1}^{n_{u}} \frac{\alpha_{v} R_{v}}{2} u_{i}^{2}-\sqrt{\alpha_{v} R_{v}} u_{i} \widetilde{Y}_{i}^{\left(t, R_{v}\right)} \\
& +\sum_{j=1}^{n_{v}} \frac{\alpha_{u} R_{u}}{2} v_{j}^{2}-\sqrt{\alpha_{u} R_{u}} v_{j} \bar{Y}_{j}^{\left(t, R_{u}\right)}
\end{aligned}
$$

The angular brackets $\langle-\rangle_{t, R}$ denote the expectation w.r.t. the posterior (3.15). Remember the definitions of the overlaps $Q_{u}$ and $Q_{v}$ in Proposition 3.4. We define

$$
F_{v}(t, R):=\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}, \quad F_{u}(t, R):=2 \rho_{u} \varphi^{\prime}\left(\alpha_{v} \rho_{u} \mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right) .
$$

Let $m_{u} \in\left[0, \rho_{u}\right]$. We consider the following first-order ODEs with initial value $\epsilon \in[0,+\infty)^{2}$ :

$$
\begin{align*}
g^{\prime} & =\left(m_{u}, F_{v}(t, g)\right), & g(0) & =\epsilon ;  \tag{3.16}\\
g^{\prime} & =\left(F_{u}(t, g), F_{v}(t, g)\right), & g(0) & =\epsilon \tag{3.17}
\end{align*}
$$

The next proposition sums up useful properties on the solutions of these two ODEs, i.e., our two kinds of interpolation paths. The proof is given in Appendix 3.C.

Proposition 3.5. For all $\epsilon \in[0,+\infty)^{2}$, there exists a unique global solution, denoted $R(\cdot, \epsilon):[0,1] \rightarrow[0,+\infty)^{2}$, to the initial value problem

$$
g^{\prime}=\left(F_{u}(t, g), F_{v}(t, g)\right), g(0)=\epsilon
$$

$R(\cdot, \epsilon)$ is continuously differentiable and the image of its derivative $R^{\prime}(\cdot, \epsilon)$ is $R^{\prime}([0,1], \epsilon) \subseteq\left[0, \rho_{u}\right] \times\left[0, \rho_{v}\right]$. Besides, for all $t \in[0,1], R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $[0,+\infty)^{2}$ into its image whose Jacobian determinant is greater than, or equal to, one;

$$
\begin{equation*}
\forall \epsilon \in[0,+\infty)^{2}: \operatorname{det} J_{R(t,)}(\epsilon) \geq 1 \tag{3.18}
\end{equation*}
$$

where $J_{R(t,)}$ denotes the Jacobian matrix of $R(t, \cdot)$. Let $m_{u} \in\left[0, \rho_{u}\right]$. The same statement holds true if we instead consider the initial value problem

$$
g^{\prime}=\left(m_{u}, F_{v}(t, g)\right), g(0)=\epsilon .
$$

### 3.3.3 Lower bound on $\liminf _{n} f_{n}$

Let $m_{u} \in\left[0, \rho_{u}\right]$ and $\epsilon:=\left(\epsilon_{u}, \epsilon_{v}\right) \in(0,+\infty)^{2}$. We choose as interpolation path the unique solution $R(\cdot, \epsilon)$ to (3.16). Then, $R_{u}^{\prime}(t, \epsilon)=m_{u}$ and $R_{v}^{\prime}(t, \epsilon)=\mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon}$. Plugging this choice in the sum-rule of Proposition 3.4 yields

$$
\begin{align*}
f_{n}=O(\|\epsilon\|)+o_{n}(1)+\phi_{\mathrm{RS}} & \left(m_{u}, \int_{0}^{1} d t R_{v}^{\prime}(1, \epsilon) ; \Theta\right) \\
& +\frac{\alpha_{u} \alpha_{v}}{2} \int_{0}^{1} d t \mathbb{E}\left\langle Q_{u}\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon}\right)\right\rangle_{t, \epsilon}, \tag{3.19}
\end{align*}
$$

where we used that $\varphi$ is Lipschitz continuous so

$$
\varphi\left(\alpha_{v} \rho_{u} R_{v}(1, \epsilon)\right)=O\left(\epsilon_{v}\right)+\varphi\left(\alpha_{v} \rho_{u} \int_{0}^{1} d t R_{v}^{\prime}(1, \epsilon)\right)
$$

and

$$
\varphi\left(\alpha_{u} \rho_{v} R_{u}(1, \epsilon)\right)=\varphi\left(\alpha_{u} \rho_{v}\left(\epsilon_{u}+m_{u}\right)\right)=O\left(\epsilon_{u}\right)+\varphi\left(\alpha_{u} \rho_{v} m_{u}\right) .
$$

We use that $\int_{0}^{1} d t R_{v}^{\prime}(1, \epsilon) \in\left[0, \rho_{v}\right]$ (see Proposition 3.4) to lower bound (3.19),

$$
\begin{equation*}
f_{n} \geq O(\|\epsilon\|)+o_{n}(1)+\inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right)+\frac{\alpha_{u} \alpha_{v}}{2} \mathcal{R}(\epsilon) \tag{3.20}
\end{equation*}
$$

where

$$
\begin{equation*}
\mathcal{R}(\epsilon):=\int_{0}^{1} d t \mathbb{E}\left\langle Q_{u}\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon}\right)\right\rangle_{t, \epsilon} . \tag{3.21}
\end{equation*}
$$

If the overlap $Q_{v}$ concentrates on its expectation then the remainder $\mathcal{R}(\epsilon)$ in the lower bound (3.20) vanishes. However, proving such concentration is only possible after integrating on a well-chosen set of "perturbation" $\epsilon$. This integration over $\epsilon$ smoothens the phase transitions that might appear for particular choices of $\epsilon$ when $n$ goes to infinity. Let $\eta$ be a positive real number and $s_{n}:=n^{-\eta}$. From now on, $\epsilon \in \mathcal{S}_{n}:=\left[s_{n}, 2 s_{n}\right]^{2}$. Integrating w.r.t. $\epsilon$ on both sides of (3.20) yields

$$
\begin{equation*}
f_{n}=\int_{\mathcal{S}_{n}} f_{n} \frac{d \epsilon}{s_{n}^{2}} \geq o_{n}(1)+\inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right)+\frac{\alpha_{u} \alpha_{v}}{2} \int_{\mathcal{S}_{n}} \mathcal{R}(\epsilon) \frac{d \epsilon}{s_{n}^{2}}, \tag{3.22}
\end{equation*}
$$

where we use that $o_{n}(1)$ in Proposition 3.4 vanishes uniformly in $\epsilon$ and that $O(\|\epsilon\|) \leq C\|\epsilon\|$ with $C$ a constant independent of $n$ and $\epsilon$ (see Proposition 3.4) so

$$
\left|\int_{\mathcal{S}_{n}} O(\|\epsilon\|) \frac{d \epsilon}{s_{n}^{2}}\right| \leq \int_{\mathcal{S}_{n}} C 2 s_{n} \frac{d \epsilon}{s_{n}^{2}}=C 2 s_{n}=o_{n}(1)
$$

By Jensen's inequality and the upper bound $\left|Q_{u}\right| \leq\|\mathbf{U}\|\|\mathbf{u}\| / n_{u}=\rho_{u}$, we have

$$
\begin{equation*}
\left|\int_{\mathcal{S}_{n}} \mathcal{R}(\epsilon) \frac{d \epsilon}{s_{n}^{2}}\right| \leq \rho_{u} \int_{0}^{1} d t \sqrt{\int_{\mathcal{S}_{n}} \frac{d \epsilon}{s_{n}^{2}} \mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon}\right)^{2}\right\rangle_{t, \epsilon}} . \tag{3.23}
\end{equation*}
$$

Next, we fix $t \in[0,1]$ and use the change of variables $\epsilon \rightarrow R:=\left(R_{u}, R_{v}\right)=R(t, \epsilon)$. The latter is justified because $R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $[0,+\infty)$ to its image (see Proposition 3.5). We get $\forall t \in[0,1]$ :

$$
\begin{align*}
& \int_{\mathcal{S}_{n}} \frac{d \epsilon}{s_{n}^{2}} \mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon_{v}}\right)^{2}\right\rangle_{t, \epsilon}=\int_{R\left(t, \mathcal{S}_{n}\right)} \frac{d R_{u} d R_{v}}{s_{n}^{2}} \frac{\mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R}}{\left|\operatorname{det} J_{R(t, \cdot)}\left(R^{-1}(t, R)\right)\right|} \\
& \leq \int_{s_{n}}^{2 s_{n}+\rho_{v}} \frac{d R_{v}}{s_{n}^{2}} \int_{s_{n}}^{2 s_{n}+\rho_{u}} d R_{u} \mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \tag{3.24}
\end{align*}
$$

The inequality follows from the integrand being nonnegative, the lower bound (3.18) for the Jacobian determinant, and $R\left(t, \mathcal{S}_{n}\right) \subseteq\left[s_{n}, 2 s_{n}+\rho_{u}\right] \times\left[s_{n}, 2 s_{n}+\rho_{v}\right]$.We now apply Proposition 3.6- an important result on the concentration of the overlap $Q_{v}$ that follows this proof - with $M_{u}=2+\rho_{u}, M_{v}=2+\rho_{v}, a=s_{n}, b=2 s_{n}+\rho_{u}$ and $\delta=s_{n} n^{\frac{2 \eta-1}{3}}$ (we further assume $\eta<1 / 2$ ). Then, for $n$ large enough, there exists $M>0$ such that $\forall t \in[0,1], \forall R_{v} \in\left[s_{n}, 2 s_{n}+\rho_{v}\right]$ :

$$
\int_{s_{n}}^{2 s_{n}+\rho_{u}} d R_{u} \mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \leq \frac{M}{n^{\frac{1-2 \eta}{3}}} .
$$

Combining this inequality with (3.23) and (3.24) yields

$$
\left|\int_{\mathcal{S}_{n}} \mathcal{R}(\epsilon) \frac{d \epsilon}{s_{n}^{2}}\right| \leq M^{\prime} n^{\frac{4 n}{3}-\frac{1}{6}}
$$

where $M^{\prime}:=\rho_{u} \sqrt{\left(1+\rho_{v}\right) M}$. This upper bound vanishes for $n$ large as long as $\eta$ is less than $1 / 8$. Then, passing to the limit inferior on both sides of (3.22) gives $\liminf _{n \rightarrow+\infty} f_{n} \geq \inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right)$. As this is true for all $m_{u} \in\left[0, \rho_{u}\right]$, we finally obtain

$$
\liminf _{n \rightarrow+\infty} f_{n} \geq \sup _{m_{u} \in\left[0, \rho_{u}\right]} \inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right)
$$

### 3.3.4 Concentration of the overlap $Q_{v}$

We rely on the following concentration result to prove the matching bounds. By symmetry, it is clear that a similar result holds for $Q_{u}$.

Proposition 3.6. Let $M_{u}$ and $M_{v}$ be positive real numbers. For $n$ large enough, there exists a constant $M$ such that $\forall(a, b) \in\left(0, M_{u}\right)^{2}: a<\min \{1, b\}, \forall \delta \in(0, a)$, $\forall R_{v} \in\left[0, M_{v}\right], \forall t \in[0,1]:$

$$
\int_{a}^{b} \mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R_{u} \leq M\left(\frac{1}{\delta^{2} n}-\frac{\ln (a)}{n}+\frac{\delta}{a-\delta}\right) .
$$

The proof is technical and is given in Appendix 3.D. It follows the same step than similar concentration results on the overlaps of inference problems [29], [37], [69, [87]. The differences with the proof in 69] are due to the entries of both $\mathbf{U}$ and $\mathbf{V}$ being not independent anymore. It mainly impacts the proof that the free entropy $\ln \mathcal{Z}_{t, R} / n$ f concentrates on its mean, which we need in our proof of the overlap concentration. We now use Lévy's lemma [99, Corollary 5.4] to show that $\ln \mathcal{Z}_{t, R} / n$ concentrates on its expectation with respect to ( $\mathbf{U}, \mathbf{V}$ ). This requires verifying that $(\widetilde{\mathbf{U}}, \widetilde{\mathbf{V}}) \mapsto \ln \mathcal{Z}_{t, R} / n$ is Lipschitz continuous with respect to $\widetilde{\mathbf{U}}:=\mathrm{U} / \sqrt{\rho_{u} n_{u}}$ on the $\left(n_{u}-1\right)$-sphere and $\widetilde{\mathbf{V}}:=\mathrm{V} / \sqrt{\rho_{v} n_{v}}$ on the $\left(n_{v}-1\right)$-sphere. The other difference is that the concentration in [69, Lemma 3.1] holds under the assumption that the prior of the i.i.d. entries of $\mathbf{V}$ is compactly supported. Here, knowing that the norm of $\mathbf{V}$ scales like $\sqrt{n}$ is in fact enough to guarantee Proposition 3.6 .

### 3.3.5 Matching upper bound on $\lim \sup _{n} f_{n}$

Let $\epsilon:=\left(\epsilon_{u}, \epsilon_{v}\right) \in(0,+\infty)^{2}$. We choose as interpolation path the unique solution $R(\cdot, \epsilon)$ to (3.17). Then, $R_{u}^{\prime}(t, \epsilon)=2 \rho_{u} \varphi^{\prime}\left(\alpha_{v} \rho_{u} \mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon}\right)$ and $R_{v}^{\prime}(t, \epsilon)=\mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon}$. Note that $\varphi: m \in[0,+\infty) \mapsto(m-\ln (1+m)) / 2$ is convex and Lipschitz continuous so

$$
\varphi\left(\alpha_{v} \rho_{u} R_{v}(1, \epsilon)\right)=O\left(\epsilon_{v}\right)+\varphi\left(\int_{0}^{1} d t \alpha_{v} \rho_{u} R_{v}^{\prime}(t, \epsilon)\right) \leq O\left(\epsilon_{v}\right)+\int_{0}^{1} d t \varphi\left(\alpha_{v} \rho_{u} R_{v}^{\prime}(t, \epsilon)\right) .
$$

A similar inequality holds for $\varphi\left(\alpha_{u} \rho_{v} R_{u}(1, \epsilon)\right)$. First writing the sum-rule of Proposition 3.4 for this specific interpolation path, and then making use of the upper bounds on $\varphi\left(\alpha_{v} \rho_{u} R_{v}(1, \epsilon)\right)$ and $\varphi\left(\alpha_{u} \rho_{v} R_{u}(1, \epsilon)\right)$, yields

$$
\begin{align*}
f_{n} \leq O(\|\epsilon\|) & +o_{n}(1)+\int_{0}^{1} d t \phi_{\mathrm{RS}}\left(R_{u}^{\prime}(t, \epsilon), R_{v}^{\prime}(t, \epsilon) ; \Theta\right) \\
& +\frac{\alpha_{u} \alpha_{v}}{2} \int_{0}^{1} d t \mathbb{E}\left\langle\left(Q_{u}-R_{u}^{\prime}(t, \epsilon)\right)\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon}\right)\right\rangle_{t, \epsilon} \tag{3.25}
\end{align*}
$$

Fix $(t, \epsilon) \in[0,1] \in(0,+\infty)^{2}$ and define $h: m_{v} \in\left[0, \rho_{v}\right] \mapsto \phi_{\mathrm{RS}}\left(R_{u}^{\prime}(t, \epsilon), m_{v} ; \Theta\right)$. As $R_{u}^{\prime}(t, \epsilon)=2 \rho_{u} \varphi^{\prime}\left(\alpha_{v} \rho_{u} R_{v}^{\prime}(t, \epsilon)\right)$, we have $h^{\prime}\left(R_{v}^{\prime}(t, \epsilon)\right)=0$ and the unique global minima of the strictly convex function $h$ is reached at $R_{v}^{\prime}(t, \epsilon) \in\left[0, \rho_{v}\right]$. Therefore,

$$
\begin{aligned}
\phi_{\mathrm{RS}}\left(R_{u}^{\prime}(t, \epsilon), R_{v}^{\prime}(t, \epsilon) ; \Theta\right) & =\inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(R_{u}^{\prime}(t, \epsilon), m_{v} ; \Theta\right) \\
& \leq \sup _{m_{u} \in\left[0, \rho_{u}\right]} \inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right) .
\end{aligned}
$$

[^15]Plugging this upper bound back in (3.25) gives

$$
\begin{align*}
& f_{n} \leq O(\|\epsilon\|)+o_{n}(1)+\sup _{m_{u} \in\left[0, \rho_{u}\right]} \inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right) \\
& \quad+\frac{\alpha_{u} \alpha_{v}}{2} \int_{0}^{1} d t \mathbb{E}\left\langle\left(Q_{u}-R_{u}^{\prime}(t, \epsilon)\right)\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon}\right)\right\rangle_{t, \epsilon} . \tag{3.26}
\end{align*}
$$

We get rid of the remainder exactly as in the proof of the lower bound on $\lim \inf _{n} f_{n}$. After integrating (3.26) over $\epsilon \in \mathcal{S}_{n}:=\left[s_{n}, 2 s_{n}\right]^{2}\left(s_{n}:=n^{-\eta}\right.$ with $\left.\eta>0\right)$, we obtain

$$
\begin{equation*}
f_{n}:=\int_{\mathcal{S}_{n}} f_{n} \frac{d \epsilon}{s_{n}^{2}} \leq o_{n}(1)+\sup _{m_{u} \in\left[0, \rho_{u}\right]} \inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right)+\frac{\alpha_{u} \alpha_{v} \mathcal{R}}{2} \tag{3.27}
\end{equation*}
$$

where $\mathcal{R}$ stands for the remainder,

$$
\mathcal{R}:=\int_{0}^{1} d t \int_{\mathcal{S}_{n}} \frac{d \epsilon}{s_{n}^{2}} \mathbb{E}\left\langle\left(Q_{u}-R_{u}^{\prime}(t, \epsilon)\right)\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon}\right)\right\rangle_{t, \epsilon}
$$

We can upper bound the absolute value of $\mathcal{R}$ by $C n^{\frac{4 \eta}{3}-\frac{1}{6}}$ for some positive constant $C$ and $n$ large enough. It is done exactly as in the proof of the lower bound. We have $\left|Q_{u}-R_{u}^{\prime}(t, \epsilon)\right| \leq 2 \rho_{u}$ and the change of variables $\epsilon \rightarrow R=R(t, \epsilon)$ is still justified by $R(t, \cdot)$ being a $C^{1}$-diffeomorphism from $[0,+\infty)$ to its image (see Proposition 3.5). As long as $\eta$ is less than $1 / 8$, the remainder vanishes when $n$ goes to infinity and passing to the limit superior on both sides of the inequality (3.27) yields the desired upper bound,

$$
\limsup _{n \rightarrow+\infty} f_{n} \leq \sup _{m_{u} \in\left[0, \rho_{u}\right]} \inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right)
$$

We have shown that

$$
\begin{aligned}
\sup _{m_{u} \in\left[0, \rho_{u}\right]} \inf _{m_{v} \in\left[0, \rho_{v}\right]} \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right) & \leq \liminf _{n \rightarrow+\infty} f_{n} \\
& \leq \limsup _{n \rightarrow+\infty} f_{n} \leq \sup _{m_{u} \in\left[0, \rho_{u}\right] m_{v} \in\left[0, \rho_{v}\right]} \inf \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right),
\end{aligned}
$$

hence $\lim _{n \rightarrow+\infty} f_{n}=\sup _{m_{u} \in\left[0, \rho_{u}\right] m_{v} \in\left[0, \rho_{v}\right]} \inf \phi_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right)$.

### 3.4 Conclusion

We conclude with a few comments on close connections with models of spin glasses. The symmetric version of the present problem can be seen to be perfectly equivalent to the spherical version of the Sherrington-Kirkpatrick spin-glass with an extra ferromagnetic interaction, on its Nishimori line. This model was introduced and solved long ago by a "spectral method" using Wigner's semicircle law [100]. Although this analysis is not completely rigorous, it can be made so (hence providing a proof of the replica formula by avoiding the replica trick entirely). For the non-symmetric inference problem considered in this paper, it is presumably
also possible to use a spectral method (using Ginibre's circle law [101]), instead of an interpolation, to arrive at the expression of the mutual information. However, it has to be noted that the interpolation method presented here readily extends to rank-one tensor problems. Indeed, the present analysis can be combined with [69] to treat the spherical tensors.

We already pointed out that the mutual informations for spherically distributed and i.i.d. Gaussian signal vectors are the same. This is perhaps not so surprising since, roughly speaking, a standard Gaussian vector concentrates on a sphere. However, this argument fails when naively applied to the spherical spin-glass model of statistical mechanics. It is well-known that the spherical and Gaussian spin-glass models are not equivalent (this goes back to [102], see [88], [103] for interesting recent developments). From this perspective, it is not obvious that in inference the two distributions lead to the same asymptotic normalized mutual information.

## Appendices

## 3.A Proof of Lemma 3.1

Let $\mathbf{X} \sim P_{x}$ a $n$-dimensional random vector uniformly distributed on the sphere of radius $\sqrt{n}$. We are interested in the normalized mutual information between $\mathbf{X}$ and $\widetilde{\mathbf{Y}}=\sqrt{m} \mathbf{X}+\widetilde{\mathbf{Z}}$ in the high-dimensional limit, where $m>0$ and the entries of the noise $\widetilde{\mathbf{Z}}$ are independent standard Gaussian random variables. We first link the normalized mutual information to the average free entropy

$$
\widetilde{f}_{n}:=\frac{1}{n} \mathbb{E} \ln \int d P_{x}(\mathbf{x}) e^{-\mathcal{H}(\mathbf{x}, \tilde{\mathbf{Y}})}
$$

where $\mathcal{H}(\mathbf{x}, \widetilde{\mathbf{Y}}):=\sum_{i=1}^{n} \frac{m}{2} x_{i}^{2}-\sqrt{m} x_{i} \widetilde{Y}_{i}$. We have

$$
\begin{align*}
\frac{I(\mathbf{X} ; \tilde{\mathbf{Y}})}{n} & :=\frac{h(\widetilde{\mathbf{Y}})}{n}-\frac{h(\widetilde{\mathbf{Y}} \mid \mathbf{X})}{n} \\
& =-\frac{\mathbb{E} \ln \int d P_{x}(\mathbf{x}) e^{-\mathcal{H}(\mathbf{x}, \tilde{\mathbf{Y}})-\frac{\|\tilde{\mathbf{Y}}\|^{2}}{2}}}{n}+\frac{\mathbb{E} \ln e^{-\frac{\|\tilde{\widetilde{Z}}\|^{2}}{2}}}{n}=\frac{m}{2}-\widetilde{f}_{n} . \tag{3.28}
\end{align*}
$$

Therefore, proving Lemma 3.1 is equivalent to proving that

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \tilde{f}_{n}=\frac{m}{2}-\frac{\ln (1+m)}{2} . \tag{3.29}
\end{equation*}
$$

We use a classical interpolation scheme to prove (3.29). For $t \in[0,1]$, consider the estimation of the $n$-dimensional standard Gaussian random vector $\widetilde{\mathbf{X}}$ from the observations

$$
\left\{\begin{align*}
\widetilde{\mathbf{Y}}^{(t)} & :=\sqrt{m(1-t) n} \frac{\widetilde{\mathbf{X}}}{\|\tilde{\mathbf{X}}\|}+\widetilde{\mathbf{Z}}  \tag{3.30}\\
\mathbf{Y}^{(t)} & :=\sqrt{m t} \widetilde{\mathbf{X}}+\mathbf{Z}
\end{align*}\right.
$$

where the noises $\mathbf{Z} \in \mathbb{R}^{n}, \widetilde{\mathbf{Z}} \in \mathbb{R}^{n}$ have i.i.d. entries with respect to $\mathcal{N}(0,1)$. The associated interpolating Hamiltonian is

$$
\begin{align*}
& \mathcal{H}_{t}\left(\widetilde{\mathbf{x}} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right):=\sum_{i=1}^{n} \frac{m(1-t) n}{2} \frac{\widetilde{x}_{i}^{2}}{\|\widetilde{\mathbf{x}}\|^{2}}-\sqrt{m(1-t) n} \frac{\widetilde{x}_{i}}{\|\widetilde{\mathbf{x}}\|} \widetilde{Y}_{i}^{(t)} \\
&+\sum_{i=1}^{n} \frac{m t}{2} \widetilde{x}_{i}^{2}-\sqrt{m t} \widetilde{x}_{i} Y_{i}^{(t)} \tag{3.31}
\end{align*}
$$

Define the interpolating free entropy $\widetilde{f}_{n}(t):=\frac{1}{n} \mathbb{E} \ln \mathcal{Z}_{t}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right)$ where

$$
\begin{equation*}
\mathcal{Z}_{t}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right):=\int \frac{d \widetilde{\mathbf{x}}}{\sqrt{2 \pi}^{n}} e^{-\frac{\|\widetilde{\|}\|^{2}}{2}} e^{-\mathcal{H}_{t}\left(\widetilde{\mathbf{x}} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right)} \tag{3.32}
\end{equation*}
$$

Note that $\sqrt{n} \frac{\widetilde{\mathbf{X}}}{\|\tilde{\mathbf{X}}\|}$ has the same distribution than $\mathbf{X}$, i.e., it is uniformly distributed on the $(n-1)$-sphere of radius $\sqrt{n}$. Then, the interpolating free entropy at $t=0$ is equal to to the average free entropy $\widetilde{f}_{n}$ whose limit we want to compute, $\widetilde{f}_{n}(0)=\widetilde{f}_{n}$. At $t=1$, the integral (3.32) is a simple Gaussian integral and we find that $\widetilde{f}_{n}(1)=\frac{m}{2}-\frac{\ln (1+m)}{2}$. Hence, we have

$$
\begin{equation*}
\left|\frac{m}{2}-\frac{\ln (1+m)}{2}-\widetilde{f}_{n}\right|=\left|\int_{0}^{1} \widetilde{f}_{n}^{\prime}(t) d t\right| \leq \int_{0}^{1}\left|\widetilde{f}_{n}^{\prime}(t)\right| d t \tag{3.33}
\end{equation*}
$$

Computing $\widetilde{f}_{n}^{\prime}(t)$ is done much like in the proof of Lemma 3.8 where we compute the derivative of the average free entropy (3.12). We obtain

$$
\begin{equation*}
\widetilde{f}_{n}^{\prime}(t)=\frac{m}{2} \mathbb{E}\left\langle\frac{\widetilde{\mathbf{x}}^{\top} \widetilde{\mathbf{X}}}{\|\widetilde{\mathbf{x}}\|\|\widetilde{\mathbf{X}}\|}\left(\frac{\|\widetilde{\mathbf{x}}\|\| \| \widetilde{\mathbf{X}} \|}{n}-1\right)\right\rangle_{t} \tag{3.34}
\end{equation*}
$$

where the angular brackets $\langle-\rangle_{t}$ denote the expectation w.r.t. the posterior distribution of $\widetilde{\mathbf{X}}$ given $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right)$,

$$
d P\left(\widetilde{\mathbf{x}} \mid \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right)=\frac{1}{\mathcal{Z}_{t}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right)} \frac{d \widetilde{\mathbf{x}}}{\sqrt{2 \pi}^{n}} e^{-\frac{\|\tilde{\mathbf{x}}\|^{2}}{2}} e^{-\mathcal{H}_{t}\left(\widetilde{\mathbf{x}} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t)}\right)}
$$

We split $\widetilde{f}_{n}^{\prime}(t)$ in two pieces,

$$
f_{n}^{\prime}(t)=\frac{m}{2} \mathbb{E}\left\langle\frac{\widetilde{\mathbf{x}}^{\top} \widetilde{\mathbf{X}}}{\|\widetilde{\mathbf{x}}\|\|\widetilde{\mathbf{X}}\|} \frac{\|\widetilde{\mathbf{x}}\|}{\sqrt{n}}\left(\frac{\|\widetilde{\mathbf{X}}\|}{\sqrt{n}}-1\right)\right\rangle_{t}+\frac{m}{2} \mathbb{E}\left\langle\frac{\widetilde{\mathbf{x}}^{\top} \widetilde{\mathbf{X}}}{\|\widetilde{\mathbf{x}}\|\|\widetilde{\mathbf{X}}\|}\left(\frac{\|\widetilde{\mathbf{x}}\|}{\sqrt{n}}-1\right)\right\rangle_{t}
$$

Applying Cauchy-Schwarz inequality separately to these two expectations, we get

$$
\begin{align*}
\left|\widetilde{f_{n}^{\prime}}(t)\right| & \leq \frac{m}{2} \sqrt{\mathbb{E}\left\langle\frac{\left(\widetilde{\mathbf{x}}^{\top} \widetilde{\mathbf{X}}\right)^{2}}{\|\widetilde{\mathbf{x}}\|^{2}\|\widetilde{\mathbf{X}}\|^{2}} \frac{\|\widetilde{\mathbf{x}}\|^{2}}{n}\right\rangle_{t} \mathbb{E}\left[\left(\frac{\|\widetilde{\mathbf{X}}\|}{\sqrt{n}}-1\right)^{2}\right]} \\
& +\frac{m}{2} \sqrt{\mathbb{E}\left\langle\frac{\left(\widetilde{\mathbf{x}}^{\top} \widetilde{\mathbf{X}}\right)^{2}}{\left.\| \widetilde{\widetilde{\mathbf{x}}\left\|^{2}\right\| \widetilde{\mathbf{X}} \|^{2}}\right\rangle_{t}^{\mathbb{E}}\left\langle\left(\frac{\|\widetilde{\mathbf{x}}\|}{\sqrt{n}}-1\right)^{2}\right\rangle_{t}}\right.} \\
& \leq \frac{m}{2} \sqrt{\mathbb{E}\left\langle\frac{\|\widetilde{\mathbf{x}}\|^{2}}{n}\right\rangle_{t} \mathbb{E}\left[\left(\frac{\|\widetilde{\mathbf{X}}\|}{\sqrt{n}}-1\right)^{2}\right]+\frac{m}{2} \sqrt{\mathbb{E}\left\langle\left(\frac{\|\widetilde{\mathbf{x}}\|}{\sqrt{n}}-1\right)^{2}\right\rangle_{t}}} \\
& =m \sqrt{\mathbb{E}\left[\left(\frac{\|\widetilde{\mathbf{X}}\|}{\sqrt{n}}-1\right)^{2}\right]} . \tag{3.35}
\end{align*}
$$

The second inequality follows from $\left|\widetilde{\mathbf{x}}^{\top} \widetilde{\mathbf{X}}\right| \leq\|\widetilde{\mathbf{x}}\|\|\widetilde{\mathbf{X}}\|$ by Cauchy-Schwarz inequality.
The subsequent equality is an application of the Nishimori identity (Lemma 2.1),

$$
\mathbb{E}\left\langle\frac{\|\widetilde{\mathbf{x}}\|^{2}}{n}\right\rangle_{t}=\mathbb{E}\left[\frac{\|\widetilde{\mathbf{X}}\|^{2}}{n}\right]=1 \quad \text { and } \quad \mathbb{E}\left\langle\left(\frac{\|\widetilde{\mathbf{x}}\|}{\sqrt{n}}-1\right)^{2}\right\rangle_{t}=\mathbb{E}\left[\left(\frac{\|\widetilde{\mathbf{X}}\|}{\sqrt{n}}-1\right)^{2}\right]
$$

The upper bound $(3.35)$ on the absolute value of the derivative of the interpolating free entropy is valid for all $t \in[0,1]$. Plugging it back in (3.33) gives

$$
\begin{equation*}
\left|\frac{m}{2}-\frac{\ln (1+m)}{2}-\tilde{f}_{n}\right| \leq \frac{m}{\sqrt{n}} \sqrt{\mathbb{E}\left[(\|\widetilde{\mathbf{X}}\|-\sqrt{n})^{2}\right]} . \tag{3.36}
\end{equation*}
$$

There exists a constant $C$ such that $\mathbb{P}(|\|\widetilde{\mathbf{X}}\|-\sqrt{n}| \geq a) \leq 2 e^{-C a^{2}}$ for all $a \geq 0$ (see $\sqrt[99]{ }$, Theorem 3.1.1]). This directly implies that $\mathbb{E}\left[(\|\widetilde{\mathbf{X}}\|-\sqrt{n})^{2}\right] \leq 2 / C$, thus concluding the proof of (3.29).

## 3.B Sum-rule of Proposition 3.4

Remember that, without loss of generality, we assume that $\lambda=1$.
Lemma 3.7 (Average interpolating free entropy at $t=0$ and $t=1$ ). Define the scalar overlaps $Q_{u}:=\frac{\mathbf{u}^{\top} \mathbf{U}}{n_{u}}$ and $Q_{v}:=\frac{\mathbf{v}^{\top} \mathbf{U}}{n_{v}}$. Assume that both $R_{u}(t, \epsilon)$ and $R_{v}(t, \epsilon)$ are uniformly bounded in $(t, \epsilon) \in[0,1] \times[0,+\infty)^{2}$. The average interpolating free entropy $f_{n}(t, \epsilon)$ defined by (3.12) satisfies

$$
\begin{align*}
& f_{n}(0, \epsilon)=f_{n}(0,0)+O(\|\epsilon\|) ;  \tag{3.37}\\
& f_{n}(1, \epsilon)=\alpha_{u} \varphi\left(\alpha_{v} \rho_{u} R_{v}(1, \epsilon)\right)+\alpha_{v} \varphi\left(\alpha_{u} \rho_{v} R_{u}(1, \epsilon)\right)+o_{n}(1) ; \tag{3.38}
\end{align*}
$$

where $O_{n}(1)$ is a quantity that vanishes uniformly in $\epsilon$ as $n$ gets large, and $O(\|\epsilon\|)$ is a quantity whose absolute value is upper bounded by $C\|\epsilon\|$ for some constant $C$ independent of both $n$ and $\epsilon$.

Proof. By definition, $f_{n}(0, \epsilon):=\frac{1}{n} \mathbb{E} \ln \mathcal{Z}_{0, \epsilon}\left(\mathbf{Y}^{(0)}, \widetilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}\right.$ ) where

$$
\mathcal{Z}_{0, \epsilon}\left(\mathbf{Y}^{(0)}, \tilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}\right):=\int d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}_{0, \epsilon}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(0)}, \tilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}\right)}
$$

$\mathcal{H}_{0, \epsilon}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(0)}, \widetilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}\right.$ ) being the Hamiltonian (3.11) evaluated at $t=0$. Remembering that $R_{u}(0, \epsilon)=\epsilon_{u}, R_{v}(0, \epsilon)=\epsilon_{v}$, and replacing $\mathbf{Y}^{(0)}, \widetilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}$ by their expressions on the right-hand side of (3.9), we obtain

$$
\begin{equation*}
\mathcal{Z}_{0, \epsilon}\left(\mathbf{Y}^{(0)}, \tilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}\right)=\int d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{v}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})} \tag{3.39}
\end{equation*}
$$

where

$$
\begin{aligned}
& \mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}):= \sum_{i=1}^{n_{u}} \\
& \sum_{j=1}^{n_{v}} \frac{u_{i}^{2} v_{j}^{2}}{2 n}-\frac{1}{n} u_{i} U_{i} v_{j} V_{j}-\frac{u_{i} v_{j} Z_{i j}}{\sqrt{n}} \\
&+\sum_{i=1}^{n_{u}} \frac{\alpha_{v} \epsilon_{v}}{2} u_{i}^{2}-\alpha_{v} \epsilon_{v} u_{i} U_{i}-\sqrt{\alpha_{v} \epsilon_{v}} u_{i} \widetilde{Z}_{i} \\
&+\sum_{j=1}^{n_{v}} \frac{\alpha_{u} \epsilon_{u}}{2} v_{j}^{2}-\alpha_{u} \epsilon_{u} v_{j} V_{j}-\sqrt{\alpha_{u} \epsilon_{u}} v_{j} \bar{Z}_{j} .
\end{aligned}
$$

Using (3.39), the partial derivative of $\epsilon:=\left(\epsilon_{u}, \epsilon_{v}\right) \mapsto f_{n}(0, \epsilon)$ with respect to $\epsilon_{u}$ reads

$$
\begin{align*}
\left.\frac{\partial f_{n}}{\partial \epsilon_{u}}\right|_{t=0, \epsilon} & =-\frac{1}{n} \mathbb{E}\left[\frac{\int d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) \frac{\partial \mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})}{\partial \epsilon_{u}} e^{-\mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})}}{\int d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}}{ }_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})}\right. \\
& =-\frac{1}{n} \mathbb{E}\left\langle\frac{\partial \mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})}{\partial \epsilon_{u}}\right\rangle_{t=0, \epsilon} \\
& =-\frac{\alpha_{u}}{2 n} \sum_{j=1}^{n_{v}} \mathbb{E}\left\langle v_{j}^{2}\right\rangle_{0, \epsilon}+\frac{\alpha_{u}}{n} \sum_{j=1}^{n_{v}} \mathbb{E}\left\langle v_{j} V_{j}\right\rangle_{0, \epsilon}+\frac{1}{2 n} \sqrt{\frac{\alpha_{u}}{\epsilon_{u}}} \sum_{j=1}^{n_{v}} \mathbb{E}\left\langle v_{j} \bar{Z}_{j}\right\rangle_{0, \epsilon} . \tag{3.40}
\end{align*}
$$

We now simplify the expectation $\mathbb{E}\left\langle v_{j} \bar{Z}_{j}\right\rangle_{0, \epsilon}$ by integrating by parts with respect to the standard Gaussian random variable $\bar{Z}_{j}$ (see Lemma 2.2,

$$
\begin{align*}
& \mathbb{E}\left\langle v_{j} \bar{Z}_{j}\right\rangle_{0, \epsilon}=\mathbb{E}\left[\bar{Z}_{j} \frac{\int d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) v_{j} e^{-\mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{v}, \mathbf{Z}, \tilde{\mathbf{Z}}, \overline{\mathbf{Z}})}}{\mathcal{Z}_{0, \epsilon}\left(\mathbf{Y}^{(0)}, \widetilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}\right)}\right] \\
& =-\mathbb{E}\left[\frac{\int d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) v_{j} e^{-\mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \tilde{\mathbf{Z}}, \overline{\mathbf{Z}})} \frac{\partial \mathcal{Z}_{0, \epsilon}\left(\mathbf{Y}^{(0)}, \tilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}\right)}{\partial \bar{Z}_{j}}}{\left(\mathcal{Z}_{0, \epsilon}\left(\mathbf{Y}^{(0)}, \widetilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}\right)\right)^{2}}\right] \\
& -\mathbb{E}\left[\frac{\int d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) v_{j} \frac{\partial \mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \tilde{\mathbf{Z}}, \overline{\mathbf{Z}})}{\partial \bar{Z}_{j}} e^{-\mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \tilde{z}, \overline{\mathbf{Z}})}}{\mathcal{Z}_{0, \epsilon}\left(\mathbf{Y}^{(0)}, \widetilde{\mathbf{Y}}^{(0, \epsilon)}, \overline{\mathbf{Y}}^{(0, \epsilon)}\right)}\right] \\
& =\mathbb{E}\left[\left\langle v_{j}\right\rangle_{0, \epsilon}\left\langle\frac{\partial \mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})}{\partial \bar{Z}_{j}}\right\rangle_{0, \epsilon}\right] \\
& -\mathbb{E}\left[\left\langle v_{j} \frac{\partial \mathcal{H}_{0, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})}{\partial \bar{Z}_{j}}\right\rangle_{0, \epsilon}\right] \\
& =-\sqrt{\alpha_{u} \epsilon_{u}} \mathbb{E}\left[\left\langle v_{j}\right\rangle_{0, \epsilon}^{2}\right]+\sqrt{\alpha_{u} \epsilon_{u}} \mathbb{E}\left[\left\langle v_{j}^{2}\right\rangle_{0, \epsilon}\right] . \tag{3.41}
\end{align*}
$$

Plugging (3.41) back in (3.40) yields

$$
\begin{equation*}
\left.\frac{\partial f_{n}}{\partial \epsilon_{u}}\right|_{t=0, \epsilon}=\frac{\alpha_{u}}{n} \sum_{j=1}^{n_{v}} \mathbb{E}\left\langle v_{j} V_{j}\right\rangle_{0, \epsilon}-\frac{\alpha_{u}}{2 n} \sum_{j=1}^{n_{v}} \mathbb{E}\left[\left\langle v_{j}\right\rangle_{0, \epsilon}^{2}\right]=\frac{\alpha_{u}}{2} \frac{n_{v}}{n} \mathbb{E}\left\langle Q_{v}\right\rangle_{0, \epsilon}, \tag{3.42}
\end{equation*}
$$

where the second equality is due to the Nishimori identity $\mathbb{E}\left\langle v_{j}\right\rangle_{0, \epsilon}^{2}=\mathbb{E}\left[\left\langle v_{j}\right\rangle_{0, \epsilon} V_{j}\right]$ (see Lemma 2.1). We have just shown that $\forall \epsilon \in[0,+\infty)^{2}$ :

$$
\left.\frac{\partial f_{n}}{\partial \epsilon_{u}}\right|_{t=0, \epsilon}=\frac{\alpha_{u}}{2} \frac{n_{v}}{n} \mathbb{E}\left\langle Q_{v}\right\rangle_{0, \epsilon},
$$

and we can similarly prove that $\left.\frac{\partial f_{n}}{\partial \epsilon_{v}}\right|_{t=0, \epsilon}=\frac{\alpha_{v}}{2} \frac{n_{u}}{n} \mathbb{E}\left\langle Q_{u}\right\rangle_{0, \epsilon}$. By Cauchy-Schwarz inequality, $\left|Q_{u}\right| \leq\|\mathbf{u}\|\|\mathbf{U}\| / n_{u}=\rho_{u}$ and $\left|Q_{v}\right| \leq\|\mathbf{v}\|\|\mathbf{V}\| / n_{v}=\rho_{v}$. Therefore,

$$
\left.\left|\frac{\partial f_{n}}{\partial \epsilon_{u}}\right|_{0, \epsilon} \right\rvert\, \leq \frac{\alpha_{u}}{2} \frac{n_{v}}{n} \rho_{v} \quad \text { and } \left.\quad\left|\frac{\partial f_{n}}{\partial \epsilon_{v}}\right|_{0,\left(0, \epsilon_{v}\right)} \right\rvert\, \leq \frac{\alpha_{v}}{2} \frac{n_{u}}{n} \rho_{u} .
$$

By the mean-value theorem, it follows that

$$
\begin{aligned}
\left|f_{n}(0, \epsilon)-f_{n}(0,0)\right| & \leq\left|f_{n}(0, \epsilon)-f_{n}\left(0,\left(0, \epsilon_{v}\right)\right)\right|+\left|f_{n}\left(0,\left(0, \epsilon_{v}\right)\right)-f_{n}(0,0)\right| \\
& \leq \frac{\alpha_{u} \rho_{v}}{2} \frac{n_{v}}{n}\left|\epsilon_{u}\right|+\frac{\alpha_{v} \rho_{u}}{2} \frac{n_{u}}{n}\left|\epsilon_{v}\right| .
\end{aligned}
$$

This last upper bound concludes the proof of (3.37) as $\left(n_{u} / n_{n}, n_{v} / n\right) \rightarrow\left(\alpha_{u}, \alpha_{v}\right)$.
At $t=1, \mathbf{Y}^{(1)}:=\mathbf{Z}$ is pure noise while $\widetilde{\mathbf{Y}^{(1, \epsilon)}}:=\sqrt{\alpha_{v} R_{v}(1, \epsilon)} \mathbf{U}+\widetilde{\mathbf{Z}}$ and $\overline{\mathbf{Y}}^{(1, \epsilon)}:=\sqrt{\alpha_{u} R_{u}(1, \epsilon)} \mathbf{V}+\overline{\mathbf{Z}}$ are two decoupled channels similar to the one described in Lemma 3.1. Therefore, we have $f_{n}(1, \epsilon)=\frac{n_{u}}{n} \widetilde{f}_{n_{u}}+\frac{n_{v}}{n} \widetilde{f}_{n_{v}}$ where

$$
\widetilde{f}_{n_{u}}:=\frac{1}{n_{u}} \mathbb{E} \ln \int d P_{u}(\mathbf{u}) e^{-\sum_{i=1}^{n_{u}} \frac{\alpha_{v} R_{v}(1, \epsilon)}{2} u_{i}^{2}-\sqrt{\alpha_{v} R_{v}(1, \epsilon)} u_{i} \tilde{Y}_{i}^{(1, \epsilon)}}
$$

and

$$
\tilde{f}_{n_{v}}:=\frac{1}{n_{v}} \mathbb{E} \ln \int d P_{v}(\mathbf{v}) e^{-\sum_{i=1}^{n_{v}} \frac{\alpha_{u} R u(1, \epsilon)}{2} v_{i}^{2}-\sqrt{\alpha_{u} R_{u}(1, \epsilon)} v_{i} \bar{Y}_{i}^{(1, \epsilon)}}
$$

are the average free entropies associated with the two aforementioned channels. In the proof of Lemma 3.1, we ultimately show that

$$
\left|\widetilde{f}_{n_{u}}-\varphi\left(\alpha_{v} \rho_{u} R_{v}(1, \epsilon)\right)\right| \leq \frac{2 \alpha_{v} R_{v}(1, \epsilon)}{C \sqrt{n_{u}}}
$$

and

$$
\left|\widetilde{f}_{n_{v}}-\varphi\left(\alpha_{u} \rho_{v} R_{u}(1, \epsilon)\right)\right| \leq \frac{2 \alpha_{u} R_{u}(1, \epsilon)}{C \sqrt{n_{v}}}
$$

where $C$ is a constant independent of $\epsilon$. These two upper bounds together with the assumption on the uniform boundedness of $R_{u}, R_{v}$ imply (3.38),

$$
f_{n}(1, \epsilon)=\frac{n_{u}}{n} \widetilde{f}_{n_{u}}+\frac{n_{v}}{n} \widetilde{f}_{n_{v}}=\alpha_{u} \varphi\left(\alpha_{v} \rho_{u} R_{v}(1, \epsilon)\right)+\alpha_{v} \varphi\left(\alpha_{u} \rho_{v} R_{u}(1, \epsilon)\right)+o_{n}(1) .
$$

Lemma 3.8 (Derivative of the average interpolating free entropy). Define the scalar overlaps $Q_{u}:=\frac{\mathbf{u}^{\top} \mathbf{U}}{n_{u}}$ and $Q_{v}:=\frac{\mathbf{v}^{\top} \mathbf{U}}{n_{v}}$. Let $R_{u}^{\prime}(\cdot, \epsilon)$ and $R_{v}^{\prime}(\cdot, \epsilon)$ be the derivatives of $R_{u}(\cdot, \epsilon)$ and $R_{v}(\cdot, \epsilon)$, respectively. Assume that both $R_{u}^{\prime}(t, \epsilon)$ and $R_{v}^{\prime}(t, \epsilon)$ are uniformly bounded for $(t, \epsilon) \in[0,1] \times[0,+\infty)^{2}$. Then, the partial derivative of the average interpolating free entropy (3.12) with respect to $t$ is $\forall(t, \epsilon) \in[0,1] \times[0,+\infty)^{2}:$
$f_{n}^{\prime}(t, \epsilon)=-\frac{\alpha_{u} \alpha_{v}}{2} \mathbb{E}\left\langle\left(Q_{u}-R_{u}^{\prime}(t, \epsilon)\right)\left(Q_{v}-R_{v}^{\prime}(t, \epsilon)\right)\right\rangle_{t, \epsilon}+\frac{\alpha_{u} \alpha_{v}}{2} R_{u}^{\prime}(t, \epsilon) R_{v}^{\prime}(t, \epsilon)+o_{n}(1)$
where $o_{n}(1)$ vanishes uniformly in $(t, \epsilon)$ as $n$ goes to infinity.
Proof. The conditional probability density function of $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)$ given $(\mathbf{U}, \mathbf{V})$ is

$$
P_{\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right) \mid(\mathbf{U}, \mathbf{V})}(\mathbf{y}, \widetilde{\mathbf{y}}, \overline{\mathbf{y}} \mid \mathbf{u}, \mathbf{v}):=\frac{e^{-\frac{\|\mathbf{y}\|^{2}+\|\tilde{\mathbf{y}}\|^{2}+\|\tilde{y}\|^{2}}{2}-\mathcal{H}_{t, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{y}, \widetilde{\mathbf{y}}, \overline{\mathbf{y}})}}{{\sqrt{2 \pi^{n} n_{v}+n_{u}+n_{v}}}^{n^{2}}}
$$

High-dimensional rank-one nonsymmetric matrix estimation: the spherical 60
where

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{y}, \widetilde{\mathbf{y}}, \overline{\mathbf{y}}):=\sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} & \frac{1-t}{2 n} u_{i}^{2} v_{j}^{2}-\sqrt{\frac{1-t}{n}} u_{i} v_{j} y_{i j} \\
& +\sum_{i=1}^{n_{u}} \frac{\alpha_{v} R_{v}(t, \epsilon)}{2} u_{i}^{2}-\sqrt{\alpha_{v} R_{v}(t, \epsilon)} u_{i} \widetilde{y}_{i} \\
& +\sum_{j=1}^{n_{v}} \frac{\alpha_{u} R_{u}(t, \epsilon)}{2} v_{j}^{2}-\sqrt{\alpha_{u} R_{u}(t, \epsilon)} v_{j} \bar{y}_{j} \tag{3.43}
\end{align*}
$$

Therefore, the average interpolating free entropy (3.12) satisfies

$$
\begin{align*}
f_{n}(t, \epsilon) & =\frac{1}{n} \mathbb{E}\left[\mathbb{E}\left[\ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right) \mid \mathbf{U}, \mathbf{V}\right]\right] \\
& =\frac{1}{n} \mathbb{E}\left[\int d \mathbf{y} d \widetilde{\mathbf{y}} d \overline{\mathbf{y}} \frac{e^{-\frac{\|\mathbf{y}\|^{2}+\|\tilde{\mathbf{y}}\|^{2}+\|\overline{\mathbf{y}}\|^{2}}{2}-\mathcal{H}_{t, \epsilon}(\mathbf{U}, \mathbf{V} ; \mathbf{y}, \tilde{\mathbf{y}}, \overline{\mathbf{y}})}}{\sqrt{2 \pi}^{n_{u} n_{v}+n_{u}+n_{v}}} \ln \mathcal{Z}_{t, \epsilon}(\mathbf{y}, \widetilde{\mathbf{y}}, \overline{\mathbf{y}})\right] . \tag{3.44}
\end{align*}
$$

where we remind that $\mathcal{Z}_{t, \epsilon}(\mathbf{y}, \widetilde{\mathbf{y}}, \overline{\mathbf{y}}):=\int d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}_{t, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{y}, \tilde{\mathbf{y}}, \overline{\mathbf{y}})}$. Taking the derivative of (3.44) with respect to $t$, we directly obtain that

$$
\begin{align*}
f_{n}^{\prime}(t, \epsilon)= & -\frac{1}{n} \mathbb{E}\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{U}, \mathbf{V} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right] \\
& -\frac{1}{n} \mathbb{E}\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right\rangle_{t, \epsilon} \tag{3.45}
\end{align*}
$$

where

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}^{\prime}(\mathbf{u}, \mathbf{v} ; \mathbf{y}, \widetilde{\mathbf{y}}, \overline{\mathbf{y}}):= & \frac{\partial \mathcal{H}_{t, \epsilon}(\mathbf{u}, \mathbf{v} ; \mathbf{y}, \widetilde{\mathbf{y}}, \overline{\mathbf{y}})}{\partial t} \\
= & \sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}}-\frac{u_{i}^{2} v_{j}^{2}}{2 n}+\frac{u_{i} v_{j} y_{i j}^{(t)}}{2 \sqrt{n(1-t)}} \\
& +\sum_{i=1}^{n_{u}} \frac{\alpha_{v} R_{v}^{\prime}(t, \epsilon)}{2} u_{i}^{2}-\frac{R_{v}^{\prime}(t, \epsilon)}{2} \sqrt{\frac{\alpha_{v}}{R_{v}(t, \epsilon)}} u_{i} \widetilde{y}_{i}^{(t, \epsilon)} \\
& +\sum_{j=1}^{n_{v}} \frac{\alpha_{u} R_{u}^{\prime}(t, \epsilon)}{2} v_{j}^{2}-\frac{R_{u}^{\prime}(t, \epsilon)}{2} \sqrt{\frac{\alpha_{u}}{R_{u}(t, \epsilon)}} v_{j} \bar{y}_{j}^{(t, \epsilon)} . \tag{3.46}
\end{align*}
$$

If we evaluate (3.46) at $(\mathbf{u}, \mathbf{v}, \mathbf{y}, \widetilde{\mathbf{y}}, \overline{\mathbf{y}})=\left(\mathbf{U}, \mathbf{V}, \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)$, we get

$$
\begin{align*}
& \mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{U}, \mathbf{V} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)=\sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} \frac{U_{i} V_{j} Z_{i j}}{2 \sqrt{n(1-t)}} \\
& \quad-\sum_{i=1}^{n_{u}} \frac{R_{v}^{\prime}(t, \epsilon)}{2} \sqrt{\frac{\alpha_{v}}{R_{v}(t, \epsilon)}} U_{i} \widetilde{Z}_{i}-\sum_{j=1}^{n_{v}} \frac{R_{u}^{\prime}(t, \epsilon)}{2} \sqrt{\frac{\alpha_{u}}{R_{u}(t, \epsilon)}} V_{j} \bar{Z}_{j} . \tag{3.47}
\end{align*}
$$

Note that $\mathbb{E}\left[U_{i} V_{j} Z_{i j}\right]=\mathbb{E}\left[U_{i} \widetilde{Z}_{i}\right]=\mathbb{E}\left[V_{j} \bar{Z}_{j}\right]=0$ for all $i, j$. The second expectation on the right-hand side of (3.45) is then easily shown to be zero thanks to the Nishimori identity (see Lemma 2.1),

$$
\mathbb{E}\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right\rangle_{t, \epsilon}=\mathbb{E}\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{U}, \mathbf{V} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right]=0
$$

Therefore, the expression for $f_{n}^{\prime}(t, \epsilon)$ simplifies to

$$
\begin{align*}
f_{n}^{\prime}(t, \epsilon)= & -\frac{1}{n} \mathbb{E}\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{U}, \mathbf{V} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right] \\
=- & \sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} \frac{\mathbb{E}\left[U_{i} V_{j} Z_{i j} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right]}{2 n \sqrt{n(1-t)}} \\
& +\sum_{i=1}^{n_{u}} \frac{R_{v}^{\prime}(t, \epsilon)}{2 n} \sqrt{\frac{\alpha_{v}}{R_{v}(t, \epsilon)}} \mathbb{E}\left[U_{i} \widetilde{Z}_{i} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right] \\
& +\sum_{j=1}^{n_{v}} \frac{R_{u}^{\prime}(t, \epsilon)}{2 n} \sqrt{\frac{\alpha_{u}}{R_{u}(t, \epsilon)}} \mathbb{E}\left[V_{j} \bar{Z}_{j} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right] . \tag{3.48}
\end{align*}
$$

We can simplify the three kind of expectations on the right-hand side of (3.48) with Gaussian integration by parts w.r.t. $Z_{i j}, \widetilde{Z}_{i}$ or $\bar{Z}_{j}$. We have

$$
\begin{aligned}
& \mathbb{E}\left[U_{i} V_{j} Z_{i j} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right]=\mathbb{E}\left[U_{i} V_{j} \frac{\partial \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)}{\partial Z_{i j}}\right] \\
& \quad=-\mathbb{E}\left[U_{i} V_{j}\left\langle\frac{\partial \mathcal{H}_{t, \epsilon}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)}{\partial Z_{i j}}\right\rangle_{t, \epsilon}\right]=\sqrt{\frac{1-t}{n}} \mathbb{E}\left\langle u_{i} U_{i} v_{j} V_{j}\right\rangle_{t, \epsilon} .
\end{aligned}
$$

In a similar way,

$$
\mathbb{E}\left[U_{i} \widetilde{Z}_{i} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right]=\sqrt{\alpha_{v} R_{v}(t, \epsilon)} \mathbb{E}\left\langle u_{i} U_{i}\right\rangle_{t, \epsilon}
$$

and

$$
\mathbb{E}\left[V_{j} \bar{Z}_{j} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \overline{\mathbf{Y}}^{(t, \epsilon)}\right)\right]=\sqrt{\alpha_{u} R_{u}(t, \epsilon)} \mathbb{E}\left\langle v_{j} V_{j}\right\rangle_{t, \epsilon} .
$$

Hence, we have

$$
\begin{equation*}
f_{n}^{\prime}(t, \epsilon)=-\frac{1}{2} \frac{n_{u}}{n} \frac{n_{v}}{n} \mathbb{E}\left\langle Q_{u} Q_{v}\right\rangle_{t, \epsilon}+\frac{n_{u}}{n} \frac{\alpha_{v} R_{v}^{\prime}(t, \epsilon)}{2} \mathbb{E}\left\langle Q_{u}\right\rangle_{t, \epsilon}+\frac{n_{v}}{n} \frac{\alpha_{u} R_{u}^{\prime}(t, \epsilon)}{2} \mathbb{E}\left\langle Q_{v}\right\rangle_{t, \epsilon} . \tag{3.49}
\end{equation*}
$$

Remember that $\left(n_{u} / n_{n}, n_{v} / n\right) \rightarrow\left(\alpha_{u}, \alpha_{v}\right)$. Besides, by Cauchy-Schwarz inequality, $\left|Q_{u}\right| \leq\|\mathbf{u}\|\|\mathbf{U}\| / n_{u}=\rho_{u}$ and $\left|Q_{v}\right| \leq\|\mathbf{v}\|\|\mathbf{V}\| / n_{v}=\rho_{v}$. It follows that
$f_{n}^{\prime}(t, \epsilon)=-\frac{\alpha_{u} \alpha_{v}}{2} \mathbb{E}\left\langle\left(Q_{u}-R_{u}^{\prime}(t, \epsilon)\right)\left(Q_{v}-R_{v}^{\prime}(t, \epsilon)\right)\right\rangle_{t, \epsilon}+\frac{\alpha_{u} \alpha_{v}}{2} R_{u}^{\prime}(t, \epsilon) R_{v}^{\prime}(t, \epsilon)+o_{n}(1)$
where $o_{n}(1)$ is a quantity that vanishes uniformly for $(t, \epsilon) \in[0,1] \times[0,+\infty)^{2}$ when $n \rightarrow+\infty$.

## 3.C Properties of the interpolation paths

This appendix is dedicated to the proof of Proposition 3.5 in Section 3.3. Before giving the proof, let us recall a few definitions. For $t \in[0,1]$ and $R:=\left(R_{u}, R_{v}\right) \in$ $[0,+\infty)^{2}$, consider the problem of estimating ( $\mathbf{U}, \mathbf{V}$ ) from the observations

$$
\begin{cases}\mathbf{Y}^{(t)} & =\sqrt{\frac{1-t}{n}} \mathbf{U} \mathbf{V}^{\top}+\mathbf{Z} \\ \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)} & =\sqrt{\alpha_{v} R_{v}} \mathbf{U}+\widetilde{\mathbf{Z}} \\ \overline{\mathbf{Y}}^{\left(t, R_{u}\right)} & =\sqrt{\alpha_{u} R_{u}} \mathbf{V}+\overline{\mathbf{Z}}\end{cases}
$$

where $\mathbf{U} \sim P_{u}, \mathbf{V} \sim P_{v}$ and the entries of $\mathbf{Z} \in \mathbb{R}^{n_{u} \times n_{v}}, \widetilde{\mathbf{Z}} \in \mathbb{R}^{n_{u}}, \overline{\mathbf{Z}} \in \mathbb{R}^{n_{v}}$ are i.i.d. with respect to $\mathcal{N}(0,1)$. The posterior distribution of ( $\mathbf{U}, \mathbf{V})$ given $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)$ is

$$
d P\left(\mathbf{u}, \mathbf{v} \mid \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right):=\frac{d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}_{t, R}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)}}{\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)},
$$

where $\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)$ is the normalization factor and $\mathcal{H}_{t, R}$ denotes the associated interpolating Hamiltonian,

$$
\begin{aligned}
\mathcal{H}_{t, R}\left(\mathbf{u}, \mathbf{v} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right):= & \sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} \frac{1-t}{2 n} u_{i}^{2} v_{j}^{2}-\sqrt{\frac{1-t}{n}} u_{i} v_{j} Y_{i j}^{(t)} \\
& +\sum_{i=1}^{n_{u}} \frac{\alpha_{v} R_{v}}{2} u_{i}^{2}-\sqrt{\alpha_{v} R_{v}} u_{i} \widetilde{Y}_{i}^{\left(t, R_{v}\right)} \\
& +\sum_{j=1}^{n_{v}} \frac{\alpha_{u} R_{u}}{2} v_{j}^{2}-\sqrt{\alpha_{u} R_{u}} v_{j} \bar{Y}_{j}^{\left(t, R_{u}\right)}
\end{aligned}
$$

The angular brackets $\langle-\rangle_{t, R}$ denote the expectation w.r.t. this posterior. Define

$$
F_{v}(t, R):=\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R} \quad \text { and } \quad F_{u}(t, R):=2 \rho_{u} \varphi^{\prime}\left(\alpha_{v} \rho_{u} \mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right),
$$

where $Q_{u}:=\mathbf{u}^{\top} \mathbf{U} / n_{u}$ and $Q_{v}:=\mathbf{v}^{\top} \mathbf{V} / n_{v}$.
Proposition 3.5. For all $\epsilon \in[0,+\infty)^{2}$, there exists a unique global solution, denoted $R(\cdot, \epsilon):[0,1] \rightarrow[0,+\infty)^{2}$, to the initial value problem

$$
g^{\prime}=\left(F_{u}(t, g), F_{v}(t, g)\right), g(0)=\epsilon
$$

$R(\cdot, \epsilon)$ is continuously differentiable and the image of its derivative $R^{\prime}(\cdot, \epsilon)$ is $R^{\prime}([0,1], \epsilon) \subseteq\left[0, \rho_{u}\right] \times\left[0, \rho_{v}\right]$. Besides, for all $t \in[0,1], R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $[0,+\infty)^{2}$ into its image whose Jacobian determinant is greater than, or equal to, one;

$$
\begin{equation*}
\forall \epsilon \in[0,+\infty)^{2}: \operatorname{det} J_{R(t,)}(\epsilon) \geq 1 \tag{3.18}
\end{equation*}
$$

where $J_{R(t, \cdot)}$ denotes the Jacobian matrix of $R(t, \cdot)$. Let $m_{u} \in\left[0, \rho_{u}\right]$. The same statement holds true if we instead consider the initial value problem

$$
g^{\prime}=\left(m_{u}, F_{v}(t, g)\right), g(0)=\epsilon
$$

Proof. We limit ourselves to the proof for the ODE $g^{\prime}=\left(F_{u}(t, g), F_{v}(t, g)\right)$, the one for $g^{\prime}=\left(m_{u}, F_{v}(t, g)\right)$ is simpler and follows the same arguments.

By the Nishimori identity and the Cauchy-Schwarz inequality, we have

$$
0 \leq \frac{\mathbb{E}\left\|\langle\mathbf{v}\rangle_{t, R}\right\|^{2}}{n_{v}}=\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R} \leq \frac{\mathbb{E}\langle\|\mathbf{v}\|\|\mathbf{V}\|\rangle_{t, R}}{n_{v}}=\rho_{v}
$$

hence $\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R} \in\left[0, \rho_{v}\right]$ for all $(t, R) \in[0,1] \times[0,+\infty)^{2}$. Thus, the function $F:(t, R) \mapsto\left(F_{u}(t, R), F_{v}(t, R)\right)$ is defined on all $[0,1] \times[0,+\infty)^{2}$ and takes value in $\left[0, \rho_{u}\right] \times\left[0, \rho_{v}\right]$.

To prove the existence and uniqueness of a solution to the initial value problem $g^{\prime}=\left(F_{u}(t, g), F_{v}(t, g)\right), g(0)=\epsilon$, we invoke the Picard-Lindelöf theorem 104, Theorem 1.1]. To do so, we have to check that $F$ is continuous in $t$ and uniformly Lipschitz continuous in $R$ (meaning that the Lipschitz constant does not depend on $t$ ). We can show the continuity in $t$ thanks to the dominated convergence theorem (continuity under the integral sign). Again with an application of the dominated convergence theorem, we can prove that $F$ is continuously differentiable with respect to $R \in(0,+\infty)^{2}$ (differentiability under the integral sign). To check the uniform Lipschitzianity, we show that the Jacobian matrix $J_{F(t,)}(R)$ of $F(t, \cdot)$ is uniformly bounded for $(t, R) \in[0,1] \times(0,+\infty)^{2}$. We have

$$
J_{F(t,)}(R)=\left[\begin{array}{cc}
c(t, R) & c(t, R) \\
1 & 1
\end{array}\right]\left[\begin{array}{cc}
\left.\frac{\partial F_{v}}{\partial R_{u}}\right|_{t, R} & 0 \\
0 & \left.\frac{\partial F_{v}}{\partial R_{v}}\right|_{t, R}
\end{array}\right]
$$

where $c(t, R):=2 \alpha_{v} \rho_{u}^{2} \varphi^{\prime \prime}\left(\alpha_{v} \rho_{u} F_{v}(t, R)\right) \in\left[0, \alpha_{v} \rho_{u}^{2}\right]$ and

$$
\begin{aligned}
\frac{\partial F_{v}}{\partial R_{u}} & =\frac{\alpha_{v}}{n_{v}} \sum_{i=1}^{n_{v}} \sum_{j=1}^{n_{v}} \mathbb{E}\left[\left(\left\langle v_{i} v_{j}\right\rangle_{t, R}-\left\langle v_{i}\right\rangle_{t, R}\left\langle v_{j}\right\rangle_{t, R}\right)^{2}\right], \\
\frac{\partial F_{v}}{\partial R_{v}} & =\frac{\alpha_{v}}{n_{v}} \sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} \mathbb{E}\left[\left(\left\langle u_{i} v_{j}\right\rangle_{t, R}-\left\langle u_{i}\right\rangle_{t, R}\left\langle v_{j}\right\rangle_{t, R}\right)^{2}\right] .
\end{aligned}
$$

Both $\partial F_{v} / \partial R_{u}, \partial F_{v} / \partial R_{v}$ are clearly nonnegative. If ( $\mathbf{u}, \mathbf{v}$ ) are jointly distributed w.r.t. the posterior distribution (3.15) then $\|\mathbf{u}\|=\sqrt{\rho_{u} n_{u}}$ and $\|\mathbf{v}\|=\sqrt{\rho_{v} n_{v}}$, hence $\partial F_{v} / \partial R_{u} \leq 4 \alpha_{v} \rho_{v}^{2} n_{v}$ and $\partial F_{v} / \partial R_{v} \leq 4 \alpha_{v} \rho_{u} \rho_{v} n_{u}$. Thus, the entries of $J_{F(t,)}(R)$ are uniformly bounded in $(t, R)$.

By the Picard-Lindelöf theorem, for all $\epsilon=\left(\epsilon_{u}, \epsilon_{v}\right) \in[0,+\infty)^{2}$ there exists a unique solution $R(\cdot, \epsilon):[0, \delta] \rightarrow[0,+\infty)^{2}$ to the initial value problem $g^{\prime}=F(t, g)$, $g(0)=\epsilon$, where $[0, \delta] \subseteq[0,1]$ is the maximal interval of existence of the solution. The function $F$ takes its values in $\left[0, \rho_{u}\right] \times\left[0, \rho_{v}\right]$ so $R([0, \delta], \epsilon) \subseteq\left[\epsilon_{u}, \epsilon_{u}+\delta \rho_{u}\right] \times$ $\left[\epsilon_{v}, \epsilon_{v}+\delta \rho_{v}\right]$ which means that $\delta=1$ (the solution never leaves the domain of definition of $F$ ).

Each initial value $\epsilon \in[0,+\infty)^{2}$ is tied to a unique solution $R(\cdot, \epsilon)$, hence $\epsilon \mapsto R(t, \epsilon)$ is injective. It is also continuously differentiable [104, Theorem 3.1] with Jacobian determinant given by Liouville's formula 104, Corollary 3.1],

$$
\operatorname{det} J_{R(t,)}(\epsilon)=\left.\exp \int_{0}^{t} d s\left(\frac{\partial F_{u}}{\partial R_{u}}+\frac{\partial F_{v}}{\partial R_{v}}\right)\right|_{s, R(s, \epsilon)}
$$

This Jacobian determinant is greater than, or equal to, one because $\partial F_{u} / \partial R_{u}$ and $\partial F_{v} / \partial R_{v}$ are nonnegative. For $\partial F_{v} / \partial R_{v}$, this follows from our previous computations. For $\partial F_{u} / \partial R_{u}$, we simply remark that $\partial F_{u} / \partial R_{u}=\frac{\alpha_{v} \rho_{u}^{2}}{\left(1+\alpha_{v} \rho_{u} F_{v}(t, R)\right)^{2}} \partial F_{v} / \partial R_{u}$ where $\partial F_{v} / \partial R_{u} \geq 0$. Therefore, the Jacobian determinant is bounded away from 0 uniformly in $\epsilon$ and, by the inverse function theorem, the injective function $\epsilon \mapsto R(t, \epsilon)$ is a $C^{1}$-diffeomorphism from $[0,+\infty)^{2}$ onto its image.

## 3.D Concentration of the overlap

Remember that we denote by angular brackets $\langle-\rangle_{t, R}$ the expectation with respect to the posterior distribution (3.15) and we define the scalar overlaps $Q_{u}:=\mathbf{u}^{\top} \mathbf{U} / n_{u}$, $Q_{v}:=\mathbf{v}^{\top} \mathbf{V} / n_{v}$. The end goal of this appendix is proving the following proposition on the concentration of the overlap $Q_{v}$.

Proposition 3.6. Let $M_{u}$ and $M_{v}$ be positive real numbers. For $n$ large enough, there exists a constant $M$ such that $\forall(a, b) \in\left(0, M_{u}\right)^{2}: a<\min \{1, b\}, \forall \delta \in(0, a)$, $\forall R_{v} \in\left[0, M_{v}\right], \forall t \in[0,1]:$

$$
\int_{a}^{b} \mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R_{u} \leq M\left(\frac{1}{\delta^{2} n}-\frac{\ln (a)}{n}+\frac{\delta}{a-\delta}\right)
$$

The proof of Proposition 3.6 is carried out mostly as in 87. The main difference is that we don't need to assume that the marginals of the prior $P_{v}$ have a support bounded uniformly with $n$. It will be enough that the norm of a vector distributed with respect to $P_{v}$ scales likes $\sqrt{n}$. The concentration of the overlap around its expectation follows from the concentration of the quantity

$$
\begin{equation*}
\mathcal{L}=\frac{1}{n} \sum_{j=1}^{n_{v}} \frac{\alpha_{u}}{2} v_{j}^{2}-\alpha_{u} v_{j} V_{j}-\frac{1}{2} \sqrt{\frac{\alpha_{u}}{R_{u}}} v_{j} \bar{Z}_{j} . \tag{3.50}
\end{equation*}
$$

We first prove a lemma that links the fluctuations of $Q_{v}$ to those of $\mathcal{L}$.
Lemma 3.9. $\forall(t, R) \in[0,1] \times(0,+\infty)^{2}$, we have

$$
\begin{gather*}
\mathbb{E}\langle\mathcal{L}\rangle_{t, R}=-\frac{\alpha_{u}}{2} \frac{n_{v}}{n} \mathbb{E}\left\langle Q_{v}\right\rangle_{t, R},  \tag{3.51}\\
\mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \leq \frac{4}{\alpha_{u}^{2}}\left(\frac{n}{n_{v}}\right)^{2} \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} . \tag{3.52}
\end{gather*}
$$

Proof. Fix $(t, R) \in[0,1] \times(0,+\infty)^{2}$. By the definition (3.50) of $\mathcal{L}$, we have

$$
\begin{equation*}
\mathbb{E}\langle\mathcal{L}\rangle_{t, R}=\frac{1}{n} \sum_{j=1}^{n_{v}} \frac{\alpha_{u}}{2} \mathbb{E}\left\langle v_{j}^{2}\right\rangle_{t, R}-\alpha_{u} \mathbb{E}\left[\left\langle v_{j}\right\rangle_{t, R} V_{j}\right]-\frac{1}{2} \sqrt{\frac{\alpha_{u}}{R_{u}}} \mathbb{E}\left[\left\langle v_{j}\right\rangle_{t, R} \bar{Z}_{j}\right] \tag{3.53}
\end{equation*}
$$

and
$\mathbb{E}\left\langle Q_{v} \mathcal{L}\right\rangle_{t, R}=\frac{1}{n} \sum_{j=1}^{n_{v}} \frac{\alpha_{u}}{2} \mathbb{E}\left\langle Q_{v} v_{j}^{2}\right\rangle_{t, R}-\alpha_{u} \mathbb{E}\left[\left\langle Q_{v} v_{j}\right\rangle_{t, R} V_{j}\right]-\frac{1}{2} \sqrt{\frac{\alpha_{u}}{R_{u}}} \mathbb{E}\left[\left\langle Q_{v} v_{j}\right\rangle_{t, R} \bar{Z}_{j}\right]$.

Integrating by parts with respect to the Gaussian random variable $\bar{Z}_{j}$, the rightmost expectations in (3.53) and (3.54) satisfy

$$
\begin{align*}
\mathbb{E}\left[\left\langle v_{j}\right\rangle_{t, R} \bar{Z}_{j}\right] & =\sqrt{\alpha_{u} R_{u}} \mathbb{E}\left[\left\langle v_{j}^{2}\right\rangle_{t, R}\right]-\sqrt{\alpha_{u} R_{u}} \mathbb{E}\left[\left\langle v_{j}\right\rangle_{t, R}^{2}\right],  \tag{3.55}\\
\mathbb{E}\left[\left\langle Q_{v} v_{j}\right\rangle_{t, R} \bar{Z}_{j}\right] & =\sqrt{\alpha_{u} R_{u}} \mathbb{E}\left[\left\langle Q_{v} v_{j}^{2}\right\rangle_{t, R}\right]-\sqrt{\alpha_{u} R_{u}} \mathbb{E}\left[\left\langle Q_{v} v_{j}\right\rangle_{t, R}\left\langle v_{j}\right\rangle_{t, R}\right] . \tag{3.56}
\end{align*}
$$

Plugging (3.55) back in (3.53) yields

$$
\mathbb{E}\langle\mathcal{L}\rangle_{t, R}=\frac{\alpha_{u}}{n} \sum_{j=1}^{n_{v}} \frac{1}{2} \mathbb{E}\left[\left\langle v_{j}\right\rangle_{t, R}^{2}\right]-\mathbb{E}\left[\left\langle v_{j}\right\rangle_{t, R} V_{j}\right]=-\frac{\alpha_{u}}{2} \frac{n_{v}}{n} \mathbb{E}\left\langle Q_{v}\right\rangle_{t, R},
$$

where the second equality follows from the Nishimori identity $\mathbb{E}\left\langle v_{j}\right\rangle_{t, R}^{2}=\mathbb{E}\left\langle v_{j}\right\rangle_{t, R} V_{j}$. This ends the proof of (3.51). Plugging (3.56) back in (3.54), it comes

$$
\begin{align*}
\mathbb{E}\left\langle Q_{v} \mathcal{L}\right\rangle_{t, R} & =\frac{\alpha_{u}}{n} \sum_{j=1}^{n_{v}} \frac{1}{2} \mathbb{E}\left[\left\langle Q_{v} v_{j}\right\rangle_{t, R}\left\langle v_{j}\right\rangle_{t, R}\right]-\mathbb{E}\left[\left\langle Q_{v} v_{j}\right\rangle_{t, R} V_{j}\right] \\
& =\frac{\alpha_{u}}{n} \sum_{j=1}^{n_{v}} \frac{1}{2} \mathbb{E}\left[\left\langle Q_{v}\right\rangle_{t, R}\left\langle v_{j} V_{j}\right\rangle_{t, R}\right]-\mathbb{E}\left[\left\langle Q_{v} v_{j}\right\rangle_{t, R} V_{j}\right] \\
& =\alpha_{u} \frac{n_{v}}{n}\left(\frac{1}{2} \mathbb{E}\left[\left\langle Q_{v}\right\rangle_{t, R}^{2}\right]-\mathbb{E}\left\langle Q_{v}^{2}\right\rangle_{t, R}\right) . \tag{3.57}
\end{align*}
$$

The second equality follows once again from the Nishimori identity,

$$
\begin{aligned}
\mathbb{E}\left[\left\langle Q_{v} v_{j}\right\rangle_{t, R}\left\langle v_{j}\right\rangle_{t, R}\right] & =\frac{1}{n_{v}} \sum_{i=1}^{n_{v}} \mathbb{E}\left[\left\langle v_{i} V_{i} v_{j}\right\rangle_{t, R}\left\langle v_{j}\right\rangle_{t, R}\right] \\
& =\frac{1}{n_{v}} \sum_{i=1}^{n_{v}} \mathbb{E}\left[V_{i}\left\langle v_{i}\right\rangle_{t, R} V_{j}\left\langle v_{j}\right\rangle_{t, R}\right]=\mathbb{E}\left[\left\langle Q_{v}\right\rangle_{t, R}\left\langle v_{j} V_{j}\right\rangle_{t, R}\right] .
\end{aligned}
$$

We combine (3.57) and (3.51) to obtain

$$
\begin{aligned}
& \mathbb{E}\left\langle Q_{v}\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)\right\rangle_{t, R}=\mathbb{E}\left\langle Q_{v} \mathcal{L}\right\rangle_{t, R}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R} \mathbb{E}\langle\mathcal{L}\rangle_{t, R} \\
&=\frac{\alpha_{u}}{2} \frac{n_{v}}{n}\left(\mathbb{E}\left[\left\langle Q_{v}\right\rangle_{t, R}^{2}\right]-2 \mathbb{E}\left\langle Q_{v}^{2}\right\rangle_{t, R}+\mathbb{E}\left[\left\langle Q_{v}\right\rangle_{t, R}\right]^{2}\right) \\
&=-\frac{\alpha_{u}}{2} \frac{n_{v}}{n}\left(\mathbb{E}\left\langle\left(Q_{v}-\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R}+\mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R}\right) .
\end{aligned}
$$

The last identity directly implies

$$
\begin{aligned}
& \frac{\alpha_{u}}{2} \frac{n_{v}}{n} \mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \\
& \leq\left|\mathbb{E}\left\langle Q_{v}\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)\right\rangle_{t, R}\right| \\
&=\left|\mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)\right\rangle_{t, R}\right| \\
& \leq \sqrt{\mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \cdot \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}}
\end{aligned}
$$

The upper bound (3.52) on the fluctuation of $Q_{v}$ follows simply from this last upper bound.

## 3.D. 1 Concentration of $\mathcal{L}$ around its expectation

To prove concentration results on $\mathcal{L}$, it is useful to work with the free entropy

$$
F_{n}(t, R):=\frac{1}{n} \ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)
$$

and its average

$$
f_{n}(t, R):=\mathbb{E} F_{n}(t, R)=\frac{1}{n} \mathbb{E}\left[\ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)\right]
$$

where $\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)$ is the normalization factor of the posterior distribution (3.15). In Appendix 3.E, we prove that the free entropy concentrates around its expectation when $n \rightarrow+\infty$.

Proposition 3.10 (Thermal fluctuations of $\mathcal{L}$ ). For $n$ large enough, we have for all positive real numbers $a<b, t \in[0,1]$ and $R_{v} \in[0,+\infty)$ :

$$
\begin{equation*}
\int_{a}^{b} d R_{u} \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \leq \frac{\alpha_{u} \alpha_{v} \rho_{v}}{n}\left(\frac{\ln (b / a)}{2}+1\right) \tag{3.58}
\end{equation*}
$$

Proof. Fix $(n, t) \in \mathbb{N}^{*} \times[0,1]$. Note that $\forall R \in(0,+\infty)^{2}$ :

$$
\begin{equation*}
\left.\frac{\partial f_{n}}{\partial R_{u}}\right|_{t, R}=-\frac{1}{n} \mathbb{E}\left[\left\langle\frac{\partial \mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)}{\partial R_{u}}\right\rangle_{t, R}\right]=-\mathbb{E}\langle\mathcal{L}\rangle_{t, R} \tag{3.59}
\end{equation*}
$$

Further differentiating, we obtain

$$
\begin{align*}
\left.\frac{\partial^{2} f_{n}}{\partial R_{u}^{2}}\right|_{t, R} & =\mathbb{E}\left[\left\langle\mathcal{L} \frac{\partial \mathcal{H}_{t, R}}{\partial R_{u}}\right\rangle_{t, R}\right]-\mathbb{E}\left[\langle\mathcal{L}\rangle_{t, R}\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial R_{u}}\right\rangle_{t, R}\right]-\mathbb{E}\left\langle\frac{\partial \mathcal{L}}{\partial R_{u}}\right\rangle_{t, R} \\
& =n \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\frac{1}{4 R_{u}} \sqrt{\frac{\alpha_{u}}{R_{u}}} \frac{\mathbb{E}\left[\langle\mathbf{v}\rangle_{t, R}^{\top} \overline{\mathbf{Z}}\right]}{n} \tag{3.60}
\end{align*}
$$

It follows directly from (3.60) that

$$
\begin{equation*}
\mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}=\left.\frac{1}{n} \frac{\partial^{2} f_{n}}{\partial R_{u}^{2}}\right|_{t, R}+\frac{1}{4 R_{u}} \sqrt{\frac{\alpha_{u}}{R_{u}}} \frac{\mathbb{E}\left[\langle\mathbf{v}\rangle_{t, R}^{\top} \overline{\mathbf{Z}}\right]}{n^{2}} \tag{3.61}
\end{equation*}
$$

We start with upper bounding the integral over the second summand on the right-hand side of $(3.61)$. Integrating by parts w.r.t. $\bar{Z}_{j} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ gives

$$
\begin{equation*}
\frac{1}{4 R_{u}} \sqrt{\frac{\alpha_{u}}{R_{u}}} \frac{\mathbb{E}\left[\langle\mathbf{v}\rangle_{t, R} \overline{\mathbf{Z}}\right]}{n^{2}}=\frac{\alpha_{u}}{4 R_{u}} \frac{\mathbb{E}\left\langle\|\mathbf{v}\|^{2}\right\rangle_{t, R}-\mathbb{E}\left\|\langle\mathbf{v}\rangle_{t, R}\right\|^{2}}{n^{2}} \leq \frac{\alpha_{u} \rho_{v}}{4 R_{u}} \frac{n_{v}}{n^{2}}, \tag{3.62}
\end{equation*}
$$

where we use that $\|\mathbf{v}\|=\sqrt{\rho_{v} n_{v}}$. Therefore,

$$
\begin{equation*}
\int_{a}^{b} \frac{d R_{u}}{4 R_{u}} \sqrt{\frac{\alpha_{u}}{R_{u}}} \frac{\mathbb{E}\left[\langle\mathbf{v}\rangle_{t, R}^{\top} \overline{\mathbf{Z}}\right]}{n^{2}} \leq \frac{n_{v}}{n^{2}} \frac{\alpha_{u} \rho_{v} \ln (b / a)}{4} \tag{3.63}
\end{equation*}
$$

It remains to upper bound $\left.\int_{a}^{b} \frac{d R_{u}}{n} \frac{\partial^{2} f_{n}}{\partial R_{u}^{2}}\right|_{t, R}=\left.\frac{1}{n} \frac{\partial f_{n}}{\partial R_{u}}\right|_{t, R_{u}=b, R_{v}}-\left.\frac{1}{n} \frac{\partial f_{n}}{\partial R_{u}}\right|_{t, R_{u}=a, R_{v}}$. Note that $\forall R \in[0,+\infty)^{2}$ :

$$
\begin{equation*}
\left.\frac{\partial f_{n}}{\partial R_{u}}\right|_{t, R}=-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}=\frac{\alpha_{u}}{2} \frac{n_{v}}{n} \mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}=\frac{\alpha_{u}}{2} \frac{n_{v}}{n} \frac{\mathbb{E}\left[\left\|\langle\mathbf{v}\rangle_{t, R}\right\|^{2}\right]}{n_{v}}, \tag{3.64}
\end{equation*}
$$

where the first equality follows from (3.59), the second from Lemma 3.9, and the third from the Nishimori identity $\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}=\mathbb{E}\left[\langle\mathbf{v}\rangle_{t, R}^{\top} \mathbf{V}\right] / n_{v}=\mathbb{E}\left[\left\|\langle\mathbf{v}\rangle_{t, R}\right\|^{2}\right] / n_{v}$. Making use of (3.64) and Jensen's inequality, it comes $\forall R \in[0,+\infty)^{2}$ :

$$
\begin{equation*}
0 \leq\left.\frac{\partial f_{n}}{\partial R_{u}}\right|_{t, R} \leq \frac{\alpha_{u}}{2} \frac{n_{v}}{n} \frac{\mathbb{E}\left\langle\|\mathbf{v}\|^{2}\right\rangle_{t, R}}{n_{v}}=\frac{\alpha_{u} \rho_{v}}{2} \frac{n_{v}}{n} . \tag{3.65}
\end{equation*}
$$

Combining both (3.63) and (3.65), we finally get

$$
\begin{equation*}
\int_{a}^{b} d R_{u} \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \leq \frac{1}{n} \frac{n_{v}}{n} \frac{\alpha_{u} \rho_{v}}{2}\left(\frac{\ln (b / a)}{2}+1\right) \tag{3.66}
\end{equation*}
$$

Proposition 3.11 (Quenched fluctuations of $\mathcal{L}$ ). Let $M_{u}, M_{v}>0$. For $n$ large enough, there exists a constant $M$ such that $\forall(a, b) \in\left(0, M_{u}\right)^{2}: a<\min \{1, b\}$, $\forall \delta \in(0, a), \forall R_{v} \in\left[0, M_{v}\right], \forall t \in[0,1]:$

$$
\begin{equation*}
\int_{a}^{b} d R_{u} \mathbb{E}\left\langle\left(\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \leq M\left(\frac{1}{\delta^{2} n}-\frac{\ln (a)}{n}+\frac{\delta}{a-\delta}\right) \tag{3.67}
\end{equation*}
$$

Proof. Fix $(n, t) \in \mathbb{N}^{*} \times[0,1]$. For all $R \in(0,+\infty)^{2}$, we have:

$$
\begin{align*}
\left.\frac{\partial F_{n}}{\partial R_{u}}\right|_{t, R} & =-\langle\mathcal{L}\rangle_{t, R},  \tag{3.68}\\
\left.\frac{\partial^{2} F_{n}}{\partial R_{u}^{2}}\right|_{t, R} & =n\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\frac{1}{4 R_{u}} \sqrt{\frac{\alpha_{u}}{R_{u}}} \frac{\langle\mathbf{v}\rangle_{t, R}^{\top} \overline{\mathbf{Z}}}{n},  \tag{3.69}\\
\left.\frac{\partial f_{n}}{\partial R_{u}}\right|_{t, R} & =-\mathbb{E}\langle\mathcal{L}\rangle_{t, R},  \tag{3.70}\\
\left.\frac{\partial^{2} f_{n}}{\partial R_{u}^{2}}\right|_{t, R} & =n \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\frac{1}{4 R_{u}} \sqrt{\frac{\alpha_{u}}{R_{u}}} \frac{\mathbb{E}\left[\langle\mathbf{v}\rangle_{t, R}^{\top} \overline{\mathbf{Z}}\right]}{n} . \tag{3.71}
\end{align*}
$$

By the Cauchy-Schwarz inequality, the right-most term in (3.69) satisfies

$$
\begin{align*}
\left|\frac{1}{4 R_{u}} \sqrt{\frac{\alpha_{u}}{R_{u}}} \frac{\langle\mathbf{v}\rangle_{t, R}^{\top} \overline{\mathbf{Z}}}{n}\right| & \leq \frac{1}{4 R_{u}} \sqrt{\frac{\alpha_{u}}{R_{u}}} \frac{\left\|\langle\mathbf{v}\rangle_{t, R}\right\|\|\overline{\mathbf{Z}}\|}{n} \\
& \leq \frac{1}{4 R_{u}} \sqrt{\frac{\alpha_{u}}{R_{u}}} \frac{\langle\|\mathbf{v}\|\rangle_{t, R}\|\overline{\mathbf{Z}}\|}{n} \leq \frac{1}{4 R_{u}} \sqrt{\frac{\alpha_{u} \rho_{v}}{R_{u}} \frac{n_{v}}{n}} \frac{\|\overline{\mathbf{Z}}\|}{\sqrt{n}} . \tag{3.72}
\end{align*}
$$

We now define for all $R_{u} \in(0,+\infty)$ :

$$
\begin{align*}
& F\left(R_{u}\right):=F_{n}\left(t,\left(R_{u}, R_{v}\right)\right)-\sqrt{\alpha_{u} \rho_{v} R_{u} \frac{n_{v}}{n}} \frac{\|\overline{\mathbf{Z}}\|}{\sqrt{n}}  \tag{3.73}\\
& f\left(R_{u}\right):=f_{n}\left(t,\left(R_{u}, R_{v}\right)\right)-\sqrt{\alpha_{u} \rho_{v} R_{u} \frac{n_{v}}{n}} \frac{\mathbb{E}\|\overline{\mathbf{Z}}\|}{\sqrt{n}} . \tag{3.74}
\end{align*}
$$

$F$ is convex on $(0,+\infty)$ as it is twice differentiable with a nonnegative second derivative by (3.69) and (3.72). The same holds for $f$. Note that $\forall R_{u} \in(0,+\infty)$ :

$$
\begin{aligned}
F\left(R_{u}\right)-f\left(R_{u}\right) & =F_{n}\left(t,\left(R_{u}, R_{v}\right)\right)-f_{n}\left(t,\left(R_{u}, R_{v}\right)\right)-\sqrt{\alpha_{u} \rho_{v} R_{u} \frac{n_{v}}{n}} \frac{\|\overline{\mathbf{Z}}\|-\mathbb{E}\|\overline{\mathbf{Z}}\|}{\sqrt{n}}, \\
F^{\prime}\left(R_{u}\right)-f^{\prime}\left(R_{u}\right) & =-\left(\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)-\frac{1}{2} \sqrt{\frac{\alpha_{u} \rho_{v}}{R_{u}}} \frac{n_{v}}{n} \frac{\|\overline{\mathbf{Z}}\|-\mathbb{E}\|\overline{\mathbf{Z}}\|}{\sqrt{n}}
\end{aligned}
$$

It follows from Lemma 2.8 (applied to the convex functions $G=F, g=f$ ) and these last two identities that $\forall R_{u} \in(0,+\infty), \forall \delta \in\left(0, R_{u}\right)$ :

$$
\begin{aligned}
\left|\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right| \leq & \frac{1}{2} \sqrt{\frac{\alpha_{u} \rho_{v}}{R_{u}} \frac{n_{v}}{n}} \frac{|\|\overline{\mathbf{Z}}\|-\mathbb{E}\|\overline{\mathbf{Z}}\||}{\sqrt{n}}+C_{\delta}\left(R_{u}\right) \\
& +\frac{1}{\delta} \sum_{x \in\{-\delta, 0, \delta\}}\left|F\left(R_{u}+x\right)-f\left(R_{u}+x\right)\right| \\
\leq & \sqrt{\alpha_{u} \rho_{v} \frac{n_{v}}{n}}\left(\frac{1}{2 \sqrt{R_{u}}}+3 \sqrt{R_{u}}\right) \frac{|\|\overline{\mathbf{Z}}\|-\mathbb{E}\|\overline{\mathbf{Z}}\||}{\sqrt{n}}+C_{\delta}\left(R_{u}\right) \\
& \quad+\frac{1}{\delta} \sum_{x \in\{-\delta, 0, \delta\}}\left|F_{n}\left(t,\left(R_{u}+x, R_{v}\right)\right)-f_{n}\left(t,\left(R_{u}+x, R_{v}\right)\right)\right|,
\end{aligned}
$$

where $C_{\delta}(r):=f^{\prime}(r+\delta)-f^{\prime}(r-\delta)$ is nonnegative ( $f$ is convex). We now use the inequality $\left(\sum_{i=1}^{m} v_{i}\right)^{2} \leq m \sum_{i=1}^{m} v_{i}^{2}$ to obtain $\forall R_{u} \in(0,+\infty), \forall \delta \in\left(0, R_{u}\right)$ :

$$
\begin{align*}
\mathbb{E}\left[\left(\langle\mathcal{L}\rangle_{t, R}\right.\right. & \left.\left.-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right] \leq 5 \alpha_{u} \rho_{v} \frac{n_{v}}{n}\left(\frac{1}{4 R_{u}}+3+9 R_{u}\right) \frac{\operatorname{Var}\|\overline{\mathbf{Z}}\|}{n}+5 C_{\delta}\left(R_{u}\right)^{2} \\
& +\frac{5}{\delta^{2}} \sum_{x \in\{-\delta, 0, \delta\}} \mathbb{E}\left[\left(F_{n}\left(t,\left(R_{u}+x, R_{v}\right)\right)-f_{n}\left(t,\left(R_{u}+x, R_{v}\right)\right)\right)^{2}\right] . \tag{3.75}
\end{align*}
$$

The next step is to bound the integral of the three summands on the right-hand side of (3.75). By [99, Theorem 3.1.1], there exists $C_{1}$ such that $\mathbb{V a r}\|\overline{\mathbf{Z}}\| \leq C_{1}$ independently of the dimension $n_{v}$. Then,

$$
\begin{equation*}
\int_{a}^{b} d R_{u} \frac{n_{v}}{n}\left(\frac{1}{4 R_{u}}+3+9 R_{u}\right) \frac{\operatorname{Var}\|\overline{\mathbf{Z}}\|}{n} \leq \frac{n_{v}}{n}\left(\frac{\ln (b / a)}{4}+3 b+\frac{9}{2} b^{2}\right) \frac{C_{1}}{n} . \tag{3.76}
\end{equation*}
$$

Note that $C_{\delta}\left(R_{u}\right)=\left|C_{\delta}\left(R_{u}\right)\right| \leq\left|f^{\prime}\left(R_{u}+\delta\right)\right|+\left|f^{\prime}\left(R_{u}-\delta\right)\right|$. For all $R_{u} \in(0,+\infty)$ :

$$
\begin{equation*}
\left|f^{\prime}\left(R_{u}\right)\right| \leq\left|\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right|+\frac{1}{2} \sqrt{\frac{\alpha_{u} \rho_{v}}{R_{u}} \frac{n_{v}}{n}} \frac{\mathbb{E}}{\|\overline{\mathbf{Z}}\|} \sqrt{n} \leq \frac{n_{v}}{n} \frac{\sqrt{\alpha_{u} \rho_{v}}}{2}\left(\sqrt{\alpha_{u} \rho_{v}}+\frac{1}{\sqrt{R_{u}}}\right) \tag{3.77}
\end{equation*}
$$

where the second inequality in (3.77) follows from $\left|\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right| \leq \alpha_{u} \rho_{v} n_{v} / 2 n$ (see (3.64) and (3.65) and $\mathbb{E}\|\overline{\mathbf{Z}}\| \leq \mathbb{E}\left[\|\overline{\mathbf{Z}}\|^{2}\right]^{1 / 2}=\sqrt{n_{v}}$. Thus, for the second summand on the right-hand side of (3.75), we obtain $\forall \delta \in(0, a)$ :

$$
\begin{align*}
\int_{a}^{b} d R_{u} C_{\delta}\left(R_{u}\right)^{2} \leq & \frac{n_{v}}{n} \sqrt{\alpha_{u} \rho_{v}}\left(\sqrt{\alpha_{u} \rho_{v}}+\frac{1}{\sqrt{a-\delta}}\right) \int_{a}^{b} d R_{u} C_{\delta}\left(R_{u}\right) \\
= & \frac{n_{v}}{n} \sqrt{\alpha_{u} \rho_{v}}\left(\sqrt{\alpha_{u} \rho_{v}}+\frac{1}{\sqrt{a-\delta}}\right) \\
& \cdot(f(b+\delta)-f(b-\delta)-(f(a+\delta)-f(a-\delta))) \\
\leq & \delta\left(\frac{n_{v}}{n}\right)^{2} \alpha_{u} \rho_{v}\left(\sqrt{\alpha_{u} \rho_{v}}+\frac{1}{\sqrt{a-\delta}}\right)^{2} . \tag{3.78}
\end{align*}
$$

The last inequality is a simple application of the mean value theorem. We finally turn to the third summand. By Proposition 3.12 in Appendix 3.E, there exists a positive constant $C_{2}$ depending only on $a, b$ and $M_{v}$ such that $\forall t \in[0,1]$, $\forall\left(R_{u}, R_{v}\right) \in(0, b+a) \times\left(0, M_{v}\right):$

$$
\begin{equation*}
\mathbb{E}\left[\left(F_{n}(t, R)-f_{n}(t, R)\right)^{2}\right] \leq \frac{C_{2}}{n} \tag{3.79}
\end{equation*}
$$

Using (3.79), we see that the third summand satisfies $\forall \delta \in(0, a)$ :

$$
\begin{equation*}
\int_{a}^{b} d R_{u} \frac{5}{\delta^{2}} \sum_{x \in\{-\delta, 0, \delta\}} \mathbb{E}\left[\left(F_{n}\left(t,\left(R_{u}+x, R_{v}\right)\right)-f_{n}\left(t,\left(R_{u}+x, R_{v}\right)\right)\right)^{2}\right] \leq \frac{15 C_{2}}{\delta^{2} n} b \tag{3.80}
\end{equation*}
$$

To end the proof, we integrate 3.75 over $R_{u} \in[a, b]$ and use the three upper bounds (3.76), (3.78) and (3.80).

## 3.D. 2 Concentration of $Q_{v}$ around its expectation

Proof of Proposition 3.6. Using the upper bound (3.52) in Lemma 3.9 and the Cauchy-Schwarz inequality yields

$$
\int_{a}^{b} \mathbb{E}\left\langle\left(Q_{v}-\mathbb{E}\left\langle Q_{v}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R_{u} \leq \frac{4}{\alpha_{u}^{2}}\left(\frac{n}{n_{v}}\right)^{2} \int_{a}^{b} \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R_{u}
$$

We then use Propositions 3.10 and 3.11 to upper bound

$$
\begin{aligned}
\int_{a}^{b} \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R_{u}=\int_{a}^{b} \mathbb{E}\langle(\mathcal{L} & \left.\left.-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R_{u} \\
& +\int_{a}^{b} \mathbb{E}\left[\left(\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right] d R_{u}
\end{aligned}
$$

hence concluding the proof.

## 3.E Concentration of the free entropy

Consider the inference problem (3.14). Once the observations $\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}$ and $\overline{\mathbf{Y}}^{\left(t, R_{u}\right)}$ have been replaced by their definitions, the associated Hamiltonian reads

$$
\begin{aligned}
\mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}):=\sum_{i=1}^{n_{u}} & \sum_{j=1}^{n_{v}} \frac{(1-t)}{2 n} u_{i}^{2} v_{j}^{2}-\frac{1-t}{n} u_{i} v_{j} U_{i} V_{j}-\sqrt{\frac{1-t}{n}} u_{i} v_{j} Z_{i j} \\
& +\sum_{i=1}^{n_{u}} \frac{\alpha_{v} R_{v}}{2} u_{i}^{2}-\alpha_{v} R_{v} u_{i} U_{i}-\sqrt{\alpha_{v} R_{v}} u_{i} \widetilde{Z}_{i} \\
& +\sum_{j=1}^{n_{v}} \frac{\alpha_{u} R_{u}}{2} v_{j}^{2}-\alpha_{u} R_{u} v_{j} V_{j}-\sqrt{\alpha_{u} R_{u}} v_{j} \bar{Z}_{j} .
\end{aligned}
$$

In this section, we show that the free entropy

$$
\frac{1}{n} \ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)=\frac{1}{n} \ln \left(\int d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \tilde{\mathbf{Z}}, \overline{\mathbf{Z}})}\right)
$$

concentrates around its expectation. In what follows, we write $\frac{1}{n} \ln \mathcal{Z}_{t, R}$, omitting the arguments, to shorten notations.

Proposition 3.12 (Concentration of the free entropy). Let $M$ be a positive number. There exists a positive constant $C$ such that for any $R \in[0,+\infty)^{2}$ whose Euclidean norm is bounded by $M$ we have

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{t, R}}{n}-\mathbb{E}\left[\frac{\ln \mathcal{Z}_{t, R}}{n}\right]\right)^{2}\right] \leq \frac{C}{n} . \tag{3.81}
\end{equation*}
$$

Proof. To lighten notations, we drop the subscripts to the angular brackets $\langle-\rangle_{t, R}$ that denote an expectation w.r.t. the posterior (3.15). First, we show that the free entropy concentrates on its conditional expectation given $\mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}$. To do so, we see $g\left(\mathbf{U} / \sqrt{\rho_{u} n_{u}}\right):=\ln \mathcal{Z}_{t, R} / n$ as a function of $\mathbf{U} / \sqrt{\rho_{u} n_{u}}$ and we work conditionally to $\mathbf{V}$, $\mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}$. We normalize by $\sqrt{\rho_{u} n_{u}}$ so that $\mathbf{U} / \sqrt{\rho_{u} n_{u}}$ is uniformly distributed on the ( $n_{u}-1$ )-sphere of radius 1 and we can apply Lévy's lemma on the concentration of uniform measure on the sphere. For the reader's convenience, we reproduce the statement of this lemma given in [105, Corollary 5.4] (see this reference for a proof).

Lemma 3.13 (Lévy's lemma). Let $\mathcal{S}^{n-1}$ be the $(n-1)$-sphere of radius 1. Let $f: \mathcal{S}^{n-1} \rightarrow \mathbb{R}$ be Lipschitz continuous with Lipschitz constant L, and let $\mathbf{X}$ be a uniform random vector in $\mathcal{S}^{n-1}$. Then,

$$
\mathbb{P}(|f(\mathbf{X})-\mathbb{E} f(\mathbf{X})| \geq L t) \leq \exp \left(\pi-n t^{2} / 4\right)
$$

By Jensen's inequality, we have

$$
\begin{align*}
& \frac{1}{n}\left\langle\mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \widetilde{\mathbf{U}}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})-\mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})\right\rangle_{\tilde{\mathbf{U}}} \\
& \quad \leq g\left(\frac{\mathbf{U}}{\sqrt{\rho_{u} n_{u}}}\right)-g\left(\frac{\widetilde{\mathbf{U}}}{\sqrt{\rho_{u} n_{u}}}\right) \\
& \quad \leq \frac{1}{n}\left\langle\mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \widetilde{\mathbf{U}}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})-\mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})\right\rangle_{\mathbf{U}} \tag{3.82}
\end{align*}
$$

The subscript $\tilde{\mathbf{U}}$ (resp. $\mathbf{U}$ ) of the angular brackets on the left-hand side (resp. righthand side) of (3.82) means that ( $\mathbf{u}, \mathbf{v}$ ) is distributed according to the posterior $\propto d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \widetilde{\mathbf{U}}, \mathbf{V}, \mathbf{z}, \tilde{\mathbf{z}}, \overline{\mathbf{Z}})}\left(\right.$ resp. $\left.\propto d P_{u}(\mathbf{u}) d P_{v}(\mathbf{v}) e^{-\mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{z}, \widetilde{\mathbf{z}}, \overline{\mathbf{Z}})}\right)$. Note that

$$
\begin{aligned}
\mid \mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \widetilde{\mathbf{U}}, & \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})-\mathcal{H}_{t, R}(\mathbf{u}, \mathbf{v} ; \mathbf{U}, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}) \mid \\
& =\left|\frac{1-t}{n} \sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} v_{j} V_{j} u_{i}\left(U_{i}-\widetilde{U}_{i}\right)+\alpha_{v} R_{v} \sum_{i=1}^{n_{u}} u_{i}\left(U_{i}-\widetilde{U}_{i}\right)\right| \\
& =\left|\frac{1-t}{n} \mathbf{v}^{\top} \mathbf{V}+\alpha_{v} R_{v}\right| \cdot\left|\mathbf{u}^{\top}(\mathbf{U}-\widetilde{\mathbf{U}})\right| \\
& \leq\left(\frac{n_{v}}{n} \rho_{v}+\alpha_{v} R_{v}\right) \rho_{u} n_{u}\left\|\frac{\mathbf{U}}{\sqrt{\rho_{u} n_{u}}}-\frac{\widetilde{\mathbf{U}}}{\sqrt{\rho_{u} n_{u}}}\right\|
\end{aligned}
$$

Combining this last inequality with (3.82) yields

$$
\left|g\left(\frac{\mathbf{U}}{\sqrt{\rho_{u} n_{u}}}\right)-g\left(\frac{\widetilde{\mathbf{U}}}{\sqrt{\rho_{u} n_{u}}}\right)\right| \leq \rho_{u} \frac{n_{u}}{n}\left(\rho_{v} \frac{n_{v}}{n}+\alpha_{v} R_{v}\right)\left\|\frac{\mathbf{U}}{\sqrt{\rho_{u} n_{u}}}-\frac{\tilde{\mathbf{U}}}{\sqrt{\rho_{u} n_{u}}}\right\|,
$$

i.e., $g$ is Lipschitz continuous with Lipschitz constant $L=\rho_{u} \frac{n_{u}}{n}\left(\rho_{v} \frac{n_{v}}{n}+\alpha_{v} R_{v}\right)$. Lemma 3.13 then directly implies

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{t, R}}{n}-\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}\right]\right)^{2}\right] \leq \frac{4 L^{2} e^{\pi}}{n_{u}}=\frac{C_{1}}{n} \tag{3.83}
\end{equation*}
$$

where $C_{1}:=4 e^{\pi} \rho_{u}^{2} \frac{n_{u}}{n}\left(\rho_{v} \frac{n_{v}}{n}+\alpha_{v} R_{v}\right)^{2}$.
In a similar way, we can show that the conditional expectation of the free entropy given $\mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}$ concentrates on its conditional expectation given $\mathbf{Z}, \widetilde{\mathbf{Z}}$, $\overline{\mathbf{Z}}$, that is,

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{V}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}\right]-\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}\right]\right)^{2}\right] \leq \frac{C_{2}}{n} \tag{3.84}
\end{equation*}
$$

where $C_{2}=4 e^{\pi} \rho_{v}^{2} \frac{n_{v}}{n}\left(\rho_{u} \frac{n_{u}}{n}+\alpha_{u} R_{u}\right)^{2}$.
Finally, we show that the conditional expectation of the free entropy given $\mathbf{Z}, \widetilde{\mathbf{Z}}$, $\overline{\mathbf{Z}}$ concentrates on its expectation. To do so, we see $g(\mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}):=\mathbb{E}\left[\ln \mathcal{Z}_{t, R} / n \mid \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}\right]$
as a function of the Gaussian noises $\mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}$. By the Gaussian-Poincaré inequality (Proposition 2.7), we have

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}\right]-\mathbb{E}\left[\frac{\ln \mathcal{Z}_{t, R}}{n}\right]\right)^{2}\right] \leq \mathbb{E}\|\nabla g(\mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}})\|^{2} \tag{3.85}
\end{equation*}
$$

where the squared norm of the gradient of $g$ is simply

$$
\|\nabla g\|^{2}=\sum_{i, j}\left|\partial g / \partial Z_{i, j}\right|^{2}+\sum_{i}\left|\partial g / \partial \widetilde{Z}_{i}\right|^{2}+\sum_{j}\left|\partial g / \partial \bar{Z}_{j}\right|^{2}
$$

Each of these partial derivatives takes the form $\frac{\partial g}{\partial x}=-\frac{1}{n} \mathbb{E}\left[\left.\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial x}\right\rangle \right\rvert\, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}\right]$ and

$$
\left|\frac{\partial \mathcal{H}_{t, R}}{\partial Z_{i j}}\right|=\left|\sqrt{\frac{1-t}{n}} u_{i} v_{j}\right|,\left|\frac{\partial \mathcal{H}_{t, R}}{\partial \widetilde{Z}_{i}}\right|=\left|\sqrt{\alpha_{v} R_{v}} u_{i}\right|,\left|\frac{\partial \mathcal{H}_{t, R}}{\partial \bar{Z}_{j}}\right|=\left|\sqrt{\alpha_{u} R_{u}} v_{j}\right| .
$$

On one hand, by Jensen's inequality, we have

$$
\begin{equation*}
\sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} \mathbb{E}\left|\frac{\partial g}{\partial Z_{i j}}\right|^{2} \leq \frac{1}{n^{3}} \sum_{i=1}^{n_{u}} \sum_{j=1}^{n_{v}} \mathbb{E}\left[\left\langle u_{i}^{2} v_{j}^{2}\right\rangle\right]=\frac{n_{u}}{n} \frac{n_{v}}{n} \frac{\rho_{u} \rho_{v}}{n} \tag{3.86}
\end{equation*}
$$

On the other hand, still by Jensen's inequality, we have

$$
\begin{equation*}
\sum_{i=1}^{n_{u}} \mathbb{E}\left|\frac{\partial g}{\partial \widetilde{Z}_{i}}\right|^{2} \leq \frac{\alpha_{v} R_{v}}{n^{2}} \sum_{i=1}^{n_{u}} \mathbb{E}\left[\left\langle u_{i}^{2}\right\rangle\right]=\frac{n_{u}}{n} \frac{\alpha_{v} \rho_{u} R_{v}}{n} \tag{3.87}
\end{equation*}
$$

and

$$
\begin{equation*}
\sum_{j=1}^{n_{v}} \mathbb{E}\left|\frac{\partial g}{\partial \bar{Z}_{j}}\right|^{2} \leq \frac{\alpha_{u} R_{u}}{n^{2}} \sum_{j=1}^{n_{v}} \mathbb{E}\left[\left\langle v_{j}^{2}\right\rangle\right]=\frac{n_{v}}{n} \frac{\alpha_{u} \rho_{v} R_{u}}{n} \tag{3.88}
\end{equation*}
$$

Plugging (3.86), (3.87), and (3.88) back in (3.85) yields

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{Z}, \widetilde{\mathbf{Z}}, \overline{\mathbf{Z}}\right]-\mathbb{E}\left[\frac{\ln \mathcal{Z}_{t, R}}{n}\right]\right)^{2}\right] \leq \frac{C_{3}}{n} \tag{3.89}
\end{equation*}
$$

where $C_{3}:=\frac{n_{u}}{n} \frac{n_{v}}{n} \rho_{u} \rho_{v}+\frac{n_{u}}{n} \alpha_{v} \rho_{u} R_{v}+\frac{n_{v}}{n} \alpha_{u} \rho_{v} R_{u}$. Note that $C_{1}+C_{2}+C_{3} \xrightarrow[n \rightarrow+\infty]{ } C$,

$$
C:=\alpha_{u} \alpha_{v}\left(4 e^{\pi} \alpha_{v} \rho_{u}^{2}\left(\rho_{v}+R_{v}\right)^{2}+4 e^{\pi} \alpha_{u} \rho_{v}^{2}\left(\rho_{u}+R_{u}\right)^{2}+\rho_{u} \rho_{v}+\rho_{u} R_{v}+\rho_{v} R_{u}\right) .
$$

This limit is combined with the inequalities (3.83), (3.84), and (3.89) to obtain (3.81).

## 3.F Formula for the asymptotic MMSE

In the whole appendix the values of the positive hyperparameters $\alpha_{u}, \alpha_{v}, \rho_{u}$ and $\rho_{v}$ are fixed. Then, we define $\forall\left(m_{u}, m_{v}, \lambda\right) \in\left[0, \rho_{u}\right] \times\left[0, \rho_{v}\right] \times(0,+\infty)$ :

$$
\begin{aligned}
i\left(m_{u}, m_{v}, \lambda\right):=i_{\mathrm{RS}}\left(m_{u}, m_{v} ; \Theta\right)= & \frac{\lambda \alpha_{u} \alpha_{v}}{2}\left(\rho_{u}-m_{u}\right)\left(\rho_{v}-m_{v}\right) \\
& +\alpha_{u} \frac{\ln \left(1+\lambda \alpha_{v} \rho_{u} m_{v}\right)}{2}+\alpha_{v} \frac{\ln \left(1+\lambda \alpha_{u} \rho_{v} m_{u}\right)}{2} .
\end{aligned}
$$

Lemma 3.14. Let $\mathcal{D}:=\left[0, \rho_{u}\right] \times(0,+\infty)$. For every pair $\left(m_{u}, \lambda\right) \in \mathcal{D}$ there exists a unique $m_{v}^{*}\left(m_{u}, \lambda\right) \in\left[0, \rho_{v}\right]$ such that

$$
i\left(m_{u}, m_{v}^{*}\left(m_{u}, \lambda\right), \lambda\right)=\sup _{m_{v} \in\left[0, \rho_{v}\right]} i\left(m_{u}, m_{v}, \lambda\right),
$$

and it is given by the formula

$$
m_{v}^{*}\left(m_{u}, \lambda\right)=\left\{\begin{array}{cll}
\frac{m_{u}}{\lambda \alpha_{v} \rho_{u}\left(\rho_{u}-m_{u}\right)} & \text { if } & 0 \leq m_{u} \leq m_{u}(\lambda) \\
\rho_{v} & \text { if } & m_{u}(\lambda)<m_{u} \leq \rho_{u}
\end{array},\right.
$$

where $\forall \lambda \in\{0,+\infty\}$ :

$$
m_{u}(\lambda):=\rho_{u}\left(1-\frac{1}{1+\lambda \alpha_{v} \rho_{u} \rho_{v}}\right) .
$$

The function $m_{v}^{*}: \mathcal{D} \mapsto\left[0, \rho_{v}\right]$ is continuous and continuously differentiable on $\mathcal{D} \backslash\left\{\left(m_{u}(\lambda), \lambda\right): \lambda>0\right\}$. Finally,

$$
\begin{equation*}
\forall \lambda \in(0,+\infty), \forall m_{u} \in\left[0, m_{u}(\lambda)\right]:\left.\frac{\partial i}{\partial m_{v}}\right|_{m_{u}, m_{v}^{*}\left(m_{u}, \lambda\right), \lambda}=0 . \tag{3.90}
\end{equation*}
$$

Proof. Fix $\left(m_{u}, \lambda\right) \in \mathcal{D}$. Let $f: m_{v} \in\left[0, \rho_{v}\right] \mapsto i\left(m_{u}, m_{v}, \lambda\right) . f$ is continuously twice differentiable on $\left[0, \rho_{v}\right]$ with derivatives

$$
f^{\prime}\left(m_{v}\right)=\frac{\lambda \alpha_{u} \alpha_{v}}{2}\left(\frac{\rho_{u}}{1+\lambda \alpha_{v} \rho_{u} m_{v}}-\rho_{u}+m_{u}\right), f^{\prime \prime}\left(m_{v}\right)=-\frac{\lambda^{2} \alpha_{u} \alpha_{v}^{2} \rho_{u}^{2}}{2\left(1+\lambda \alpha_{v} \rho_{u} m_{v}\right)^{2}} .
$$

Note that

$$
f^{\prime}\left(m_{v}\right)=0 \Leftrightarrow m_{v}=\frac{m_{u}}{\lambda \alpha_{v} \rho_{u}\left(\rho_{u}-m_{u}\right)} .
$$

It is easy to check that the solution to $f^{\prime}\left(m_{v}\right)=0$ lies in $\left[0, \rho_{v}\right]$ if, and only if, $m_{u} \in\left[0, m_{u}(\lambda)\right]$ where $m_{u}(\lambda)$ is defined in the lemma. Besides, $f$ is strictly concave as $f^{\prime \prime}<0$. Therefore, $f$ has a unique global maximizer that is given by the unique solution to $f^{\prime}\left(m_{v}\right)=0$ if $m_{u} \in\left[0, m_{u}(\lambda)\right]$ and is equal to $\rho_{v}$ if $m_{u} \in\left[m_{u}(\lambda), \rho_{u}\right]$. The formula and properties of $m_{v}^{*}: \mathcal{D} \mapsto\left[0, \rho_{v}\right]$ directly follow.

Lemma 3.15. Define $h(\lambda):=\inf _{m_{u} \in\left[0, \rho_{u}\right]} \sup _{m_{v} \in\left[0, \rho_{v}\right]} i\left(m_{u}, m_{v}, \lambda\right)$. The function $h$ is continuously differentiable on $(0,+\infty)$ and for all $\lambda \in(0,+\infty)$ :

$$
\begin{align*}
h(\lambda) & =i\left(m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda\right),  \tag{3.91}\\
h^{\prime}(\lambda) & =\frac{\alpha_{u} \alpha_{v}}{2}\left(\rho_{u} \rho_{v}-m_{u}^{*}(\lambda) m_{v}^{*}(\lambda)\right) ; \tag{3.92}
\end{align*}
$$

where $\left(m_{u}^{*}(\lambda), m_{v}^{*}(\lambda)\right)$ is the unique solution to the extremization that defines $h$ and is given by the formulas

$$
\begin{align*}
& m_{u}^{*}(\lambda)=\left\{\begin{array}{cll}
0 & \text { if } & 0<\lambda \leq 1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}} \\
\frac{\lambda^{2} \alpha_{u} \alpha_{v} \rho_{v}^{2} \rho_{u}^{2}-1}{\lambda \alpha_{u} \rho_{v}\left(1+\lambda \alpha_{u} \rho_{v} \rho_{u}\right)} & \text { if } & \lambda>1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}
\end{array} ;\right.  \tag{3.93}\\
& m_{v}^{*}(\lambda)=\left\{\begin{array}{cl}
0 & \text { if } \\
\frac{\lambda^{2} \alpha_{u} \alpha_{v} \rho_{v}^{2} \rho_{u}^{2}-1}{\lambda \alpha_{v} \rho_{u}\left(1+\lambda \alpha_{u} \rho_{v} \rho_{u}\right)} & \text { if } \\
1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}
\end{array} .\right. \tag{3.94}
\end{align*}
$$

Proof. By Lemma 3.14, $h(\lambda)=\inf _{m_{u} \in\left[0, \rho_{u}\right]} g\left(m_{u}, \lambda\right)$ where $\forall \lambda \in(0,+\infty)$ :

$$
g\left(m_{u}, \lambda\right):=i\left(m_{u}, m_{v}^{*}\left(m_{u}, \lambda\right), \lambda\right) .
$$

By continuity of $i(\cdot, \cdot, \lambda)$ and $m_{v}^{*}(\cdot, \lambda), g(\cdot, \lambda)$ is continuous on $\left[0, \rho_{u}\right]$. Besides, $g(\cdot, \lambda)$ is increasing on $\left[m_{u}(\lambda), \rho_{u}\right]$ as $\forall m_{u} \in\left[m_{u}(\lambda), \rho_{u}\right]$ :

$$
g\left(m_{u}, \lambda\right):=i\left(m_{u}, \rho_{v}, \lambda\right)=\alpha_{u} \frac{\ln \left(1+\lambda \alpha_{v} \rho_{u} \rho_{v}\right)}{2}+\alpha_{v} \frac{\ln \left(1+\lambda \alpha_{u} \rho_{v} m_{u}\right)}{2} .
$$

Thus, we can restrict the infimum to the interval $\left[0, m_{u}(\lambda)\right]$ in the definition of $h$,

$$
h(\lambda)=\inf _{m_{u} \in\left[0, m_{u}(\lambda)\right]} g\left(m_{u}, \lambda\right) .
$$

For all $m_{u} \in\left[0, m_{u}(\lambda)\right]$ :

$$
\begin{aligned}
& g\left(m_{u}, \lambda\right):=i\left(m_{u}, \frac{m_{u}}{\lambda \alpha_{v} \rho_{u}\left(\rho_{u}-m_{u}\right)}, \lambda\right) \\
&= \frac{\alpha_{u}}{2 \rho_{u}}\left(\lambda \alpha_{v} \rho_{v} \rho_{u}^{2}-m_{u}\left(1+\lambda \alpha_{v} \rho_{v} \rho_{u}\right)\right) \\
&+\frac{\alpha_{u}}{2} \ln \left(\frac{\rho_{u}}{\rho_{u}-m_{u}}\right) \\
&+\frac{\alpha_{v}}{2} \ln \left(1+\lambda \alpha_{u} \rho_{v} m_{u}\right), \\
&\left.\frac{\partial g}{\partial m_{u}}\right|_{m_{u}, \lambda}=-\frac{\alpha_{u}}{2 \rho_{u}}\left(1+\lambda \alpha_{v} \rho_{v} \rho_{u}\right)+\frac{\alpha_{u}}{2} \frac{1}{\rho_{u}-m_{u}}+\frac{\alpha_{v}}{2} \frac{\lambda \alpha_{u} \rho_{v}}{1+\lambda \alpha_{u} \rho_{v} m_{u}} \\
&= a\left(m_{u}, \lambda\right) m_{u}\left(\frac{1-\lambda^{2} \alpha_{u} \alpha_{v} \rho_{v}^{2} \rho_{u}^{2}}{\lambda \alpha_{u} \rho_{v}\left(1+\lambda \alpha_{v} \rho_{v} \rho_{u}\right)}+m_{u}\right),
\end{aligned}
$$

where

$$
a\left(m_{u}, \lambda\right):=\frac{\lambda \alpha_{u}^{2} \rho_{v}\left(1+\lambda \alpha_{v} \rho_{v} \rho_{u}\right)}{2 \rho_{u}\left(\rho_{u}-m_{u}\right)\left(1+\lambda \alpha_{u} \rho_{v} m_{u}\right)} .
$$

Note that $\forall \lambda \in(0,+\infty), \forall m_{u} \in\left[0, m_{u}(\lambda)\right]: a\left(m_{u}, \lambda\right)>0$. If $\lambda \leq 1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}$ then 0 is the unique global minimizer of $g(\cdot, \lambda)$ on $\left[0, m_{u}(\lambda)\right]$. Instead, if $\lambda>1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}$ then $\left.\frac{\partial g}{\partial m_{u}}\right|_{m_{u}, \lambda}$ has a nonzero root given by

$$
\frac{\lambda^{2} \alpha_{u} \alpha_{v} \rho_{v}^{2} \rho_{u}^{2}-1}{\lambda \alpha_{u} \rho_{v}\left(1+\lambda \alpha_{v} \rho_{v} \rho_{u}\right)}=m_{u}(\lambda)-\frac{1}{\lambda \alpha_{u} \rho_{v}\left(1+\lambda \alpha_{v} \rho_{v} \rho_{u}\right)} \in\left(0, m_{u}(\lambda)\right) .
$$

We can easily check that this root is the unique global minimizer of $g(\cdot, \lambda)$ on $\left[0, m_{u}(\lambda)\right]$. Hence, we have just shown that

$$
m_{u}^{*}(\lambda):=\left\{\begin{array}{ccl}
0 & \text { if } & 0<\lambda \leq 1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}} \\
\frac{\lambda^{2} \alpha_{u} \alpha_{0} \rho_{v}^{2} \rho_{u}^{2}-1}{\lambda \alpha_{u} \rho_{v}\left(1+\lambda \alpha_{v} \rho_{v} \rho_{u}\right)} & \text { if } & \lambda>1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}
\end{array}\right.
$$

is the unique global minimizer of $g(\cdot, \lambda)$ on $\left[0, m_{u}(\lambda)\right]$ (and, in fact, $\left.\left[0, \rho_{u}\right]\right)$ and that $\forall \lambda \in(0,+\infty):\left.\frac{\partial g}{\partial m_{u}}\right|_{m_{u}^{*}(\lambda), \lambda}=0$. Define

$$
m_{v}^{*}(\lambda):=m_{v}^{*}\left(m_{u}^{*}(\lambda), \lambda\right)=\left\{\begin{array}{cll}
0 & \text { if } & 0<\lambda \leq 1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}  \tag{3.95}\\
\frac{\lambda^{2} \alpha_{u} \alpha_{v} \rho_{\lambda}^{2} \rho_{u}^{2}-1}{\lambda \alpha_{v} \rho_{u}\left(1+\lambda \rho_{u} \rho_{v} \rho_{u}\right)} & \text { if } & \lambda>1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}
\end{array} .\right.
$$

It follows from Lemma 3.14 that $\forall \lambda \in(0,+\infty)$ :

$$
\begin{equation*}
h(\lambda)=g\left(m_{u}^{*}(\lambda), \lambda\right)=i\left(m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda\right) . \tag{3.96}
\end{equation*}
$$

By Lemma 3.14, $m_{v}^{*}(\cdot, \lambda)$ is continuously differentiable on $\left[0, m_{u}(\lambda)\right]$ so for all $m_{u} \in\left[0, m_{u}(\lambda)\right):$

$$
\begin{aligned}
\left.\frac{\partial g}{\partial m_{u}}\right|_{m_{u}, \lambda} & =\left.\frac{\partial i}{\partial m_{u}}\right|_{m_{u}, m_{v}^{*}\left(m_{u}, \lambda\right), \lambda}+\left.\left.\frac{\partial i}{\partial m_{v}}\right|_{m_{u}, m_{v}^{*}\left(m_{u}, \lambda\right), \lambda} \cdot \frac{\partial m_{v}^{*}}{\partial m_{u}}\right|_{m_{u}, \lambda} \\
& =\left.\frac{\partial i}{\partial m_{u}}\right|_{m_{u}, m_{v}^{*}\left(m_{u}, \lambda\right), \lambda}
\end{aligned}
$$

where we use (3.90) to obtain the last equality. Evaluating the latter partial derivative at $\left(m_{u}, \lambda\right)=\left(m_{u}^{*}(\lambda), \lambda\right)$ yields

$$
\left.\frac{\partial i}{\partial m_{u}}\right|_{m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda}=\left.\frac{\partial g}{\partial m_{u}}\right|_{m_{u}^{*}(\lambda), \lambda}=0
$$

as we have previously shown in this proof that $\left.\frac{\partial g}{\partial m_{u}}\right|_{m_{u}^{*}(\lambda), \lambda}=0$. All in all, we have shown that $\forall \lambda \in(0,+\infty)$ :

$$
\begin{equation*}
\left.\frac{\partial i}{\partial m_{u}}\right|_{m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda}=\left.\frac{\partial i}{\partial m_{v}}\right|_{m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda}=0 . \tag{3.97}
\end{equation*}
$$

Combining (3.96) and the fact that $m_{u}^{*}, m_{v}^{*}$ are continuously differentiable on $(0,+\infty) \backslash\left\{1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}\right\}$, we obtain that $h$ is continuously differentiable on $(0,+\infty) \backslash$ $\left\{1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}\right\}$ and for all $\lambda \in(0,+\infty) \backslash\left\{1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}\right\}$ :

$$
\begin{aligned}
& h^{\prime}(\lambda)=\left.\frac{\partial i}{\partial \lambda}\right|_{m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda}+\left.\left.\frac{d m_{u}^{*}}{d \lambda}\right|_{\lambda} \cdot \frac{\partial i}{\partial m_{u}}\right|_{m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda}+\left.\left.\frac{d m_{v}^{*}}{d \lambda}\right|_{\lambda} \cdot \frac{\partial i}{\partial m_{v}}\right|_{m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda} \\
&=\left.\frac{\partial i}{\partial \lambda}\right|_{m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda} \\
&= \frac{\alpha_{u} \alpha_{v}}{2}\left(\rho_{u}-m_{u}^{*}(\lambda)\right)\left(\rho_{v}-m_{v}^{*}(\lambda)\right) \\
& \quad+\frac{\alpha_{u} \alpha_{v} \rho_{u} m_{v}^{*}(\lambda)}{2\left(1+\lambda \alpha_{v} \rho_{u} m_{v}^{*}(\lambda)\right)}+\frac{\alpha_{u} \alpha_{v} \rho_{v} m_{u}^{*}(\lambda)}{2\left(1+\lambda \alpha_{u} \rho_{v} m_{u}^{*}(\lambda)\right)} \\
&=\frac{\alpha_{u} \alpha_{v}}{2}\left(\rho_{u}-m_{u}^{*}(\lambda)\right)\left(\rho_{v}-m_{v}^{*}(\lambda)\right) \quad \\
& \quad+\frac{m_{v}^{*}(\lambda)}{\lambda}\left(\left.\frac{\partial i}{\partial m_{v}}\right|_{m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda}+\frac{\lambda \alpha_{u} \alpha_{v}}{2}\left(\rho_{u}-m_{u}^{*}(\lambda)\right)\right) \\
& \quad+\frac{m_{u}^{*}(\lambda)}{\lambda}\left(\left.\frac{\partial i}{\partial m_{u}}\right|_{m_{u}^{*}(\lambda), m_{v}^{*}(\lambda), \lambda}+\frac{\lambda \alpha_{u} \alpha_{v}}{2}\left(\rho_{v}-m_{v}^{*}(\lambda)\right)\right) \\
&= \frac{\alpha_{u} \alpha_{v}}{2}\left(\rho_{u} \rho_{v}-m_{u}^{*}(\lambda) m_{v}^{*}(\lambda)\right) .
\end{aligned}
$$

The second and last inequalities follow from (3.97). Note that the function $\lambda \mapsto \frac{\alpha_{u} \alpha_{v}}{2}\left(\rho_{u} \rho_{v}-m_{u}^{*}(\lambda) m_{v}^{*}(\lambda)\right)$ is continuous at $\lambda=1 / \rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}$. Therefore, $h$ is continuously differentiable on $(0,+\infty)$.

We now have everything we need to prove the formula for the MMSE in the high-dimensional limit.

Theorem 3.3. Define $\lambda_{\mathrm{IT}}:=\left(\rho_{u} \rho_{v} \sqrt{\alpha_{u} \alpha_{v}}\right)^{-1}$ and for all $\lambda \in(0,+\infty)$ :

$$
\left(m_{u}^{*}(\lambda), m_{v}^{*}(\lambda)\right)=\left\{\begin{array}{cl}
(0,0) & \text { if } 0<\lambda \leq \lambda_{\mathrm{IT}} \\
\left(\frac{\lambda^{2} \alpha_{u} \alpha_{v} \rho_{v}^{2} \rho_{u}^{2}-1}{\lambda \lambda_{u} \rho_{v}\left(1+\lambda \alpha_{v} \rho_{v} \rho_{u}\right)}, \frac{\lambda^{2} \alpha_{u} \alpha_{v} \rho_{\rho}^{2} \rho_{u}^{2}-1}{\lambda \alpha_{v} \rho_{u}\left(1+\lambda \lambda_{u} \rho_{v} \rho_{u}\right)}\right) & \text { if } \lambda>\lambda_{\mathrm{IT}}
\end{array} .\right.
$$

The pair $\left(m_{u}^{*}(\lambda), m_{v}^{*}(\lambda)\right)$ is the unique solution to the extremization over $\left(m_{u}, m_{v}\right)$ on the right-hand side of (3.6), and $\operatorname{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right)$ satisfies

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \operatorname{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right)=\rho_{u} \rho_{v}-m_{u}^{*}(\lambda) m_{v}^{*}(\lambda) \tag{3.8}
\end{equation*}
$$

Hence, the asymptotic MMSE is less than $\rho_{u} \rho_{v}$ if, and only if, $\lambda>\lambda_{\mathrm{IT}}$.
Proof. Let $n \in \mathbb{N}^{*}$. Define $h_{n}: \lambda \in(0,+\infty) \mapsto \frac{I(\mathbf{U}, \mathbf{V} ; \mathbf{Y})}{n}$ (the mutual information depends on $\lambda$ through $\mathbf{Y}$ ). We have the I-MMSE relationship [51]

$$
\begin{equation*}
h_{n}^{\prime}(\lambda)=\frac{\partial}{\partial \lambda}\left(\frac{I(\mathbf{U}, \mathbf{V} ; \mathbf{Y})}{n}\right)=\frac{n_{u}}{n} \frac{n_{v}}{n} \frac{\operatorname{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right)}{2} \tag{3.98}
\end{equation*}
$$

The function $\lambda \mapsto \operatorname{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right)$ is nonincreasing so $h_{n}$ is concave on $(0,+\infty)$. By Theorem 3.2, the sequence of continuously differentiable concave functions $\left(h_{n}\right)_{n \in \mathbb{N}^{*}}$ converges pointwise to $h: \lambda \mapsto \inf _{m_{u} \in\left[0, \rho_{u}\right]} \sup _{m_{v} \in\left[0, \rho_{v}\right]} i\left(m_{u}, m_{v}, \lambda\right)$. By Lemma 3.15, this limit $h$ is continuously differentiable. Therefore, by Griffiths' lemma 52, Appendix A], for all $\lambda \in(0,+\infty)$ :

$$
\lim _{n \rightarrow+\infty} h_{n}^{\prime}(\lambda)=h^{\prime}(\lambda)=\frac{\alpha_{u} \alpha_{v}}{2}\left(\rho_{u} \rho_{v}-m_{u}^{*}(\lambda) m_{v}^{*}(\lambda)\right) .
$$

We combine the latter with (3.98) to get

$$
\lim _{n \rightarrow+\infty} \operatorname{MMSE}_{\lambda}\left(\mathbf{U V}^{\top} \mid \mathbf{Y}\right)=\rho_{u} \rho_{v}-m_{u}^{*}(\lambda) m_{v}^{*}(\lambda)
$$

The rest of the theorem simply follows from Lemma 3.15 .

# Mutual information for low-rank even-order symmetric tensor estimation 

## 4

### 4.1 Introduction

There exist well-known unsupervised algorithms to discover structure in a 2D dataset, e.g., singular value decomposition (SVD), principal component analysis (PCA) and other spectral methods [1]. Tensors naturally handle multidimensional data and their use becomes more and more beneficial with the emergence of big data, a strong incentive to go beyond the flat matrix world. Tensor decompositions come with some advantages with respect to matrices, and have numerous applications in signal processing and machine learning, e.g., data compression, data visualization, learning probabilistic latent variables models, etc. [14], [94]. The canonical polyadic (CP) decomposition, also known as tensor rank decomposition or CP tensor factorization, is the most familiar one and represents a tensor as a minimum-length linear combination of rank-one tensors. This minimum-length defines the tensor rank. If instead the number $K$ of rank-one tensors forming the linear combination is not minimal, we talk of a $K$-term decomposition.

One approach to explore computational and statistical limits of tensor factorization is to consider a statistical model, as done in [19]. In this chapter we consider Model 1.2 that generalizes the one proposed in [19] where the estimated tensor has rank one. The model reads as follows: draw $K$ column vectors in $\mathbb{R}^{n}$, evaluate for each of them their $p^{\text {th }}$ tensor power, sum these $K$ symmetric order- $p$ tensors (this sum is exactly a $K$-term polyadic decomposition), and finally add noise to each entry of this symmetric order- $p$ tensor. Tensor factorization can then be studied as an inference problem where the task is to estimate the initial $K$ vectors that produce the informative tensor from the noisy observations of this tensor. We want to determine information theoretic limits for this task. To do so, we focus on proving variational formulas for the asymptotic normalized mutual information between the noisy observed tensor and the original $K$ vectors. Such formulas were first rigorously derived for $p=2$ and $K=1$, i.e., rank-one matrix factorization; see [66] for the case with a binary input vector, [18] for the restricted case in which no discontinuous phase transition occurs, [21] for a
single-sided bound, and finally [34] for the fully general case. The proof in 34 combines interpolation techniques with spatial coupling and an analysis of the approximate message passing (AMP) algorithm. Later, and still for $p=2, ~ \sqrt{36}$ went beyond rank-one by using a rigorous version of the cavity method. In 45 the authors applied the heuristic replica method to conjecture a formula for any $p$ and finite $K$, and proved its exactness when $p \geq 2$ and $K=1$. They also detail the AMP algorithm for tensor factorization and show how the single-letter variational expression for the mutual information allows one to give guarantees on the performance of AMP. Afterwards, [37], [87] introduced the adaptive interpolation proof technique which they applied to the case $p \geq 2, K=1$. Other proofs based on interpolations recently appeared; see [35] where $p=2, K=1$, and [106] where $p \geq 2, K=1$.

In this chapter, we prove the conjectured replica formula for any finite rank $K$ and any even order $p$ using the adaptive interpolation method. We also underline what is missing to extend the proof to odd orders. While our proof outline is similar to [37], there are two important new ingredients. First, to establish a tight lower bound on the asymptotic normalized mutual information, we have to prove the regularity of a change of variable given by the solutions to an ordinary differential equation. This is nontrivial when the rank becomes greater than one. Second, the same bound requires one to prove the concentration of the overlap (a quantity that fully characterizes the system in the high-dimensional limit). When the rank $K$ is greater than one, this overlap is a matrix. Proving its concentration requires new ideas and technical arguments to bypass difficulties that are absent in the scalar case $K=1$.

The chapter is organized as follows. In Section 4.2 we set up the precise statistical model and state our main theorem which is a single-letter variational expression for the asymptotic normalized mutual information. The adaptive interpolation method is formulated in Section 4.3. In Section 4.4 we use this adaptive interpolation to prove that the variational expression is both an upper and lower bound on the asymptotic normalized mutual information. Sections 4.5 and 4.6 contain the new and essential results which allow to go from rank-one to finite-rank tensors. Finally, the difficulties encountered for odd-order tensors are discussed in the last section, that is, Section 4.7. The reader will find in Appendix 4.A a technical calculation which is new and crucial to our proof, while the content of Appendix 4.B is more classical.

### 4.2 Low-rank symmetric tensor estimation

Statistical model for tensor estimation We now describe the statistical model that we study in this chapter. Let $n, K$ be positive integers and $P_{X}$ a probability distribution on $\mathbb{R}^{K}$. Let $X_{1}, \ldots, X_{n}$ be random column vectors in $\mathbb{R}^{K}$, independent and identically distributed (i.i.d.) with respect to $P_{X}$. We define $\mathbf{X}$ the $n \times K$ matrix whose $j^{\text {th }}$ row is equal to $X_{j}^{\top}$,

$$
\mathbf{X}:=\left[\begin{array}{llll}
X_{1} & X_{2} & \ldots & X_{n} \tag{4.1}
\end{array}\right]^{\top} .
$$

We denote by $X_{., k} \in \mathbb{R}^{n}$ the $k^{\text {th }}$ column of $\mathbf{X}$. We are interested in the symmetric order- $p$ tensor $\sum_{k=1}^{K} X_{\cdot, k}^{\otimes p}$ whose rank is at most $K$. This tensor is not directly observed. Instead, for each $p$-tuple $\underline{\boldsymbol{i}}=\left(i_{1}, \ldots, i_{p}\right)$ with $1 \leq i_{1} \leq i_{2} \leq \cdots \leq i_{p} \leq n$, we are given the noisy observation

$$
\begin{equation*}
Y_{\underline{i}}:=\sqrt{\frac{\lambda(p-1)!}{n^{p-1}}} \sum_{k=1}^{K} X_{i_{1} k} X_{i_{2} k} \ldots X_{i_{p} k}+Z_{\underline{i}}, \tag{4.2}
\end{equation*}
$$

where $\lambda>0$ is known and akin to a signal-to-noise ratio (SNR), and the noises $\left\{Z_{\underline{i}}\right\}_{1 \leq i_{1} \leq \cdots \leq i_{p} \leq n}$ are independent standard Gaussian random variables. We use the indexed noises $\left\{Z_{\underline{i}}\right\}_{1 \leq i_{1} \leq \cdots \leq i_{p} \leq n}$ to define the symmetric order- $p$ tensor $\mathbf{Z}:=\left\{Z_{i}\right\}_{1 \leq i_{1}, \ldots, i_{p} \leq n}$, meaning that for any permutation $\pi:\{1, \ldots, p\} \rightarrow\{1, \ldots, p\}$ and any $p$-tuple $\left(i_{1}, i_{2}, \ldots, i_{p}\right)$ we have

$$
Z_{i_{\pi(1)} i_{\pi(2)} \ldots i_{\pi(p)}}=Z_{i_{1} i_{2} \ldots i_{p}} .
$$

We see that being given the components (4.2) is equivalent to observing the symmetric order- $p$ tensor

$$
\begin{equation*}
\mathbf{Y}:=\sqrt{\frac{\lambda(p-1)!}{n^{p-1}}} \sum_{k=1}^{K} X_{\cdot, k}^{\otimes p}+\mathbf{Z} . \tag{4.3}
\end{equation*}
$$

Mutual information in the high-dimensional regime Our main result is the proof of a formula for the normalized mutual information between $\mathbf{X}$ and $\mathbf{Y}$ in the high-dimension limit where $n \rightarrow+\infty$ while the rank $K$ is kept fixed. This formula is given as the optimization of a potential over the cone of $K \times K$ symmetric positive semidefinite matrices $S_{K}^{+}$. Let $\widetilde{Z} \sim \mathcal{N}\left(0, I_{K}\right)$ and $X \sim P_{X}$ be two independent random vectors. We define

$$
\begin{equation*}
\psi_{P_{X}}: S \in S_{K}^{+} \mapsto \mathbb{E} \ln \int d P_{X}(x) \exp \left(X^{\top} S x+\widetilde{Z}^{\top} \sqrt{S} x-\frac{x^{\top} S x}{2}\right) \tag{4.4}
\end{equation*}
$$

that is a convex Lipschitz-continuous function (see Lemma 2.4). For two matrices $A$ and $B$ having the same dimension, the Hadamard product $A \circ B$ is the matrix of same dimension with elements given by $(A \circ B)_{i j}=A_{i j} B_{i j}$. Note that, by the Schur Product Theorem [107], the Hadamard product of two matrices in $S_{K}^{+}$is also in $S_{K}^{+}$. We now use $\psi_{P_{X}}$ to define the potential function

$$
\begin{equation*}
\phi_{p, \lambda}: S \in S_{K}^{+} \mapsto \psi_{P_{X}}\left(\lambda S^{\circ(p-1)}\right)-\frac{\lambda(p-1)}{2 p} \sum_{\ell, \ell^{\prime}=1}^{K}\left(S^{\circ p}\right)_{\ell \ell^{\prime}} \tag{4.5}
\end{equation*}
$$

where $S^{\text {ok }}$ is the $k^{\text {th }}$ Hadamard power of $S$. We prove the replica symmetric formula conjectured in 45].

Theorem 4.1 (RS formula for the normalized mutual information). Assume that the positive integer $p$ is even and the first $2 p$ moments of the distribution $P_{X}$ are
finite. Let $\Sigma_{X}:=\mathbb{E}\left[X X^{\top}\right] \in S_{K}^{+}$be the matrix of second moments of a random vector $X \sim P_{X}$. Then,

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \mathbf{Y})}{n}=\frac{\lambda}{2 p} \sum_{\ell, \ell^{\prime}=1}^{K}\left(\Sigma_{X}^{\circ p}\right)_{\ell \ell^{\prime}}-\sup _{S \in S_{K}^{+}} \phi_{p, \lambda}(S) . \tag{4.6}
\end{equation*}
$$

Remark (From now on $\lambda=1$ ). We can reduce the proof of (4.6) to the case $\lambda=1$ by properly rescaling $P_{X}$. From now on, we set $\lambda=1$ and define $\phi_{p}:=\phi_{p, 1}$.

Before proving Theorem 4.1 we introduce important quantities, adopting the statistical mechanics terminology. We denote by $\mathcal{I}$ the subset of $p$-tuples

$$
\mathcal{I}:=\left\{\underline{\boldsymbol{i}}=\left(i_{1}, i_{2}, \ldots, i_{p}\right): 1 \leq i_{1} \leq i_{2} \leq \cdots \leq i_{p} \leq n\right\} .
$$

Given the observations $\mathbf{Y}$, we define for all $\mathbf{x} \in \mathbb{R}^{n \times K}$ :

$$
\begin{equation*}
\mathcal{H}_{n}(\mathbf{x} ; \mathbf{Y}):=\sum_{\underline{i} \in \mathcal{I}} \frac{(p-1)!}{2 n^{p-1}}\left(\sum_{\ell=1}^{K} \prod_{a=1}^{p} x_{i_{a} \ell}\right)^{2}-\sum_{\underline{i} \in \mathcal{I}} \sqrt{\frac{(p-1)!}{n^{p-1}}} Y_{\underline{i}} \sum_{\ell=1}^{K} \prod_{a=1}^{p} x_{i_{a} \ell} . \tag{4.7}
\end{equation*}
$$

We refer to $\mathcal{H}_{n}(\cdot ; \mathbf{Y})$ as the Hamiltonian. Using Bayes' rule, the posterior probability density function is

$$
d P(\mathbf{x} \mid \mathbf{Y}):=\frac{1}{\mathcal{Z}_{n}(\mathbf{Y})} e^{-\mathcal{H}_{n}(\mathbf{x} ; \mathbf{Y})} \prod_{j=1}^{n} d P_{X}\left(x_{j}\right)
$$

where $\mathcal{Z}_{n}(\mathbf{Y}):=\int e^{-\mathcal{H}_{n}(\mathbf{x} ; \mathbf{Y})} \prod_{j} d P_{X}\left(x_{j}\right)$ is the normalization factor. Finally, the average free entropy is

$$
\begin{equation*}
f_{n}:=\frac{\mathbb{E} \ln \mathcal{Z}_{n}(\mathbf{Y})}{n}, \tag{4.8}
\end{equation*}
$$

and is linked to the mutual information by the identity

$$
\begin{equation*}
\frac{I(\mathbf{X} ; \mathbf{Y})}{n}=\frac{1}{2 p} \sum_{\ell, \ell^{\prime}=1}^{K}\left(\Sigma_{X}^{\circ p}\right)_{\ell \ell^{\prime}}-f_{n}+O\left(n^{-1}\right) \tag{4.9}
\end{equation*}
$$

The quantity $O\left(n^{-1}\right)$ in (4.9) is such that $n O\left(n^{-1}\right)$ is bounded uniformly in $n$. Thanks to 4.9, Theorem 4.1 follows directly from the next two bounds on the asymptotic average free entropy.

Theorem 4.2 (Lower bound on asymptotic average free entropy). Assume that the positive integer $p$ is even and the first $2 p$ moments of the distribution $P_{X}$ are finite. Then,

$$
\begin{equation*}
\liminf _{n \rightarrow+\infty} f_{n} \geq \sup _{S \in S_{K}^{+}} \phi_{p}(S) \tag{4.10}
\end{equation*}
$$

Theorem 4.3 (Upper bound on asymptotic average free entropy). Assume that the positive integer $p$ is even and the distribution $P_{X}$ has bounded support. Then,

$$
\begin{equation*}
\limsup _{n \rightarrow+\infty} f_{n} \leq \sup _{S \in S_{K}^{+}} \phi_{p}(S) . \tag{4.11}
\end{equation*}
$$

Remark. The assumption on $P_{X}$ in Theorem 4.3 is stricter than the one in Theorem 4.1. Therefore, combining Theorem 4.2 and Theorem 4.3 only proves the limit (4.6) for a distribution $P_{X}$ that has bounded support. The generalization to a distribution $P_{X}$ having finite $2 p^{\text {th }}$ moments is done by approaching $P_{X}$ with distributions having bounded support, much as it is done in [36, Section 6.2.2].

### 4.3 Adaptive path interpolation

We prove Theorems 4.2 and 4.3 thanks to the adaptive interpolation method. Let us introduce a parameter $t \in[0,1]$. We interpolate from the original channel (4.2) at $t=0$ to decoupled channels at $t=1$. In between, we follow an interpolation path $R(\cdot, \epsilon):[0,1] \rightarrow S_{K}^{+}$that is a continuously differentiable function parametrized by a "small perturbation" $\epsilon \in S_{K}^{+}$such that $R(0, \epsilon)=\epsilon$. More precisely, for $t \in[0,1]$, we observe

$$
\left\{\begin{array}{lll}
Y_{\underline{i}}^{(t)} & :=\sqrt{\frac{(1-t)(p-1)!}{n^{p-1}}} \sum_{k=1}^{K} \prod_{a=1}^{p} X_{i_{a} k}+Z_{\underline{i}}, & \underline{i} \in \mathcal{I}  \tag{4.12}\\
\widetilde{Y}_{j}^{(t, \epsilon)} & :=\sqrt{R(t, \epsilon)} X_{j}+\widetilde{Z}_{j}, \quad j \in\{1,2, \ldots, n\}
\end{array},\right.
$$

where $\mathbf{X}, \mathbf{Z}$ are defined like in Section 4.2, and $\widetilde{Z}_{1}, \ldots, \widetilde{Z}_{n}$ are $K$-dimensional column vectors that are i.i.d. with respect to $\mathcal{N}\left(0, I_{K}\right)$ and independent of $(\mathbf{X}, \mathbf{Z})$. Let $\widetilde{\mathbf{Z}}:=\left[\begin{array}{llll}\widetilde{Z}_{1} & \widetilde{Z}_{2} & \ldots & \widetilde{Z}_{n}\end{array}\right]^{\top}$ be the $n \times K$ matrix whose $j^{\text {th }}$ row is $\widetilde{Z}_{j}^{\top}$. Observing the components (4.12) is the same than being given the symmetric order- $p$ tensor $\mathbf{Y}^{(t)}$ together with the matrix $\widetilde{\mathbf{Y}}^{(t, \epsilon)}$, where

$$
\left\{\begin{array}{ll}
\mathbf{Y}^{(t)} & :=\sqrt{\frac{(1-t)(p-1)!}{n^{p-1}}} \sum_{k=1}^{K} X_{\cdot, k}^{\otimes p}+\mathbf{Z} \\
\widetilde{\mathbf{Y}}^{(t, \epsilon)} & :=\sqrt{R(t, \epsilon)} \mathbf{X}^{\top}+\widetilde{\mathbf{Z}}^{\top}
\end{array} .\right.
$$

The Hamiltonian associated with the interpolating problem at a fixed $t$ reads

$$
\begin{align*}
& \mathcal{H}_{t, \epsilon}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right) \\
& :=\sum_{\underline{i} \in \mathcal{I}} \frac{(1-t)(p-1)!}{2 n^{p-1}}\left(\sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k}\right)^{2}-\sqrt{\frac{(1-t)(p-1)!}{n^{p-1}}} Y_{\underline{i}}^{(t)} \sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k} \\
& \quad+\sum_{j=1}^{n} \frac{x_{j}^{\top} R(t, \epsilon) x_{j}}{2}-\left(\widetilde{Y}_{j}^{(t, \epsilon)}\right)^{\top} \sqrt{R(t, \epsilon)} x_{j} . \tag{4.13}
\end{align*}
$$

The interpolating average free entropy is defined similarly to the original average free entropy (4.8), that is,

$$
\begin{equation*}
f_{n}(t, \epsilon):=\frac{1}{n} \mathbb{E} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right) \tag{4.14}
\end{equation*}
$$

with $\mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right):=\int e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)} \prod_{j=1}^{n} d P_{X}\left(x_{j}\right)$. By evaluating (4.14) at both extremes we find that

$$
\left\{\begin{array}{l}
f_{n}(0, \epsilon)=f_{n}+O(\|\epsilon\|)  \tag{4.15}\\
f_{n}(1, \epsilon)=\psi_{P_{X}}(R(1, \epsilon))
\end{array}\right.
$$

where $\|\cdot\|$ denotes the Euclidean norm and $O(\|\epsilon\|)$ is such that $|O(\|\epsilon\|)| \leq \frac{\operatorname{Tr}\left(\Sigma_{x}\right)}{2}\|\epsilon\|$. In order to deal with future computations, it is useful to denote by angular brackets $\langle-\rangle_{t, \epsilon}$ the expectation with respect to the posterior distribution of $\mathbf{X}$ given $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)$, i.e.,

$$
\begin{equation*}
\langle g(\mathbf{x})\rangle_{t, \epsilon}=\int g(\mathbf{x}) \frac{\left.e^{-\mathcal{H} t, \epsilon\left(\mathbf{x} ; \mathbf{Y}^{(t)}\right)} \tilde{\mathbf{Y}}^{(t, \epsilon)}\right)}{\mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}\right)} \prod_{j=1}^{n} d P_{X}\left(x_{j}\right) \tag{4.16}
\end{equation*}
$$

Let $f_{n}^{\prime}(\cdot, \epsilon)$ be the derivative of $f_{n}(\cdot, \epsilon)$. Combining 4.15) with the fundamental theorem of calculus $f_{n}(0, \epsilon)=f_{n}(1, \epsilon)-\int_{0}^{1} f_{n}^{\prime}(t, \epsilon) d t$ yields the following sum-rule of the adaptive path interpolation.

Proposition 4.4 (Sum-rule). Assume that the first $2 p$ moments of $P_{X}$ are finite. Let $R^{\prime}(\cdot, \epsilon)$ be the derivative of the interpolation path $R(\cdot, \epsilon)$ and $\mathbf{Q}:=\mathbf{x}^{\top} \mathbf{x} / n$ the $K \times K$ overlap matrix whose entries are $Q_{\ell \ell^{\prime}}:=\frac{1}{n} \sum_{j=1}^{n} x_{j \ell} X_{j \ell^{\prime}}$. Then,

$$
\begin{align*}
f_{n}=O(\|\epsilon\|) & +O\left(n^{-1}\right)+\psi_{P_{X}}(R(1, \epsilon)) \\
& +\frac{1}{2 p} \int_{0}^{1} d t \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell^{\prime}}\right)^{p}\right\rangle_{t, \epsilon}-p\left(R^{\prime}(t, \epsilon)\right)_{\ell \ell^{\prime}} \mathbb{E}\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, \epsilon}, \tag{4.17}
\end{align*}
$$

where $O\left(n^{-1}\right) / n^{-1}$ and $O(\|\epsilon\|) /\|\epsilon\|$ are bounded uniformly in $\epsilon$ and $n$, respectively.
Proof. See Proposition 4.6 in Section 4.5 for the computation of $f_{n}^{\prime}(t, \epsilon)$.

### 4.4 Matching bounds

In this section we prove first Theorem 4.2 and then Theorem 4.3 by plugging two different choices for $R(\cdot, \epsilon)$ in the sum-rule (4.17).

### 4.4.1 Lower bound on the asymptotic average free entropy

Proof of Theorem 4.2. We obtain a lower bound on $f_{n}$ by choosing the interpolation path

$$
R(t, 0)=t S^{\circ(p-1)}
$$

where $S$ is a $K \times K$ symmetric positive semidefinite matrix. Note that $\epsilon=0$ and $R^{\prime}(t, \epsilon)=S^{\circ(p-1)}$. Under this choice, the sum-rule 4.17) reads

$$
\begin{equation*}
f_{n}=O\left(n^{-1}\right)+\phi_{p}(S)+\frac{1}{2 p} \int_{0}^{1} d t \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle h_{p}\left(S_{\ell \ell^{\prime}}, Q_{\ell \ell^{\prime}}\right)\right\rangle_{t, 0} \tag{4.18}
\end{equation*}
$$

where $h_{p}(r, q):=q^{p}-p q r^{p-1}+(p-1) r^{p}$. If $p$ is even then $h_{p}$ is nonnegative on $\mathbb{R}^{2}$ and (4.18) directly implies $f_{n} \geq \phi_{p}(S)+O\left(n^{-1}\right)$. Taking the inferior limit on both sides of this inequality, and bearing in mind that the inequality is valid for all $S \in S_{K}^{+}$, ends the proof of Theorem 4.2.

We have at our disposal a wealth of interpolation paths when considering any continuously differentiable $R(\cdot, \epsilon)$. However, to establish the lower bound (4.10), we have used a simple linear interpolation as $R^{\prime}(t, \epsilon)=S^{\circ(p-1)}$. Such an interpolation dates back to Guerra [60] and was already used by [36], [45] to derive the lower bound (4.10) for both cases $K=1$, any order $p$, and $p=2$, any finite rank $K$. Next we turn to the proof of the upper bound (4.11) and see how the flexibility in the choice of $R(\cdot, \epsilon)$ constitutes an improvement on the classical interpolation.

### 4.4.2 Matching upper bound

The sum-rule (4.17) suggests to pick an interpolation path satisfying

$$
\begin{equation*}
\forall\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}:\left(R^{\prime}(t, \epsilon)\right)_{\ell \ell^{\prime}}=\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, \epsilon}\right]^{p-1} \tag{4.19}
\end{equation*}
$$

The integral in (4.17) can then be split in two terms: one is similar to the second summand in 4.5), the other vanishes in the high-dimensional limit if the overlap concentrates. It is not obvious that there is an interpolation path satisfying (4.19) given that the angular brackets $\langle-\rangle_{t, \epsilon}$ themselves depend on $R(\cdot, \epsilon)$. Let us rewrite (4.19) explicitly as an ODE.

Let $R$ be a matrix in $S_{K}^{+}$and $t \in[0,1]$. Consider the observations

$$
\left\{\begin{array}{lll}
Y_{\underline{i}}^{(t)} & :=\sqrt{\frac{(1-t)(p-1)!}{n^{p-1}}} \sum_{k=1}^{K} \prod_{a=1}^{p} X_{i_{a} k}+Z_{\underline{i}}, & \underline{i} \in \mathcal{I}  \tag{4.20}\\
\widetilde{Y}_{j}^{(t, R)} & :=\sqrt{R} X_{j}+\widetilde{Z}_{j}, & j \in\{1, \ldots, n\}
\end{array},\right.
$$

where $\mathbf{X}, \mathbf{Z}$ and $\widetilde{\mathbf{Z}}$ are defined exactly like in the previous sections. These observations are reminiscent of the interpolating problem (4.12) and the Hamiltonian associated with (4.20) is equal to (4.13) with $R$ replacing $R(t, \epsilon)$. The angular brackets $\langle-\rangle_{t, R}$ denote the expectation with respect to the posterior distribution of $\mathbf{X}$. We define the function

$$
G_{n}: \begin{array}{cc}
{[0,1] \times S_{K}^{+}} & \rightarrow S_{K}^{+}  \tag{4.21}\\
(t, R) & \mapsto \mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]^{\circ(p-1)}
\end{array}
$$

Note that $\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}$ is a symmetric positive semidefinite matrix. Indeed, by the Nishimori identity, $\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}:=n^{-1} \mathbb{E}\left[\langle\mathbf{x}\rangle_{t, R}^{\top} \mathbf{X}\right]=n^{-1} \mathbb{E}\left[\langle\mathbf{x}\rangle_{t, R}^{\top}\langle\mathbf{x}\rangle_{t, R}\right]$. By the Schur product theorem [107], the Hadamard power $\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]^{\circ(p-1)}$ also belongs to $S_{K}^{+}$, justifying that $G_{n}$ takes its values in $S_{K}^{+}$. We see that the interpolation path $R(\cdot, \epsilon)$ satisfies 4.19) if, and only if, it is a solution to the $K(K+1) / 2$-dimensional first-order ODE

$$
R^{\prime}=G_{n}(t, R) \Leftrightarrow R^{\prime}=\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]^{\circ(p-1)} .
$$

The next proposition states that this ODE indeed admits a solution along. We also prove nontrivial properties that are needed to show the upper bound (4.11).

Proposition 4.5. For all $\epsilon \in S_{K}^{+}$, there exists a unique global solution, denoted $R(\cdot, \epsilon):[0,1] \rightarrow S_{K}^{+}$, to the initial value problem

$$
\begin{equation*}
R^{\prime}=G_{n}(t, R), R(0)=\epsilon . \tag{4.22}
\end{equation*}
$$

where $G_{n}$ is defined in (4.21). The function $R(\cdot, \epsilon)$ is continuously differentiable. Let $S_{K}^{++}$be the open cone of $K \times K$ symmetric positive definite matrices. If $p$ is even then, for all $t \in[0,1], R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $S_{K}^{++}$into $R\left(t, S_{K}^{++}\right)$whose Jacobian determinant is greater than one, i.e.,

$$
\begin{equation*}
\forall \epsilon \in S_{K}^{++}: \operatorname{det} J_{R(t, \cdot)}(\epsilon) \geq 1 \tag{4.23}
\end{equation*}
$$

Here $J_{R(t,)}$ denotes the Jacobian matrix of $R(t, \cdot)$.
Proof. $G_{n}$ is continuously differentiable on $[0,1] \times S_{K}^{+}$. Hence, by the PicardLindelöf theorem, for every $\epsilon \in S_{K}^{+}$, there exists a unique global solution $R(\cdot, \epsilon)$ to the initial value problem (4.22).

Each initial value $\epsilon \in S_{K}^{+}$is tied to a unique solution $R(\cdot, \epsilon)$, hence $\epsilon \mapsto R(t, \epsilon)$ is injective. It is also continuously differentiable [104, Theorem 3.1] with Jacobian determinant given by Liouville's formula [104, Corollary 3.1],

$$
\begin{equation*}
\operatorname{det} J_{R(t, \cdot)}(\epsilon)=\left.\exp \int_{0}^{t} d s \sum_{1 \leq \ell \leq \ell^{\prime} \leq K} \frac{\partial\left(G_{n}\right)_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right|_{s, R(s, \epsilon)} \tag{4.24}
\end{equation*}
$$

Thanks to the identity (4.24), we can show that the Jacobian determinant is greater than, or equal to, one by proving that the divergence $\left.\sum_{1 \leq \ell \leq \ell^{\prime} \leq K} \frac{\partial\left(G_{n}\right)_{\ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R}$ is nonnegative for all $(t, R) \in[0,1] \times S_{K}^{+}$. In the remaining part of the proof, we omit the subscripts of the angular brackets $\langle-\rangle_{t, R}$. By Lemma 4.13 in Appendix 4.A, we have

$$
\begin{equation*}
\left.\sum_{\ell \leq \ell^{\prime}} \frac{\partial\left(G_{n}\right)_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R}=n(p-1) \sum_{\ell, \ell^{\prime}} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\right]^{p-2} \Delta_{\ell \ell^{\prime}} \tag{4.25}
\end{equation*}
$$

where

$$
\begin{align*}
\Delta_{\ell \ell^{\prime}} & :=\mathbb{E}\left\langle\left(\frac{Q_{\ell \ell^{\prime}}+Q_{\ell^{\prime} \ell}}{2}-\left\langle\frac{Q_{\ell \ell^{\prime}}+Q_{\ell^{\prime} \ell}}{2}\right\rangle\right)^{2}\right\rangle \\
& -\mathbb{E}\left[\left(\left\langle\frac{Q_{\ell \ell^{\prime}}+Q_{\ell^{\prime} \ell}}{2}\right\rangle-\frac{\left(\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle\right)_{\ell \ell^{\prime}}}{n}\right)^{2}\right] . \tag{4.26}
\end{align*}
$$

The second expectation on the right-hand side (r.h.s.) of 4.26) satisfies

$$
\begin{aligned}
\mathbb{E}\left[\left(\left\langle\frac{Q_{\ell \ell^{\prime}}+Q_{\ell^{\prime} \ell}}{2}\right\rangle-\right.\right. & \left.\left.\frac{\left(\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle\right)_{\ell \ell^{\prime}}}{n}\right)^{2}\right] \\
& =\mathbb{E}\left\langle\frac{\left(\mathbf{x}^{\top} \mathbf{X}+\mathbf{X}^{\top} \mathbf{x}\right)_{\ell \ell^{\prime}}}{2 n}-\frac{\left(\langle\mathbf{x}\rangle^{\top} \mathbf{x}+\mathbf{x}^{\top}\langle\mathbf{x}\rangle\right)_{\ell \ell^{\prime}}}{2 n}\right\rangle^{2} \\
& \leq \mathbb{E}\left\langle\left(\frac{\left(\mathbf{x}^{\top} \mathbf{X}+\mathbf{X}^{\top} \mathbf{x}\right)_{\ell \ell^{\prime}}}{2 n}-\frac{\left(\langle\mathbf{x}\rangle^{\top} \mathbf{x}+\mathbf{x}^{\top}\langle\mathbf{x}\rangle\right)_{\ell \ell^{\prime}}}{2 n}\right)^{2}\right\rangle \\
& =\mathbb{E}\left\langle\left(\frac{\left(\mathbf{X}^{\top} \mathbf{x}+\mathbf{x}^{\top} \mathbf{X}\right)_{\ell \ell^{\prime}}}{2 n}-\frac{\left(\langle\mathbf{x}\rangle^{\top} \mathbf{X}+\mathbf{X}^{\top}\langle\mathbf{x}\rangle\right)_{\ell \ell^{\prime}}}{2 n}\right)^{2}\right\rangle \\
& =\mathbb{E}\left\langle\left(\frac{Q_{\ell^{\prime} \ell}+Q_{\ell \ell^{\prime}}}{2}-\left\langle\frac{Q_{\ell \ell^{\prime}}+Q_{\ell^{\prime} \ell}}{2}\right\rangle\right)^{2}\right\rangle
\end{aligned}
$$

where the inequality is a simple application of Jensen's inequality while the subsequent equality is due to the Nishimori identity. Note that the final upper bound is nothing but the first expectation on the r.h.s. of (4.26), hence $\Delta_{\ell \ell^{\prime}} \geq 0$ for all $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$. Besides, $\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, R}\right]^{p-2}$ is nonnegative if $p$ is even. The divergence (4.25) is thus nonnegative.

We can now prove Theorem 4.3 by choosing interpolation paths that are solutions to the ODE $R^{\prime}=G_{n}(t, R)$.

Proof of Theorem 4.3. Let $\epsilon$ be a symmetric positive definite matrix, i.e., $\epsilon \in S_{K}^{++}$. We interpolate with the unique solution $R(\cdot, \epsilon):[0,1] \mapsto S_{K}^{++}$to the initial value problem (4.22). Hence, the interpolation path $R(\cdot, \epsilon)$ satisfies 4.19) and the sum-rule (4.17) reads

$$
\begin{align*}
& f_{n}=O(\|\epsilon\|)+O\left(n^{-1}\right)+\psi_{P_{X}}(R(1, \epsilon))-\frac{p-1}{2 p} \sum_{\ell, \ell^{\prime}=1}^{K} \int_{0}^{1} d t \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, \epsilon}\right]^{p} \\
&+\int_{0}^{1} \frac{d t}{2 p} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell^{\prime}}\right)^{p}-\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, \epsilon}\right]^{p}\right\rangle_{t, \epsilon} . \tag{4.27}
\end{align*}
$$

Using first the Lipschitz continuity of $\psi_{P_{X}}$ and then its convexity (see Lemma 2.4), it comes

$$
\begin{align*}
\psi_{P_{X}}(R(1, \epsilon)) & =\psi_{P_{X}}\left(\epsilon+\int_{0}^{1} d t \mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, \epsilon}\right]^{\circ(p-1)}\right) \\
& =O(\|\epsilon\|)+\psi_{P_{X}}\left(\int_{0}^{1} d t \mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, \epsilon}\right]^{\circ(p-1)}\right) \\
& \leq O(\|\epsilon\|)+\int_{0}^{1} d t \psi_{P_{X}}\left(\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, \epsilon}\right]^{\circ(p-1)}\right), \tag{4.28}
\end{align*}
$$

where $|O(\|\epsilon\|)| \leq \frac{\operatorname{Tr} \Sigma_{X}}{2}\|\epsilon\|$. Combining (4.27) and (4.28) yields

$$
\begin{align*}
f_{n} & \leq O\left(n^{-1}\right)+O(\|\epsilon\|)+\int_{0}^{1} d t \phi_{p}\left(\mathbb{E}\langle\mathbf{Q}\rangle_{t, \epsilon}\right)+\int_{0}^{1} d t v(t, \epsilon) \\
& \leq O\left(n^{-1}\right)+O(\|\epsilon\|)+\sup _{S \in S_{K}^{+}} \phi_{p}(S)+\int_{0}^{1} d t v(t, \epsilon) \tag{4.29}
\end{align*}
$$

where

$$
\begin{equation*}
v(t, \epsilon):=\frac{1}{2 p} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell^{\prime}}\right)^{p}-\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, \epsilon}\right]^{p}\right\rangle_{t, \epsilon} . \tag{4.30}
\end{equation*}
$$

In order to prove (4.11), we have to show that the integral $\int_{0}^{1} d t v(t, \epsilon)$ in the upper bound (4.29) vanishes when $n$ diverges. This is the case if the overlap matrix $\mathbf{Q}$ concentrates around its expectation $\mathbb{E}\langle\mathbf{Q}\rangle_{t, \epsilon}$. Indeed, using that

$$
\left(Q_{\ell \ell^{\prime}}\right)^{p}-\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, \epsilon}\right]^{p}=\left(Q_{\ell \ell^{\prime}}-\mathbb{E}\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, \epsilon}\right) \sum_{k=0}^{p-1} Q_{\ell \ell^{\prime}}^{p-1-k}\left(\mathbb{E}\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, \epsilon}\right)^{k}
$$

and that the $(4 p-4)^{\text {th }}$ moments of $P_{X}$ are finite, we show that there exists a constant $C_{X}$ depending only on $P_{X}$ such that

$$
\begin{equation*}
\left|\int_{0}^{1} d t v(t, \epsilon)\right| \leq \frac{C_{X}}{2} \int_{0}^{1} d t\left(\mathbb{E}\left\langle\left\|\mathbf{Q}-\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, \epsilon}\right]\right\|^{2}\right\rangle_{t, \epsilon}\right)^{1 / 2} \tag{4.31}
\end{equation*}
$$

However, proving that the r.h.s. of (4.31) vanishes is only possible after integrating on a well-chosen set of "small perturbations" $\epsilon$ (that are the initial conditions for the ODE (4.22). In essence, the integration over $\epsilon$ smoothens the phase transitions that might appear for particular choices of $\epsilon$ when $n$ goes to infinity. We now describe the set of perturbations on which to integrate.

Let $\left(s_{n}\right)_{n \in \mathbb{N}^{*}}$ be a decreasing sequence of real numbers in $(0,1)$ and define the sequence of subsets

$$
\mathcal{E}_{n}:=\left\{\begin{array}{l|l}
\epsilon \in \mathbb{R}^{K \times K} & \begin{array}{l}
\forall \ell \neq \ell^{\prime}: \epsilon_{\ell \ell^{\prime}}=\epsilon_{\ell^{\prime} \ell} \in\left[s_{n}, 2 s_{n}\right] \\
\forall \ell: \epsilon_{\ell \ell} \in\left[2 K s_{n},(2 K+1) s_{n}\right]
\end{array} \tag{4.32}
\end{array}\right\} .
$$

Those are subsets of symmetric strictly diagonally dominant matrices with positive diagonal entries, hence they are included in $S_{K}^{++}$(see [108, Corollary 7.2.3]). As $\mathcal{E}_{n}$ is a $K(K+1) / 2$-dimensionnal hypercube whose side has length $s_{n}$, its volume is $V_{\mathcal{E}_{n}}=s_{n}^{K(K+1) / 2}$.

Remember that for every $\epsilon \in \mathcal{E}_{n}$ the interpolation path is the unique solution $R(\cdot, \epsilon):[0,1] \mapsto S_{K}^{++}$to the initial value problem (4.22). For a fixed $t \in[0,1]$, using first Cauchy-Schwarz inequality and then a change of variable $\epsilon \rightarrow R=R(t, \epsilon)$ (justified by Proposition 4.5 that states that $\epsilon \mapsto R(t, \epsilon)$ is a $C^{1}$-diffeomorphism), we obtain

$$
\begin{align*}
& \int_{\mathcal{E}_{n}} \frac{d \epsilon}{V_{\mathcal{E}_{n}}}\left(\mathbb{E}\left\langle\left\|\mathbf{Q}-\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, \epsilon}\right]\right\|^{2}\right\rangle_{t, \epsilon}\right)^{1 / 2} \leq\left(\int_{\mathcal{E}_{n}} \frac{d \epsilon}{V_{\mathcal{E}_{n}}} \mathbb{E}\left\langle\left\|\mathbf{Q}-\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, \epsilon}\right]\right\|^{2}\right\rangle_{t, \epsilon}\right)^{1 / 2} \\
&=\left(\int_{\mathcal{R}_{n, t}} \frac{d R}{V_{\mathcal{E}_{n}}\left|\operatorname{det} J_{R(t,)}(\epsilon)\right|} \mathbb{E}\left\langle\left\|\mathbf{Q}-\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]\right\|^{2}\right\rangle_{t, R}\right)^{1 / 2} \\
&\left.\leq\left(\int_{\mathcal{R}_{n, t}} \frac{d R}{V_{\mathcal{E}_{n}}} \mathbb{E}\left\langle\| \mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right] \|^{2}\right\rangle_{t, R}\right)^{1 / 2} \tag{4.33}
\end{align*}
$$

where $\mathcal{R}_{n, t}:=R\left(t, \mathcal{E}_{n}\right)$ and the angular brackets $\langle-\rangle_{t, R}$ denote the expectation with respect to the posterior distribution associated with the inference problem (4.20). The last inequality is due to the lower bound (4.23) for the Jacobian determinant of $R(t, \cdot)$. It is easier to work with the convex hulls of $\mathcal{R}_{n, t}$ that we denote by $\mathrm{C}\left(\mathcal{R}_{n, t}\right)$. These convex hulls are uniformly bounded compact sets of $S_{K}^{++}$. Indeed, every $\mathcal{R}_{n, t}$ is compact and included in the convex set

$$
\begin{equation*}
\mathcal{B}\left(\Sigma_{X}, K, p\right):=\left\{S \in S_{K}^{++}:\|S\| \leq 4 K^{3 / 2}+\operatorname{Tr}\left(\Sigma_{X}\right)^{p-1}\right\} \tag{4.34}
\end{equation*}
$$

that does not depend on $n$ and $t$ (see point (i) of Lemma 4.8 in Section 4.6). The upper bound (4.33) together with the inclusion $\mathcal{R}_{n, t} \subseteq \mathrm{C}\left(\mathcal{R}_{n, t}\right)$ directly yields

$$
\begin{equation*}
\int_{\mathcal{E}_{n}} \frac{d \epsilon}{V_{\mathcal{E}_{n}}}\left(\mathbb{E}\left\langle\left\|\mathbf{Q}-\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, \epsilon}\right]\right\|^{2}\right\rangle_{t, \epsilon}\right)^{1 / 2} \leq\left(\int_{C\left(\mathcal{R}_{n, t}\right)} \frac{d R}{V_{\mathcal{E}_{n}}} \mathbb{E}\left\langle\left\|\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right\|^{2}\right\rangle_{t, R}\right)^{1 / 2} \tag{4.35}
\end{equation*}
$$

By Theorem 4.7 in Section 4.6, there exists a positive constant $C$ which depends only on $P_{X}, K$ and $p$ such that

$$
\begin{equation*}
\int_{\mathrm{C}\left(\mathcal{R}_{n, t}\right)} d R \mathbb{E}\left\langle\left\|\mathbf{Q}-\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]\right\|^{2}\right\rangle_{t, R} \leq \frac{C}{s_{n}^{3 / 2} n^{1 / 6}} \tag{4.36}
\end{equation*}
$$

Combining (4.31), 4.35 and (4.36) gives

$$
\begin{equation*}
\left|\int_{\mathcal{E}_{n}} \frac{d \epsilon}{V_{\mathcal{E}_{n}}} \int_{0}^{1} d t v(t, \epsilon)\right| \leq \frac{C_{X}}{2} \sqrt{\frac{C}{\left(s_{n}^{9+3 K(K+1)} n\right)^{1 / 6}}} . \tag{4.37}
\end{equation*}
$$

This last upper bound vanishes if we choose a sequence $\left(s_{n}\right)_{n \in \mathbb{N}^{*}}$ such that $s_{n} \rightarrow 0$ and $s_{n}^{9+3 K(K+1)} n \rightarrow+\infty$ when $n$ diverges. Let us pick $s_{n}=n^{-\alpha}$ with $0<\alpha<(9+3 K(K+1))^{-1}$. We integrate both sides of (4.29) over $\epsilon \in \mathcal{E}_{n}$, and use that both the upper bound (4.37) and $\left|\int_{\mathcal{E}_{n}} \frac{d \epsilon}{V_{\mathcal{E}_{n}}} O(\|\epsilon\|)\right|=O\left(\max _{\epsilon \in \mathcal{E}_{n}}\|\epsilon\|\right)=O\left(s_{n}\right)$ vanish, to finally obtain

$$
f_{n}=\int_{\mathcal{E}_{n}} \frac{d \epsilon}{V_{\mathcal{E}_{n}}} f_{n} \leq \sup _{S \in S_{K}^{+}} \phi_{p}(S)+o_{n}(1) .
$$

Passing to the limit superior in this inequality ends the proof of Theorem 4.3.

### 4.5 Derivative of the average interpolating free entropy

In order to prove the sum-rule in Proposition 4.4, we need to compute the derivative with respect to $t$ of the average interpolating free entropy (4.14).

Proposition 4.6 (Derivative of the average interpolating free entropy). Assume that the first $2 p$ moments of $P_{X}$ are finite. Let $R^{\prime}(\cdot, \epsilon)$ be the derivative of the interpolation path $R(\cdot, \epsilon)$ and $\mathbf{Q}:=\mathbf{x}^{\top} \mathbf{x} / n \in \mathbb{R}^{K \times K}$ the overlap matrix whose entries are $Q_{\ell \ell^{\prime}}:=\frac{1}{n} \sum_{j=1}^{n} x_{j \ell} X_{j \ell^{\prime}}$. The derivative of the average free entropy $f_{n}(\cdot, \epsilon)$ defined in 4.14) is $\forall t \in[0,1]$ :

$$
\begin{equation*}
f_{n}^{\prime}(t, \epsilon)=-\frac{1}{2 p} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\left\langle\left(Q_{\ell \ell^{\prime}}\right)^{p}\right\rangle_{t, \epsilon}\right]+\frac{\operatorname{Tr}\left(R^{\prime}(t, \epsilon) \mathbb{E}\langle\mathbf{Q}\rangle_{t, \epsilon}\right)}{2}+O\left(n^{-1}\right) \tag{4.38}
\end{equation*}
$$

where $n \cdot O\left(n^{-1}\right)$ is bounded uniformly in $n, t$ and $\epsilon$.
Proof. The conditional probability density function of $\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}\right)$ given $\mathbf{X}$ is

$$
\begin{equation*}
P_{\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon) \mid \mathbf{X}}}(\mathbf{y}, \widetilde{\mathbf{y}} \mid \mathbf{x})=\frac{1}{\sqrt{2 \pi}^{n K+|\mathcal{T}|}} \exp \left(-\mathcal{H}_{t, \epsilon}(\mathbf{x} ; \mathbf{y}, \widetilde{\mathbf{y}})-\frac{\|\widetilde{\mathbf{y}}\|^{2}}{2}-\sum_{\underline{i} \in \mathcal{I}} \frac{y_{\underline{i}}^{2}}{2}\right), \tag{4.39}
\end{equation*}
$$

where $\mathcal{H}_{t, \epsilon}^{\prime}(\mathbf{x} ; \mathbf{y}, \widetilde{\mathbf{y}})$ is obtained from (4.13) by replacing $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)$ by $(\mathbf{y}, \widetilde{\mathbf{y}})$. Therefore, the average interpolating free entropy reads

$$
\begin{equation*}
f_{n}(t, \epsilon)=\frac{1}{n} \mathbb{E}_{\mathbf{X}}\left[\int d \mathbf{y} d \widetilde{\mathbf{y}} \frac{e^{-\mathcal{H}_{t, \epsilon}(\mathbf{X} ; \mathbf{y}, \widetilde{\mathbf{y}})-\frac{\|\widetilde{\mathbf{y}}\|^{2}}{2}-\frac{1}{2} \sum_{i \in \mathcal{I}} y_{\dot{i}}^{2}}}{\sqrt{2 \pi}^{n K+|\mathcal{I}|}} \ln \mathcal{Z}_{t, \epsilon}(\mathbf{y}, \widetilde{\mathbf{y}})\right] . \tag{4.40}
\end{equation*}
$$

Differentiating 4.40 under the integral sign yields

$$
\begin{align*}
f_{n}^{\prime}(t, \epsilon)=- & \frac{1}{n} \mathbb{E}\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right] \\
& -\frac{1}{n} \mathbb{E}\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right\rangle_{t, \epsilon} \tag{4.41}
\end{align*}
$$

where

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}^{\prime}(\mathbf{x} ; \mathbf{y}, \widetilde{\mathbf{y}}):= & \sum_{\underline{i} \in \mathcal{I}}-\frac{(p-1)!}{2 n^{p-1}}\left(\sum_{\ell=1}^{K} \prod_{a=1}^{p} x_{i_{a} \ell}\right)^{2}+\frac{1}{2} \sqrt{\frac{(p-1)!}{(1-t) n^{p-1}}} y_{\underline{i}} \sum_{\ell=1}^{K} \prod_{a=1}^{p} x_{i_{a} \ell} \\
& +\sum_{j=1}^{n} \frac{1}{2} x_{j}^{\top} \frac{d R(t, \epsilon)}{d t} x_{j}-\left(\widetilde{y}_{j}\right)^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} x_{j} . \tag{4.42}
\end{align*}
$$

The definition (4.42) comes from differentiating $\mathcal{H}_{t, \epsilon}^{\prime}(\mathbf{x} ; \mathbf{y}, \widetilde{\mathbf{y}})$ with respect to $t$. Before diving further, note that

$$
\begin{align*}
& \frac{d R(t, \epsilon)}{d t}=\sqrt{R(t, \epsilon)} \frac{d \sqrt{R(t, \epsilon)}}{d t}+\frac{d \sqrt{R(t, \epsilon)}}{d t} \sqrt{R(t, \epsilon)}  \tag{4.43}\\
& \forall v \in \mathbb{R}^{K}: v^{\top} \sqrt{R(t, \epsilon)} \frac{d \sqrt{R(t, \epsilon)}}{d t} v=v^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} \sqrt{R(t, \epsilon)} v . \tag{4.44}
\end{align*}
$$

The identities (4.43) and (4.44) can be combined to further obtain

$$
\begin{equation*}
\forall v \in \mathbb{R}^{K}: v^{\top} \sqrt{R(t, \epsilon)} \frac{d \sqrt{R(t, \epsilon)}}{d t} v=\frac{1}{2} v^{\top} \frac{d R(t, \epsilon)}{d t} v . \tag{4.45}
\end{equation*}
$$

Evaluating 4.42) at $(\mathbf{x}, \mathbf{y}, \widetilde{\mathbf{y}})=\left(\mathbf{X}, \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)$ and making use of 4.45) gives

$$
\begin{align*}
& \mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}\right) \\
&= \sum_{\underline{i} \in \mathcal{I}} \frac{1}{2} \sqrt{\frac{(p-1)!}{(1-t) n^{p-1}}} Z_{\underline{i}} \sum_{\ell=1}^{K} \prod_{a=1}^{p} X_{i_{a} \ell} \\
&+\sum_{j=1}^{n} X_{j}^{\top}\left(\frac{1}{2} \frac{d R(t, \epsilon)}{d t}-\sqrt{R(t, \epsilon)} \frac{d \sqrt{R(t, \epsilon)}}{d t}\right) X_{j}-\widetilde{Z}_{j}^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} X_{j} \\
&= \sum_{\underline{i} \in \mathcal{I}} \frac{1}{2} \sqrt{\frac{(p-1)!}{(1-t) n^{p-1}}} Z_{\underline{i}} \sum_{\ell=1}^{K} \prod_{a=1}^{p} X_{i_{a} \ell}-\sum_{j=1}^{n} \widetilde{Z}_{j}^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} X_{j} . \tag{4.46}
\end{align*}
$$

By the Nishimori identity, we have

$$
\begin{equation*}
\mathbb{E}\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right\rangle_{t, \epsilon}=\mathbb{E} \mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)=0 \tag{4.47}
\end{equation*}
$$

where we use (4.46) and $\mathbb{E} Z_{\underline{i}}=\mathbb{E} \widetilde{Z}_{j}=0$ to get the last equality. Therefore, the expression 4.41) for $f_{n}^{\prime}(t, \epsilon)$ simplifies to

$$
\begin{align*}
& f_{n}^{\prime}(t, \epsilon)=-\frac{1}{n} \mathbb{E}[ \left.\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right] \\
&=-\frac{1}{2 n} \sqrt{\frac{(p-1)!}{(1-t) n^{p-1}} \sum_{\underline{i} \in \mathcal{I}} \sum_{\ell=1}^{K} \mathbb{E}\left[Z_{\underline{i}} \prod_{a=1}^{p} X_{i_{a} \ell} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right]} \\
&+\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}\left[\widetilde{Z}_{j}^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} X_{j} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right] . \tag{4.48}
\end{align*}
$$

We simplify the two kind of expectations appearing on the r.h.s. of (4.48) in the following paragraphs a) and $\mathbf{b}$ ).
a) A Gaussian integration by parts with respect to $Z_{\underline{i}}$ gives

$$
\begin{aligned}
\mathbb{E}\left[Z_{\underline{i}} \prod_{a=1}^{p} X_{i_{a} \ell} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right] & =\mathbb{E}\left[\prod_{a=1}^{p} X_{i_{a} \ell} \frac{\partial \ln \mathcal{Z}_{t, \epsilon}\left(\widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)}{\partial Z_{\underline{\boldsymbol{i}}}}\right] \\
& =\sqrt{\frac{(1-t)(p-1)!}{n^{p-1}}} \sum_{\ell^{\prime}=1}^{K} \mathbb{E}\left\langle\prod_{a=1}^{p} x_{i_{a} \ell^{\prime}} X_{i_{a} \ell}\right\rangle_{t, \epsilon}
\end{aligned}
$$

Summing the latter identity over $\ell \in\{1, \ldots, K\}$ and $\underline{\boldsymbol{i}} \in \mathcal{I}$ yields

$$
\begin{aligned}
&-\frac{1}{2 n} \sqrt{\frac{(p-1)!}{(1-t) n^{p-1}}} \sum_{\underline{i} \in \mathcal{I}} \sum_{\ell=1}^{K} \mathbb{E}\left[Z_{\underline{i}} \prod_{a=1}^{p} X_{i_{a} \ell} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right] \\
&=-\frac{(p-1)!}{2 n^{p}} \sum_{\underline{i} \in \mathcal{I}} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle\prod_{a=1}^{p} x_{i_{a} \ell^{\prime}} X_{i_{a} \ell}\right\rangle_{t, \epsilon} .
\end{aligned}
$$

Let us simplify the right-hand side of this last equality. Note that the sum over the $p$-tuples $\underline{\boldsymbol{i}}$ such that $i_{1}<\cdots<i_{p}$ is equal to the sum over any tuple whose elements are distinct divided by $p$ ! (the cardinality of the symmetric group of degree $p$ ). This is because the summand is symmetric with respect to any permutation of the indices $\left(i_{1}, \ldots, i_{p}\right)$. However, the sum is over $\boldsymbol{i} \in \mathcal{I}$ so we also need to account for the terms that correspond to $p$-tuples having common elements (meaning that $i_{a}=i_{a^{\prime}}$ for some $a \neq a^{\prime}$ ). There are $O\left(n^{p-1}\right)$ such terms and each summand is bounded uniformly in ( $n, t, \epsilon$ ) because the first $2 p$ moments of $P_{X}$ are finite. Hence, the sum of these terms divided by $n^{p}$ is only $O\left(n^{-1}\right)$ and we have

$$
\begin{aligned}
-\frac{1}{2 n} \sqrt{\frac{(p-1)!}{(1-t) n^{p-1}}} \sum_{\underline{i} \in \mathcal{I}} & \sum_{\ell=1}^{K} \mathbb{E}\left[Z_{\underline{i}} \prod_{a=1}^{p} X_{i_{a} \ell} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right] \\
& =O\left(n^{-1}\right)-\frac{(p-1)!}{2 n^{p} p!} \sum_{\underline{i} \in\{1, \ldots, n\}^{p}}^{n} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle\prod_{a=1}^{p} x_{i_{a} \ell^{\prime}} X_{i_{a} \ell}\right\rangle_{t, \epsilon}
\end{aligned}
$$

$$
\begin{equation*}
=O\left(n^{-1}\right)-\frac{1}{2 p} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell^{\prime}}\right)^{p}\right\rangle_{t, \epsilon} . \tag{4.49}
\end{equation*}
$$

b) A Gaussian integration by parts with respect to the entries of the standard Gaussian random vector $\widetilde{Z}_{j}$ gives

$$
\begin{align*}
& \mathbb{E}\left[\widetilde{Z}_{j}^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} X_{j} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)\right] \\
&=\sum_{\ell=1}^{K} \mathbb{E}\left[\left(\frac{d \sqrt{R(t, \epsilon)}}{d t} X_{j}\right)_{\ell} \frac{\partial \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}_{t}, \widetilde{\mathbf{Y}}_{t, \epsilon}\right)}{\partial \widetilde{Z}_{j \ell}}\right] \\
&=\sum_{\ell=1}^{K} \mathbb{E}\left[\left(\frac{d \sqrt{R(t, \epsilon)}}{d t} X_{j}\right)_{\ell}\left\langle\left(\sqrt{R(t, \epsilon)} x_{j}\right)_{\ell}\right\rangle_{t, \epsilon}\right] \\
&=\mathbb{E}\left[X_{j}^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} \sqrt{R(t, \epsilon)}\left\langle x_{j}\right\rangle_{t, \epsilon}\right] \tag{4.50}
\end{align*}
$$

The last expectation on the r.h.s. of 4.50 can be further simplified thanks to the Nishimory identity (first and last equalities below) and the identity 4.45) (second equality below). We get

$$
\begin{aligned}
\mathbb{E}\left[X_{j}^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} \sqrt{R(t, \epsilon)}\left\langle x_{j}\right\rangle_{t, \epsilon}\right]=\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, \epsilon}^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} \sqrt{R(t, \epsilon)}\left\langle x_{j}\right\rangle_{t, \epsilon}\right] \\
=\frac{1}{2} \mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, \epsilon}^{\top} \frac{d R(t, \epsilon)}{d t}\left\langle x_{j}\right\rangle_{t, \epsilon}\right]=\frac{1}{2} \mathbb{E}\left[X_{j}^{\top} \frac{d R(t, \epsilon)}{d t}\left\langle x_{j}\right\rangle_{t, \epsilon}\right] .
\end{aligned}
$$

Summing the latter over $j \in\{1, \ldots, n\}$ yields

$$
\begin{align*}
\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}\left[\widetilde{Z}_{j}^{\top} \frac{d \sqrt{R(t, \epsilon)}}{d t} X_{j} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}_{t}, \widetilde{\mathbf{Y}}_{t, \epsilon}\right)\right] & =\frac{1}{2 n} \sum_{j=1}^{n} \mathbb{E}\left[X_{j}^{\top} \frac{d R(t, \epsilon)}{d t}\left\langle x_{j}\right\rangle_{t, \epsilon}\right] \\
& =\frac{1}{2} \mathbb{E}\left\langle\operatorname{Tr}\left(\frac{d R(t, \epsilon)}{d t} \frac{\mathbf{x}^{\top} \mathbf{X}}{n}\right)\right\rangle_{t, \epsilon} \\
& =\frac{1}{2} \operatorname{Tr}\left(R^{\prime}(t, \epsilon) \mathbb{E}\langle\mathbf{Q}\rangle_{t, \epsilon}\right) \tag{4.51}
\end{align*}
$$

Using the final expressions (4.49) and (4.51) back in (4.48) ends the proof.

### 4.6 Concentration of the overlap matrix

In the proof of Theorem 4.3 we need that, up to an integral over a small volume of perturbations $\epsilon \in S_{K}^{++}$, the overlap matrix $\mathbf{Q}$ concentrates around its expectation $\mathbb{E}\langle\mathbf{Q}\rangle_{t, \epsilon}$. We choose to integrate over the hypercube $\mathcal{E}_{n} \subseteq S_{K}^{++}$defined by 4.32) and that depends on a sequence $\left(s_{n}\right)_{n \in \mathbb{N}^{*}}$ of decreasing numbers in $(0,1)$.

Remember that we denote by the angular brackets $\langle-\rangle_{t, R}$ the expectation with respect to the posterior distribution

$$
\begin{equation*}
d P\left(\mathbf{x} \mid \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right):=\frac{1}{\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right)} e^{-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right)} \prod_{j=1}^{n} d P_{X}\left(x_{j}\right) \tag{4.52}
\end{equation*}
$$

where $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)$ are defined in (4.20) and

$$
\begin{aligned}
\mathcal{H}_{t, R}(\mathbf{x} ; \mathbf{y}, \widetilde{\mathbf{y}}):=\sum_{\underline{i} \in \mathcal{I}} & \frac{(1-t)(p-1)!}{2 n^{p-1}}\left(\sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k}\right)^{2}-\sqrt{\frac{(1-t)(p-1)!}{n^{p-1}}} y_{\underline{i}} \sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k} \\
& +\sum_{j=1}^{n} \frac{x_{j}^{\top} R x_{j}}{2}-\widetilde{y}_{j}^{\top} \sqrt{R} x_{j} .
\end{aligned}
$$

Then, for all $\epsilon \in \mathcal{E}_{n}$, we choose the interpolation path $R(\cdot, \epsilon):[0,1] \mapsto S_{K}^{++}$to be the unique solution to the initial value problem $R^{\prime}=\left(\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right)^{\circ(p-1)}, R(0)=\epsilon$. For a fixed $t \in[0,1]$, we denote by $C\left(\mathcal{R}_{n, t}\right)$ the convex hull of the image $\mathcal{R}_{n, t}:=R\left(t, \mathcal{E}_{n}\right)$ and we crucially rely on the following concentration result for the overlap matrix.

Theorem 4.7 (Concentration of the overlap matrix around its expectation). Assume that $P_{X}$ has bounded support. There exists a positive constant $C$ depending only on $P_{X}, K$ and $p$ such that

$$
\begin{equation*}
\int_{C\left(\mathcal{R}_{n, t}\right)} d R \mathbb{E}\left\langle\left\|\mathbf{Q}-\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]\right\|^{2}\right\rangle_{t, R} \leq \frac{C}{s_{n}^{3 / 2} n^{1 / 6}} \tag{4.53}
\end{equation*}
$$

The proof of Theorem 4.7 is inspired by [109, Theorem 3]. In the latter reference, the concentration result is given for an integral over a hypercube. In our case, the integral on the left-hand side of 4.53 is over the convex hull of $\mathcal{E}_{n}$ 's image by the function $R(t, \cdot)$. It is likely not a hypercube, even less one whose form is similar to 4.32). Therefore, we first show that the convex hulls $\mathrm{C}\left(\mathcal{R}_{n, t}\right)$ have properties that allow us to carry out a proof similar to (109.

### 4.6.1 Properties of $\mathcal{R}_{n, t}$ 's convex hull

For $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$, we denote by $E^{\left(\ell, \ell^{\prime}\right)}$ the $K \times K$ symmetric matrix whose entries are

$$
E_{i j}^{\left(\ell, \ell^{\prime}\right)}=\left\{\begin{array}{l}
1 \text { if }(i, j) \in\left\{\left(\ell, \ell^{\prime}\right),\left(\ell^{\prime}, \ell\right)\right\},  \tag{4.54}\\
0 \text { otherwise }
\end{array}\right.
$$

Lemma 4.8 (Properties of $\mathcal{R}_{n, t}$ 's convex hull). For every $R \in \mathrm{C}\left(\mathcal{R}_{n, t}\right)$ :
(i) $\|R\| \leq 4 K^{3 / 2}+\operatorname{Tr}\left(\Sigma_{X}\right)^{p-1}$;
(ii) there exists $\epsilon \in \mathcal{E}_{n}$ such that $R \succcurlyeq \epsilon$;
(iii) for every pair $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$ and real number $\delta \in\left(-s_{n}, s_{n}\right), R+\delta E^{\left(\ell, \ell^{\prime}\right)}$ is a symmetric positive definite matrix;
(iv) the $1^{s t}$-order Fréchet derivative $\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}$ and the $2^{\text {nd }}$-order Fréchet derivative $\frac{\partial^{2} \sqrt{R}}{\partial R_{\ell^{\prime}}^{2}}$ satisfy

$$
\begin{align*}
\left\|\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right\| & \leq \frac{1}{\sqrt{2 s_{n}}}  \tag{4.55}\\
\left\|\frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}}\right\| & \leq \frac{\sqrt{K}}{\left(2 s_{n}\right)^{3 / 2}} \tag{4.56}
\end{align*}
$$

Note that (i) does not depend on $n$ and $t$ while (ii-iv) do not depend on $t$.
Proof. We start by proving (i). If $R \in \mathcal{R}_{n, t}$ then there exists $\epsilon \in \mathcal{E}_{n}$ such that $R=R(t, \epsilon)$, i.e.,

$$
\begin{equation*}
R=\epsilon+\int_{0}^{t} \mathbb{E}\left[\langle\mathbf{Q}\rangle_{s, \epsilon}\right]^{\mathrm{o}(p-1)} d s \tag{4.57}
\end{equation*}
$$

The Frobenius norm is subadditive $(\|A+B\| \leq\|A\|+\|B\|)$ and submultiplicative with respect to the Hadamard product $(\|A \circ B\| \leq\|A\| \cdot\|B\|)$ so

$$
\|R\| \leq\|\epsilon\|+\int_{0}^{t}\left\|\mathbb{E}\left[\langle\mathbf{Q}\rangle_{s, \epsilon}\right]^{\circ(p-1)}\right\| d s \leq 4 K^{3 / 2}+\int_{0}^{t}\left\|\mathbb{E}\langle\mathbf{Q}\rangle_{s, \epsilon}\right\|^{p-1} d s
$$

Besides,

$$
\begin{align*}
\left\|\mathbb{E}\langle\mathbf{Q}\rangle_{s, \epsilon}\right\| \leq \mathbb{E}\langle\|\mathbf{Q}\|\rangle_{s, \epsilon} & \leq \frac{\mathbb{E}\langle\|\mathbf{x}\|\|\mathbf{X}\|\rangle_{s, \epsilon}}{n} \\
& \leq \sqrt{\frac{\mathbb{E}\|\mathbf{X}\|^{2}}{n} \frac{\mathbb{E}\left\langle\|\mathbf{x}\|^{2}\right\rangle_{s, \epsilon}}{n}}=\frac{\mathbb{E}\|\mathbf{X}\|^{2}}{n}=\operatorname{Tr}\left(\Sigma_{X}\right) \tag{4.58}
\end{align*}
$$

where the first and third inequalities follow from Cauchy-Schwarz inequality, the second inequality from the submultiplicativity of the Frobenius norm with respect to the ordinary matrix product, and the first equality from the Nishimori identity. Hence, $\|R\| \leq 4 K^{3 / 2}+\operatorname{Tr}\left(\Sigma_{X}\right)^{p-1}$ for all $R \in \mathcal{R}_{n, t}$. The bound directly extends to $C\left(\mathcal{R}_{n, t}\right)$ by definition of a convex hull.

We now prove (ii). If $R \in \mathcal{R}_{n, t}, 4.57$ ) directly implies that $R-\epsilon \succcurlyeq 0$ because, by the Nishimori identity and the Schur product theorem, $\mathbb{E}\left[\langle\mathbf{Q}\rangle_{s, \epsilon}\right]^{\circ(p-1)} \in S_{K}^{+}$for all $s \in[0,1]$. If instead $R \in C\left(\mathcal{R}_{n, t}\right)$, there exist $m \in \mathbb{N}^{*},\left(\alpha_{1}, \alpha_{2}, \ldots, \alpha_{m}\right) \in[0,1]^{m}$ and $\left(R_{1}, \ldots, R_{m}\right) \in\left(\mathcal{R}_{n, t}\right)^{M} m$ such that $\sum_{j=1}^{m} \alpha_{j}=1$ and $R=\sum_{j=1}^{m} \alpha_{j} R_{j}$. It follows direcly that $R \succcurlyeq \sum_{j=1}^{m} \alpha_{j} \epsilon_{j}$ where $\forall j \in\{1, \ldots, m\}: \mathcal{E}_{n} \ni \epsilon_{j} \preccurlyeq R_{j}$. As $\mathcal{E}_{n}$ is convex, it concludes the proof of (ii).

We now show (ii) $\Rightarrow$ (iii). Let $R \in \mathrm{C}\left(\mathcal{R}_{n, t}\right)$ and pick $\epsilon \in \mathcal{E}_{n}$ such that $R \succcurlyeq \epsilon$. For all $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$ and $\delta \in\left(-s_{n}, s_{n}\right), \epsilon+\delta E^{\left(\ell, \ell^{\prime}\right)}$ is a symmetric strictly diagonally dominant matrix with positive diagonal entries. Therefore, $\epsilon+\delta E^{\left(\ell, \ell^{\prime}\right)}$ belongs to $S_{K}^{++}$and $R+\delta E^{\left(\ell, \ell^{\prime}\right)} \succcurlyeq \epsilon+\delta E^{\left(\ell, \ell^{\prime}\right)} \succ 0$.

Finally, we prove (iv). Let $R \in \mathrm{C}\left(\mathcal{R}_{n, t}\right)$ and denote $\lambda_{\text {min }}(R)$ its minimum eigenvalue. Applying [110, Theorem 1.1] (the first upper bound in (6) to be more precise), we obtain

$$
\begin{equation*}
\left\|\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right\| \leq \frac{\left\|E^{\left(\ell, \ell^{\prime}\right)}\right\|}{2 \sqrt{\lambda_{\min }(R)}} ;\left\|\frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}}\right\| \leq \frac{\sqrt{K}\left\|E^{\left(\ell, \ell^{\prime}\right)}\right\|}{4 \lambda_{\min }(R)^{3 / 2}} \tag{4.59}
\end{equation*}
$$

Thanks to (ii) we can pick $\epsilon \in \mathcal{E}_{n}$ such that $R \succcurlyeq \epsilon$. By [111, Corollary 2], the minimum eigenvalue of $\epsilon$ is greater than $\sqrt{\alpha \beta}$ where

$$
\alpha=\min _{1 \leq k \leq K}\left\{\left|\epsilon_{k k}\right|-\sum_{j \neq k}\left|\epsilon_{k j}\right|\right\} \geq s_{n} \quad \text { and } \quad \beta=\min _{1 \leq k \leq K}\left\{\left|\epsilon_{k k}\right|-\sum_{j \neq k}\left|\epsilon_{j k}\right|\right\} \geq s_{n}
$$

Hence, $\lambda_{\min }(R) \geq \sqrt{\alpha \beta} \geq s_{n}$ which together with (4.59) yields (iv).

### 4.6.2 Concentration of $\mathcal{L}$ around its expectation

The concentration of the overlap matrix around its expectation follows from the concentration of the $K \times K$ symmetric matrix $\mathcal{L} \equiv \mathcal{L}(R)$ whose entries are

$$
\begin{equation*}
\mathcal{L}_{\ell \ell^{\prime}}:=\frac{1}{n} \sum_{j=1}^{n} \frac{1}{2} x_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}-X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}-x_{j}^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \widetilde{Z}_{j} \tag{4.60}
\end{equation*}
$$

for every $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$. This is well-defined as long as $R \in S_{K}^{++}$. To prove concentration results for $\mathcal{L}$, it is useful to work with the free entropy $\ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right) / n$ where $\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right)$ is the normalization factor of the posterior distribution (4.52). In Appendix 4.B we prove that this free entropy concentrates around its expectation when $n \rightarrow+\infty$. We define

$$
\begin{aligned}
F_{n}(t, R) & :=\frac{1}{n} \ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right) \\
f_{n}(t, R) & :=\frac{1}{n} \mathbb{E}\left[\ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)\right]=\mathbb{E} F_{n}(t, R)
\end{aligned}
$$

Proposition 4.9 (Thermal fluctuations of $\mathcal{L}$ ). Assume that $P_{X}$ has finite fourth moments. There exists a positive constant $C$, depending only on $\Sigma_{X}, K$ and $p$, such that for all $(n, t) \in \mathbb{N}^{*} \times[0,1]$ :

$$
\begin{equation*}
\int_{C\left(\mathcal{R}_{n, t}\right)} d R \mathbb{E}\left\langle\left\|\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right\|^{2}\right\rangle_{t, R} \leq \frac{C}{n s_{n}} \tag{4.61}
\end{equation*}
$$

Proof. Fix $(n, t) \in \mathbb{N}^{*} \times[0,1]$. Note that $\forall R \in S_{K}^{++}, \forall\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$ :

$$
\begin{equation*}
\left.\frac{\partial f_{n}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R}=-\frac{1}{n} \mathbb{E}\left[\left\langle\frac{\partial \mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)}{\partial R_{\ell \ell^{\prime}}}\right\rangle_{t, R}\right]=-\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R} \tag{4.62}
\end{equation*}
$$

Further differentiating, we obtain

$$
\begin{align*}
\left.\frac{\partial^{2} f_{n}}{\partial R_{\ell \ell^{\prime}}^{2}}\right|_{t, R} & =\mathbb{E}\left[\left\langle\mathcal{L}_{\ell \ell^{\prime}} \frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle_{t, R}\right]-\mathbb{E}\left[\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R}\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle_{t, R}\right]-\mathbb{E}\left\langle\frac{\partial \mathcal{L}_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right\rangle_{t, R} \\
& =n \mathbb{E}\left\langle\left(\mathcal{L}_{\ell \ell^{\prime}}-\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\mathbb{E}\left\langle\frac{\partial \mathcal{L}_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right\rangle_{t, R} \tag{4.63}
\end{align*}
$$

Combining (4.63) and the identity (4.72) for $\mathbb{E}\left\langle\partial \mathcal{L}_{\ell \ell^{\prime}} / \partial R_{\ell \ell^{\prime}}\right\rangle_{t, R}$ (see Lemma 4.10 below the proof), it comes

$$
\begin{equation*}
\mathbb{E}\left\langle\left(\mathcal{L}_{\ell \ell^{\prime}}-\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R}=\left.\frac{1}{n} \frac{\partial^{2} f_{n}}{\partial R_{\ell \ell^{\prime}}^{2}}\right|_{t, R}+\frac{1}{n} \operatorname{Tr}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\left(\Sigma_{X}-\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right) \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right) . \tag{4.64}
\end{equation*}
$$

We start by upper bounding the integral over $C\left(\mathcal{R}_{n, t}\right)$ of the second summand on the right-hand side of (4.64). Thanks to the Nishimory identity, we see that $\Sigma_{X} \succcurlyeq \mathbb{E}\langle\mathbf{Q}\rangle_{t, R}$. Indeed,

$$
\begin{aligned}
\Sigma_{X}-\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}:=\frac{\mathbb{E}\left[\mathbf{X}^{\top} \mathbf{X}\right]-\mathbb{E}\left[\langle\mathbf{x}\rangle_{t, R}^{\top} \mathbf{X}\right]}{n} & =\frac{\mathbb{E}\left[\left\langle\mathbf{x}^{\top} \mathbf{x}\right\rangle_{t, R}\right]-\mathbb{E}\left[\langle\mathbf{x}\rangle_{t, R}^{\top}\langle\mathbf{x}\rangle_{t, R}\right]}{n} \\
& =\frac{\mathbb{E}\left\langle\left(\mathbf{x}-\langle\mathbf{x}\rangle_{t, R}\right)^{\top}\left(\mathbf{x}-\langle\mathbf{x}\rangle_{t, R}\right)\right\rangle_{t, R}}{n} \succcurlyeq 0
\end{aligned}
$$

Therefore, $\left(\partial \sqrt{R} / \partial R_{\ell \ell^{\prime}}\right)\left(\Sigma_{X}-\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right)\left(\partial \sqrt{R} / \partial R_{\ell \ell^{\prime}}\right)$ is symmetric positive semidefinite and the trace on the right-hand side of (4.64) satisfies

$$
\begin{aligned}
0 \leq \operatorname{Tr}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\left(\Sigma_{X}-\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right) \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right) \leq \operatorname{Tr}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \Sigma_{X} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right) & \leq\left\|\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right\|^{2}\left\|\Sigma_{X}\right\| \\
& \leq \frac{\left\|\Sigma_{X}\right\|}{2 s_{n}}
\end{aligned}
$$

where the last inequality is due to 4.55) in Lemma 4.8. Keep in mind that $C\left(\mathcal{R}_{n, t}\right)$ is included in the ball $\mathcal{B}\left(\Sigma_{X}, K, p\right)$ defined in 4.34). Therefore, there exists a positive constant $C_{1}$ depending only on $\Sigma_{X}, K$ and $p$ such that

$$
\begin{equation*}
\int_{C\left(\mathcal{R}_{n, t}\right)} \frac{d R}{n} \operatorname{Tr}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\left(\Sigma_{X}-\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right) \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right) \leq \frac{C_{1}}{n s_{n}} \tag{4.65}
\end{equation*}
$$

We now turn to upper bounding $\left.\int_{C\left(\mathcal{R}_{n, t}\right)} \frac{d R}{n} \frac{\partial^{2} f_{n}}{\partial R_{\ell \ell^{\prime}}^{2}}\right|_{t, R}$. Define the closed convex set

$$
\begin{equation*}
C^{\left(\ell, \ell^{\prime}\right)}:=\left\{\left\{R_{k k^{\prime}}\right\}_{\left(k, k^{\prime}\right) \neq\left(\ell, \ell^{\prime}\right),\left(\ell^{\prime}, \ell\right)} \mid R \in C\left(\mathcal{R}_{n, t}\right)\right\} . \tag{4.66}
\end{equation*}
$$

For every pair $(\widetilde{R}, r) \in C^{\left(\ell, \ell^{\prime}\right)} \times \mathbb{R}$, we denote by $\widetilde{R} \cup\{r\}$ the $K \times K$ symmetric matrix whose entries are given by

$$
(\widetilde{R} \cup\{r\})_{k k^{\prime}}=\left\{\begin{array}{cl}
\widetilde{R}_{k k^{\prime}} & \text { if }\left(k, k^{\prime}\right) \neq\left(\ell, \ell^{\prime}\right),\left(\ell^{\prime}, \ell\right)  \tag{4.67}\\
r & \text { otherwise }
\end{array}\right.
$$

Because $C\left(\mathcal{R}_{n, t}\right)$ is a closed convex, there exist two functions $a, b: C^{\left(\ell, \ell^{\prime}\right)} \rightarrow \mathbb{R}$ such that $\forall \widetilde{R} \in C^{\left(\ell, \ell^{\prime}\right)}$ :
(i) $a(\widetilde{R}) \leq b(\widetilde{R})$;
(ii) $\forall r \in[a(\widetilde{R}), b(\widetilde{R})]: \widetilde{R} \cup\{r\} \in C\left(\mathcal{R}_{n, t}\right)$;
(iii) $\forall r \in \mathbb{R} \backslash[a(\widetilde{R}), b(\widetilde{R})]: \widetilde{R} \cup\{r\} \notin C\left(\mathcal{R}_{n, t}\right)$.

Therefore,

$$
\begin{align*}
\left.\int_{C\left(\mathcal{R}_{n, t}\right)} \frac{d R}{n} \frac{\partial^{2} f_{n}}{\partial R_{\ell \ell^{\prime}}^{2}}\right|_{t, R} & =\left.\int_{C^{\left(\ell, \ell^{\prime}\right)}} \frac{d \widetilde{R}}{n} \int_{a(\widetilde{R})}^{b(\widetilde{R})} d r \frac{\partial^{2} f_{n}}{\partial R_{\ell \ell^{\prime}}^{2}}\right|_{t, \tilde{R} \cup\{r\}} \\
& =\int_{C^{\left(\ell, \ell^{\prime}\right)}} \frac{d \widetilde{R}}{n}\left(\left.\frac{\partial f_{n}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, \widetilde{R} \cup\{b(\widetilde{R})\}}-\left.\frac{\partial f_{n}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, \widetilde{R} \cup\{a(\widetilde{R})\}}\right) . \tag{4.68}
\end{align*}
$$

Note that $\forall R \in S_{K}^{++}$:

$$
\begin{equation*}
\left|\frac{\partial f_{n}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R}\left|=\left|\mathbb{E}\left\langle\mathcal{L}_{\ell^{\prime}}\right\rangle_{t, R}\right| \leq\left|\mathbb{E}\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, R}\right| \leq \operatorname{Tr} \Sigma_{X},\right. \tag{4.69}
\end{equation*}
$$

where the first inequality follow from the identity (4.71) (see Lemma 4.10 below the proof) and the second from (4.58). Combining (4.68) and (4.69) yields

$$
\begin{equation*}
\left.\left|\int_{C\left(\mathcal{R}_{n, t}\right)} \frac{d R}{n} \frac{\partial^{2} f_{n}}{\partial R_{\ell \ell^{\prime}}^{2}}\right|_{t, R} \right\rvert\, \leq \frac{C_{2}}{n} \tag{4.70}
\end{equation*}
$$

where $C_{2}$ is a positive constant that depends only on $\Sigma_{X}, K$ and $p$. Integrating (4.64) over $C\left(\mathcal{R}_{n, t}\right)$, making use of the upper bounds 4.65) and 4.70), and finally summing over $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$, ends the proof.

We relied on the following lemma in the proof of Proposition 4.9.
Lemma 4.10. Assume that $P_{X}$ has finite second moments. Let $\delta_{\ell \ell^{\prime}}$ be 0 if $\ell \neq \ell^{\prime}$ and 1 otherwise. Then, $\forall(t, R) \in[0,1] \times S_{K}^{++}, \forall\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$ :

$$
\begin{align*}
\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R} & =-\left(1-\delta_{\ell \ell^{\prime}} / 2\right) \mathbb{E}\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, R}  \tag{4.71}\\
\mathbb{E}\left\langle\frac{\partial \mathcal{L}_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right\rangle_{t, R} & =\operatorname{Tr}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\left(\Sigma_{X}-\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right) \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right) \tag{4.72}
\end{align*}
$$

Proof. Fix $(t, R) \in[0,1] \times S_{K}^{++}$. By definition of $\mathcal{L}$, we have $\forall\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$ :

$$
\begin{align*}
\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R}:=\frac{1}{n} \sum_{j=1}^{n} \frac{1}{2} \mathbb{E}\left[\left\langle x_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle_{t, R}\right] & -\mathbb{E}\left[X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle_{t, R}\right] \\
& -\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \widetilde{Z}_{j}\right] . \tag{4.73}
\end{align*}
$$

We simplify the last expectation on the right-hand side of (4.73) by integrating by parts with respect to the entries of the standard Gaussian random vector $\widetilde{Z}_{j}$,

$$
\begin{align*}
\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \widetilde{Z}_{j}\right] & =\mathbb{E}\left\langle x_{j}^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R} x_{j}\right\rangle_{t, R}-\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left\langle x_{j}\right\rangle_{t, R}\right] \\
& =\frac{1}{2} \mathbb{E}\left\langle x_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle_{t, R}-\frac{1}{2} \mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle_{t, R}\right] \\
& =\frac{1}{2} \mathbb{E}\left\langle x_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle_{t, R}-\frac{1}{2} \mathbb{E}\left[X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle_{t, R}\right] . \tag{4.74}
\end{align*}
$$

The second equality follows from (4.116) and the third from the Nishimori identity. Plugging (4.74) back in (4.73), and noticing that $\partial R / \partial R_{\ell \ell^{\prime}}=E^{\left(\ell, \ell^{\prime}\right)}$, yields

$$
\begin{aligned}
\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R}=-\frac{1}{2 n} \sum_{j=1}^{n} \mathbb{E}\left[X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle_{t, R}\right] & =-\frac{1}{2} \operatorname{Tr}\left(\frac{\partial R}{\partial R_{\ell \ell^{\prime}}} \mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right) \\
& =-\left(1-\delta_{\ell \ell^{\prime}} / 2\right) \mathbb{E}\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, R},
\end{aligned}
$$

that is, 4.71). We now turn to the proof of 4.72). We have

$$
\begin{align*}
\mathbb{E}\left\langle\frac{\partial \mathcal{L}_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right\rangle_{t, R} & =-\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{\top} \frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}} \widetilde{Z}_{j}\right] \\
& =\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{\top} \frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}} \sqrt{R}\left\langle x_{j}\right\rangle_{t, R}\right]-\mathbb{E}\left\langle x_{j}^{\top} \frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}} \sqrt{R} x_{j}\right\rangle_{t, R}, \tag{4.75}
\end{align*}
$$

where the second equality follows once again from a Gaussian integration by parts with respect to $\widetilde{Z}_{j}$. Note that for all $v \in \mathbb{R}^{K}$ :

$$
\begin{equation*}
v^{\top} \frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}} \sqrt{R} v=\frac{1}{2} v^{\top}\left(\sqrt{R} \frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}}+\frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}} \sqrt{R}\right) v=-v^{\top}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right)^{2} v \tag{4.76}
\end{equation*}
$$

because of the identity
$0=\frac{\partial^{2} R}{\partial R_{\ell \ell^{\prime}}^{2}}=\frac{\partial}{\partial R_{\ell \ell^{\prime}}}\left(\sqrt{R} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}+\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\right)=2\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right)^{2}+\sqrt{R} \frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}}+\frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}} \sqrt{R}$.
We use (4.76) in 4.75) to simplify $\mathbb{E}\left\langle\partial \mathcal{L}_{\ell \ell^{\prime}} / \partial R_{\ell^{\prime}}\right\rangle_{t, R}$ further,

$$
\begin{align*}
\mathbb{E}\left\langle\frac{\partial \mathcal{L}_{\ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right\rangle_{t, R} & =\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}\left\langle x_{j}^{\top}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right)^{2} x_{j}\right\rangle_{t, R}-\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{\top}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right)^{2}\left\langle x_{j}\right\rangle_{t, R}\right] \\
& =\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}\left[X_{j}^{\top}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right)^{2} X_{j}\right]-\mathbb{E}\left[X_{j}^{\top}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right)^{2}\left\langle x_{j}\right\rangle_{t, R}\right] \\
& =\operatorname{Tr}\left(\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right)^{2} \Sigma_{X}\right)-\operatorname{Tr}\left(\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right)^{2} \mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right) \\
& =\operatorname{Tr}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\left(\Sigma_{X}-\mathbb{E}\langle\mathbf{Q}\rangle_{t, R}\right) \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right) . \tag{4.77}
\end{align*}
$$

The second equality follows from the Nishimori identity.

Proposition 4.11 (Quenched fluctuations of $\mathcal{L}$ ). Assume that $P_{X}$ has bounded support. There exists a positive constant $C$, depending only on $P_{X}, K$ and $p$, such that for all $(n, t) \in \mathbb{N}^{*} \times[0,1]$ :

$$
\begin{equation*}
\int_{C\left(\mathcal{R}_{n, t}\right)} d R \mathbb{E}\left\langle\left\|\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right\|^{2}\right\rangle_{t, R} \leq \frac{C}{s_{n}^{3} n^{1 / 3}} \tag{4.78}
\end{equation*}
$$

Proof. Fix $(n, t) \in \mathbb{N}^{*} \times[0,1]$. For all $R \in S_{K}^{++}$and $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$, we have

$$
\begin{align*}
\left.\frac{\partial F_{n}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R} & =-\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R} ;  \tag{4.79}\\
\left.\frac{\partial^{2} F_{n}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R} & =n\left\langle\left(\mathcal{L}_{\ell \ell^{\prime}}-\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R}+\frac{1}{n} \sum_{j=1}^{n}\left\langle x_{j}\right\rangle_{t, R}^{\top} \frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}} \widetilde{Z}_{j} ;  \tag{4.80}\\
\left.\frac{\partial f_{n}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R} & =-\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R} ;  \tag{4.81}\\
\left.\frac{\partial^{2} f_{n}}{\partial R_{\ell \ell^{\prime}}^{2}}\right|_{t, R} & =n \mathbb{E}\left\langle\left(\mathcal{L}_{\ell \ell^{\prime}}-\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R}+\frac{1}{n} \sum_{j=1}^{n} \mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{\top} \frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}} \widetilde{Z}_{j}\right] . \tag{4.82}
\end{align*}
$$

By assumption there exists a nonnegative real number $B_{X}$ such that $\|X\| \leq B_{X}$ almost surely if $X \sim P_{X}$. Using the upper bound 4.56) in Lemma 4.8, the second term on the right-hand side of 4.80 can be upper bounded,

$$
\begin{equation*}
\left|\frac{1}{n} \sum_{j=1}^{n}\left\langle x_{j}\right\rangle_{t, R}^{\top} \frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}} \widetilde{Z}_{j}\right| \leq \frac{1}{n} \sum_{j=1}^{n}\left\|\left\langle x_{j}\right\rangle_{t, R}\right\|\left\|\widetilde{Z}_{j}\right\|\left\|\frac{\partial^{2} \sqrt{R}}{\partial R_{\ell \ell^{\prime}}^{2}}\right\| \leq \frac{B_{X} \sqrt{K}}{\left(2 s_{n}\right)^{3 / 2} n} \sum_{j=1}^{n}\left\|\widetilde{Z}_{j}\right\| \tag{4.83}
\end{equation*}
$$

From now on, we also fix $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$ as well as $\widetilde{R} \in C^{\left(\ell, \ell^{\prime}\right)}$, where $C^{\left(\ell, \ell^{\prime}\right)}$ is the closed convex set defined in 4.66). Remember that, for every real number $r, \widetilde{R} \cup\{r\}$ is the matrix defined by (4.67), and that there exist two functions $a, b: C^{\left(\ell, \ell^{\prime}\right)} \rightarrow \mathbb{R}$ such that $\forall \widetilde{R} \in C^{\left(\ell, \ell^{\prime}\right)}$ :
(i) $a(\widetilde{R}) \leq b(\widetilde{R})$;
(ii) $\forall r \in[a(\widetilde{R}), b(\widetilde{R})]: \widetilde{R} \cup\{r\} \in C\left(\mathcal{R}_{n, t}\right)$;
(iii) $\forall r \in \mathbb{R} \backslash[a(\widetilde{R}), b(\widetilde{R})]: \widetilde{R} \cup\{r\} \notin C\left(\mathcal{R}_{n, t}\right)$.

Besides, by property (iii) in Lemma 4.8, for every $r \in\left(a(\widetilde{R})-s_{n}, b(\widetilde{R})+s_{n}\right)$ the matrix $\widetilde{R} \cup\{r\}$ is in $S_{K}^{++}$. Thus, we can define for all $r \in\left(a(\widetilde{R})-s_{n}, b(\widetilde{R})+s_{n}\right)$ :

$$
\begin{aligned}
F(r) & :=F_{n}(t, \widetilde{R} \cup\{r\})+\frac{r^{2}}{2} \frac{B_{X} \sqrt{K}}{\left(2 s_{n}\right)^{3 / 2} n} \sum_{j=1}^{n}\left\|\widetilde{Z}_{j}\right\| \\
f(r) & :=f_{n}(t, \widetilde{R} \cup\{r\})+\frac{r^{2}}{2} \frac{B_{X} \sqrt{K}}{\left(2 s_{n}\right)^{3 / 2} n} \sum_{j=1}^{n} \mathbb{E}\left\|\widetilde{Z}_{j}\right\| .
\end{aligned}
$$

$F$ is convex on $\left(a(\widetilde{R})-s_{n}, b(\widetilde{R})+s_{n}\right)$ as it is twice differentiable with a nonnegative second derivative by 4.80) and 4.83). Obviously, $f: r \mapsto \mathbb{E} F(r)$ is convex too. Note that for all $r \in\left(a(R)-s_{n}, b(R)+s_{n}\right)$ :

$$
\begin{aligned}
F(r)-f(r) & =F_{n}(t, \widetilde{R} \cup\{r\})-f_{n}(t, \widetilde{R} \cup\{r\})+\frac{r^{2}}{2} \frac{B_{X} \sqrt{K}}{\left(2 s_{n}\right)^{3 / 2} n} \sum_{j=1}^{n}\left\|\widetilde{Z}_{j}\right\|-\mathbb{E}\left\|\widetilde{Z}_{j}\right\| \\
F^{\prime}(r)-f^{\prime}(r) & =-\left(\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \widetilde{R} \cup\{r\}}-\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \widetilde{R} \cup\{r\}}\right)+r \frac{B_{X} \sqrt{K}}{\left(2 s_{n}\right)^{3 / 2} n} \sum_{j=1}^{n}\left\|\widetilde{Z}_{j}\right\|-\mathbb{E}\left\|\widetilde{Z}_{j}\right\|
\end{aligned}
$$

It follows from Lemma 2.8 (applied to the convex functions $G=F, g=f$ ) and these last two identities that $\forall(r, \delta) \in[a(\widetilde{R}), b(\widetilde{R})] \times\left(0, s_{n}\right)$ :

$$
\begin{aligned}
& \left|\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \widetilde{R} \cup\{r\}}-\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \tilde{R} \cup\{r\}}\right| \\
& \leq|r| \frac{B_{X} \sqrt{K}}{\left(2 s_{n}\right)^{3 / 2} n}\left|\sum_{j=1}^{n}\left\|\widetilde{Z}_{j}\right\|-\mathbb{E}\left\|\widetilde{Z}_{j}\right\|\right|+C_{\delta}(r)+\frac{1}{\delta} \sum_{u \in\{-\delta, 0, \delta\}}|F(r+u)-f(r+u)| \\
& \quad \leq\left(|r|+\frac{3}{2} r^{2}\right) \frac{B_{X} \sqrt{K}}{\left(2 s_{n}\right)^{3 / 2} n}\left|\sum_{j=1}^{n}\left\|\widetilde{Z}_{j}\right\|-\mathbb{E}\left\|\widetilde{Z}_{j}\right\|\right|+C_{\delta}(r) \\
& \quad+\frac{1}{\delta} \sum_{u \in\{-\delta, 0, \delta\}}\left|F_{n}(t, \widetilde{R} \cup\{r+u\})-f_{n}(t, \widetilde{R} \cup\{r+u\})\right|,
\end{aligned}
$$

where $C_{\delta}(r):=f^{\prime}(r+\delta)-f^{\prime}(r-\delta)$ is nonnegative by convexity of $f$. Using the inequality $\left(\sum_{i=1}^{5} v_{i}\right)^{2} \leq 5 \sum_{i=1}^{5} v_{i}^{2}$, we obtain that $\forall(r, \delta) \in[a(\widetilde{R}), b(\widetilde{R})] \times\left(0, s_{n}\right)$ :

$$
\begin{align*}
& \mathbb{E}\left[\left(\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \widetilde{R} \cup\{r\}}-\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \widetilde{R} \cup\{r\}}\right)^{2}\right] \\
& \quad \leq 5\left(|r|+\frac{3}{2} r^{2}\right)^{2} \frac{B_{X}^{2} K}{2 s_{n}^{3} n^{2}} \mathbb{V a r}\left(\sum_{j=1}^{n}\left\|\widetilde{Z}_{j}\right\|\right)+5 C_{\delta}(r)^{2} \\
& \quad+\frac{5}{\delta^{2}} \sum_{u \in\{-\delta, 0, \delta\}} \mathbb{E}\left[\left(F_{n}(t, \widetilde{R} \cup\{r+u\})-f_{n}(t, \widetilde{R} \cup\{r+u\})\right)^{2}\right] . \tag{4.84}
\end{align*}
$$

The next step is to bound the integral of the three summands on the right-hand side of (4.84). Remember that $\forall r \in[a(\widetilde{R}), b(\widetilde{R})]: \widetilde{R} \cup\{r\} \in C\left(\mathcal{R}_{n, t}\right)$. By property (i) in Lemma 4.8, we have $\forall r \in[a(\widetilde{R}), b(\widetilde{R})]$ :

$$
\begin{equation*}
|r| \leq\|\widetilde{R} \cup\{r\}\| \leq 4 K^{3 / 2}+\operatorname{Tr}\left(\Sigma_{X}\right)^{p-1} \tag{4.85}
\end{equation*}
$$

Besides, the standard Gaussian random vectors $\widetilde{Z}_{j}, j \in\{1, \ldots, n\}$, are independent so $\operatorname{Var}\left(\sum_{j=1}^{n}\left\|\widetilde{Z}_{j}\right\|\right)=n K$. We conclude that there exists a positive constant $C_{1}$ depending only on $P_{X}, K$ and $p$ such that $\forall \delta \in\left(0, s_{n}\right)$ :

$$
\begin{equation*}
\int_{a(\widetilde{R})}^{b(\widetilde{R})} d r 5\left(|r|+\frac{3}{2} r^{2}\right)^{2} \frac{B_{X}^{2} K}{2 s_{n}^{3} n^{2}} \operatorname{Var}\left(\sum_{j=1}^{n}\left\|\widetilde{Z}_{j}\right\|\right) \leq \frac{C_{1}}{s_{n}^{3} n} . \tag{4.86}
\end{equation*}
$$

Note that $C_{\delta}(r)=\left|C_{\delta}(r)\right| \leq\left|f^{\prime}(r+\delta)\right|+\left|f^{\prime}(r-\delta)\right|$. For all $q \in\left(a(\widetilde{R})-s_{n}, b(\widetilde{R})+s_{n}\right)$ :

$$
\begin{align*}
\left|f^{\prime}(q)\right| & \leq\left|\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \widetilde{R} \cup\{q\}}\right|+|q| \frac{B_{X} \sqrt{K}}{\left(2 s_{n}\right)^{3 / 2} n} \sum_{j=1}^{n} \mathbb{E}\left\|\widetilde{Z}_{j}\right\| \\
& \leq \operatorname{Tr}\left(\Sigma_{X}\right)+\left(s_{n}+4 K^{3 / 2}+\operatorname{Tr}\left(\Sigma_{X}\right)^{p-1}\right) \frac{B_{X} K}{\left(2 s_{n}\right)^{3 / 2}} \leq \frac{\widetilde{C}_{2}}{s_{n}^{3 / 2}}, \tag{4.87}
\end{align*}
$$

where the second inequality follows from $\left|\underset{\sim}{\mathbb{E}}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \tilde{R} \cup\{q\}}\right| \leq \operatorname{Tr} \Sigma_{X}$ (see (4.69), (4.85) and $\mathbb{E}\left\|\widetilde{Z}_{j}\right\| \leq \mathbb{E}\left[\left\|\widetilde{Z}_{j}\right\|^{2}\right]^{1 / 2}=\sqrt{K}$, and $\widetilde{C}_{2}$ is a positive constant that depends only on $P_{X}, K$ and $p$. Thus, for the second summand, we obtain $\forall \delta \in\left(0, s_{n}\right)$ :

$$
\begin{align*}
\int_{a(\widetilde{R})}^{b(\widetilde{R})} d r C_{\delta}(r)^{2} & \leq \frac{2 \widetilde{C}_{2}}{s_{n}^{3 / 2}} \int_{a(\widetilde{R})}^{b(\widetilde{R})} d r C_{\delta}(r) \\
& =2 \widetilde{C}_{2} \frac{f(b(\widetilde{R})+\delta)-f(b(\widetilde{R})-\delta)-(f(a(\widetilde{R})+\delta)-f(a(\widetilde{R})-\delta))}{s_{n}^{3 / 2}} \\
& \leq \frac{8 \delta \widetilde{C}_{2}^{2}}{s_{n}^{3}} . \tag{4.88}
\end{align*}
$$

The last inequality is a simple application of the mean value theorem. We turn to the third and last summand. $\forall \widetilde{R} \in C^{\left(\ell, \ell^{\prime}\right)}, \forall(r, \delta) \in[a(\widetilde{R}), b(\widetilde{R})] \times\left(-s_{n}, s_{n}\right)$ :

$$
\begin{aligned}
\|\widetilde{R} \cup\{r+\delta\}\| & =\left\|\widetilde{R} \cup\{r\}+\delta E^{\left(\ell, \ell^{\prime}\right)}\right\| \\
& \leq\|\widetilde{R} \cup\{r\}\|+|\delta|\left\|E^{\left(\ell, \ell^{\prime}\right)}\right\| \leq 4 K^{3 / 2}+\operatorname{Tr}\left(\Sigma_{X}\right)^{p-1}+2 .
\end{aligned}
$$

This upper bound is uniform in $n$ and $t$. Hence, by Theorem 4.14 of Appendix 4.B. there exists a positive constant $C_{3}$ depending only on $P_{X}, K$ and $p$ such that $\forall R \in C^{\left(\ell, \ell^{\prime}\right)}, \forall(r, \delta) \in[a(\widetilde{R}), b(\widetilde{R})] \times\left(-s_{n}, s_{n}\right):$

$$
\begin{equation*}
\mathbb{E}\left[\left(F_{n}(t, \widetilde{R} \cup\{r+\delta\})-f_{n}(t, \widetilde{R} \cup\{r+\delta\})\right)^{2}\right] \leq \frac{C_{3}}{n} \tag{4.89}
\end{equation*}
$$

Using first (4.89) and then 4.85, we see that $\forall \delta \in\left(0, s_{n}\right)$ :

$$
\begin{align*}
\int_{a(\widetilde{R})}^{b(\widetilde{R})} d r \frac{5}{\delta^{2}} & \sum_{u \in\{-\delta, 0, \delta\}} \mathbb{E}\left[\left(F_{n}(t, \widetilde{R} \cup\{r+u\})-f_{n}(t, \widetilde{R} \cup\{r+u\})\right)^{2}\right] \\
& \leq \frac{15 C_{3}}{\delta^{2} n}(b(\widetilde{R})-a(\widetilde{R})) \leq \frac{30 C_{3}}{\delta^{2} n}\left(4 K^{3 / 2}+\operatorname{Tr}\left(\Sigma_{X}\right)^{p-1}+2\right) \tag{4.90}
\end{align*}
$$

We now choose $\delta=s_{n}^{3 / 2} \cdot n^{-\frac{1}{3}}$. As $s_{n} \in(0,1)$, this choice satisfies $\delta \in\left(0, s_{n}\right)$. Combining (4.84) with the three upper bounds (4.86), (4.88) and (4.90) shows the existence of a positive constant $C$ depending only on $P_{X}, K$ and $p$ such that

$$
\begin{equation*}
\int_{a(\widetilde{R})}^{b(\widetilde{R})} d r \mathbb{E}\left[\left(\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \tilde{R} \cup\{r\}}-\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \tilde{R} \cup\{r\}}\right)^{2}\right] \leq \frac{C}{s_{n}^{3} n^{1 / 3}} \tag{4.91}
\end{equation*}
$$

One important fact following from our analysis is that $C$ can be chosen independently of both $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$ and $\widetilde{R} \in C^{\left(\ell, \ell^{\prime}\right)}$. Therefore, for all $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$, we have

$$
\begin{align*}
& \int_{C\left(\mathcal{R}_{n, t}\right)} d R \mathbb{E}\left[\left(\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \tilde{R} \cup\{r\}}-\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \tilde{R} \cup\{r\}}\right)^{2}\right] \\
& =\int_{C^{\left(\ell, \ell^{\prime}\right)}} d \widetilde{R} \int_{a(\widetilde{R})}^{b(\widetilde{R})} d r \mathbb{E}\left[\left(\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \tilde{R} \cup\{r\}}-\mathbb{E}\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, \tilde{R} \cup\{r\}}\right)^{2}\right] \leq \frac{C V_{C^{\left(\ell, \ell^{\prime}\right)}}}{s_{n}^{3} n^{1 / 3}}, \tag{4.92}
\end{align*}
$$

where $V_{C^{\left(\ell, \ell^{\prime}\right)}}$ denotes the volume of $C^{\left(\ell, \ell^{\prime}\right)}$. Each of the $K(K+1) / 2$ sets $C^{\left(\ell, \ell^{\prime}\right)}$ is uniformly bounded in $n$ and $t$ so the theorem follows from summing (4.92) over $\left(\ell, \ell^{\prime}\right)$.

### 4.6.3 Concentration of Q around its expectation

We now use the concentration results for $\mathcal{L}$, that is, Propositions 4.9 and 4.11, to prove Theorem 4.7. Let us first state an intermediary result on the thermal fluctuations of $\mathbf{Q}$.

Proposition 4.12 (Concentration of the overlap matrix around its expectation). Assume that $P_{X}$ has finite fourth moments. There exists a positive constant $C$ depending only on $P_{X}, K$ and $p$ such that

$$
\begin{array}{r}
\int_{C\left(\mathcal{R}_{n, t}\right)} d R \mathbb{E}\left\langle\left\|\mathbf{Q}-\langle\mathbf{Q}\rangle_{t, R}\right\|^{2}\right\rangle_{t, R} \leq \frac{C}{\sqrt{s_{n} n}} \\
\int_{C\left(\mathcal{R}_{n, t}\right)} d R \mathbb{E}\left\langle\left\|\mathbf{Q}-\frac{\langle\mathbf{x}\rangle_{t, R}^{\top}\langle\mathbf{x}\rangle_{t, R}}{n}\right\|^{2}\right\rangle_{t, R} \leq \frac{C}{\sqrt{s_{n} n}} \tag{4.94}
\end{array}
$$

Proof. Fix $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$. Note that $\forall(t, R) \in[0,1] \times S_{K}^{++}$:

$$
\begin{gather*}
\mathbb{E}\left\langle\left(Q_{\ell \ell^{\prime}}-\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R}=\frac{1}{n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[X_{i \ell^{\prime}} X_{j \ell^{\prime}}\left(\left\langle x_{i \ell} x_{j \ell}\right\rangle_{t, R}-\left\langle x_{i \ell}\right\rangle_{t, R}\left\langle x_{j \ell}\right\rangle_{t, R}\right)\right] \\
\leq \sqrt{M_{X}}\left(\frac{1}{n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[\left(\left\langle x_{i \ell} x_{j \ell}\right\rangle_{t, R}-\left\langle x_{i \ell}\right\rangle_{t, R}\left\langle x_{j \ell}\right\rangle_{t, R}\right)^{2}\right]\right)^{1 / 2}, \tag{4.95}
\end{gather*}
$$

where $M_{X}:=\frac{1}{n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[X_{i \ell^{\prime}}^{2} X_{j \ell^{\prime}}^{2}\right]$ is finite under our assumptions. Differentiating with respect to $R_{\ell \ell}$ on both sides of $\mathbb{E}\left\langle\mathcal{L}_{\ell \ell}\right\rangle_{t, R}=-\frac{1}{2} \mathbb{E}\left\langle Q_{\ell \ell}\right\rangle_{t, R}$ (see Lemma 4.10), we obtain (see 109 for the detailed computation)

$$
\begin{align*}
& \frac{1}{n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[\left(\left\langle x_{i \ell} x_{j \ell}\right\rangle_{t, R}-\left\langle x_{i \ell}\right\rangle_{t, R}\left\langle x_{j \ell}\right\rangle_{t, R}\right)^{2}\right] \\
&=2 \mathbb{E}\left\langle\left(\mathcal{L}_{\ell \ell^{\prime}}-\left\langle\mathcal{L}_{\ell \ell^{\prime}}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\frac{2}{n} \mathbb{E}\left\langle\frac{\partial \mathcal{L}_{\ell \ell}}{\partial R_{\ell \ell}}\right\rangle_{t, R} \tag{4.96}
\end{align*}
$$

By Proposition 4.9 and the inequality (4.65) combined with (4.72), there exists a positive constant $C$ depending only on $P_{X}, K$ and $p$ such that

$$
\begin{equation*}
\int_{C\left(\mathcal{R}_{n, t}\right)} \frac{d R}{n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[\left(\left\langle x_{i \ell} x_{j \ell}\right\rangle_{t, R}-\left\langle x_{i \ell}\right\rangle_{t, R}\left\langle x_{j \ell}\right\rangle_{t, R}\right)^{2}\right] \leq \frac{C}{n s_{n}} . \tag{4.97}
\end{equation*}
$$

Combining (4.95), 4.97) and the Cauchy-Schwarz inequality yields

$$
\begin{equation*}
\int_{C\left(\mathcal{R}_{n, t}\right)} d R \mathbb{E}\left\langle\left(Q_{\ell \ell^{\prime}}-\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \leq \sqrt{\frac{M_{X} V_{C\left(\mathcal{R}_{n, t}\right)} C}{n s_{n}}}, \tag{4.98}
\end{equation*}
$$

where $V_{C\left(\mathcal{R}_{n, t}\right)}$ is the volume of $C\left(\mathcal{R}_{n, t}\right)$ and is bounded uniformly in $n$ and $t$ by (i) of Lemma 4.8. This ends the proof of (4.93). The inequality (4.94) is proved in a similar way (see [109]).

Proof of Theorem 4.7. To lighten notations we drop the subscripts of the angular brackets $\langle-\rangle_{t, R}$. The concentration of $\mathbf{Q}$ can be linked to the concentration of $\mathcal{L}$ by rewriting $\operatorname{Tr} \mathbb{E}\langle\mathbf{Q}(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle)\rangle$ properly. Thanks to the identity 4.71), we have

$$
\begin{equation*}
\operatorname{Tr}(\mathbb{E}[\langle\mathbf{Q}\rangle] \mathbb{E}[\langle\mathcal{L}\rangle])=-\frac{1}{2} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle Q_{\ell \ell^{\prime}}\right\rangle \operatorname{Tr}\left(\frac{\partial R}{\partial R_{\ell \ell^{\prime}}} \mathbb{E}\langle\mathbf{Q}\rangle\right) . \tag{4.99}
\end{equation*}
$$

Plugging the definition (4.60) of $\mathcal{L}$ in $\operatorname{Tr} \mathbb{E}\langle\mathbf{Q} \mathcal{L}\rangle$, and integrating by parts with respect to the standard Gaussian random vectors $\widetilde{Z}_{j}, j \in\{1, \ldots, n\}$, we find that

$$
\begin{equation*}
\operatorname{Tr} \mathbb{E}\langle\mathbf{Q} \mathcal{L}\rangle=\frac{1}{n} \sum_{\ell, \ell^{\prime}=1}^{K} \sum_{j=1}^{n} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}} x_{j}^{\top}\right\rangle \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left\langle x_{j}\right\rangle\right]-\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}} X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle\right] . \tag{4.100}
\end{equation*}
$$

Note that $\forall\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}, \forall j \in\{1, \ldots, n\}$ :

$$
\begin{align*}
& \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}} x_{j}^{\top}\right\rangle \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left\langle x_{j}\right\rangle\right] \\
& \quad=\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle x_{j}\right\rangle^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left\langle x_{j}\right\rangle\right]+\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\left(x_{j}-\left\langle x_{j}\right\rangle\right)^{\top}\right\rangle \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left\langle x_{j}\right\rangle\right] \\
& \quad=\mathbb{E}\left[\frac{\left\langle Q_{\ell \ell^{\prime}}\right\rangle}{2}\left\langle x_{j}\right\rangle^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right]+\mathbb{E}\left[\left\langle Q_{\ell^{\prime} \ell}\right\rangle\left(X_{j}-\left\langle x_{j}\right\rangle\right)^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left\langle x_{j}\right\rangle\right] . \tag{4.101}
\end{align*}
$$

The second equality follows from (4.116) for the first expectation and the Nishimori identity for the second expectation. Plugging (4.101) back in 4.100), we get

$$
\begin{array}{rl}
\operatorname{Tr} \mathbb{E}\langle\mathbf{Q} \mathcal{L}\rangle=-\sum_{\ell, \ell^{\prime}=1}^{K} & \mathbb{E}\left\langle Q_{\ell \ell^{\prime}} \operatorname{Tr}\left(\frac{\partial R}{\partial R_{\ell \ell^{\prime}}} \mathbf{Q}\right)\right\rangle+\frac{1}{2} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle \operatorname{Tr}\left(\frac{\partial R}{\partial R_{\ell \ell^{\prime}}} \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right] \\
& +\sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell^{\prime} \ell}\right\rangle \operatorname{Tr}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left(\langle\mathbf{Q}\rangle-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right)\right] . \tag{4.102}
\end{array}
$$

Subtracting (4.99) to 4.102, we obtain
$\operatorname{Tr} \mathbb{E}\langle\mathbf{Q}(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle)\rangle$

$$
\begin{equation*}
=B-A+\sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell^{\prime} \ell}\right\rangle \operatorname{Tr}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left(\langle\mathbf{Q}\rangle-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right)\right] \tag{4.103}
\end{equation*}
$$

where

$$
A:=\sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle Q_{\ell \ell^{\prime}} \operatorname{Tr}\left(\frac{\partial R}{\partial R_{\ell \ell^{\prime}}}(\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle)\right)\right\rangle
$$

and

$$
B:=\frac{1}{2} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle \operatorname{Tr}\left(\frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left(\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}-\mathbb{E}\langle\mathbf{Q}\rangle\right)\right)\right]
$$

Note that $\partial R / \partial R_{\ell \ell^{\prime}}=E^{\left(\ell, \ell^{\prime}\right)}$ where $E^{\left(\ell, \ell^{\prime}\right)}$ is defined in (4.54), hence

$$
\begin{aligned}
A= & \mathbb{E}\left\langle\operatorname{Tr}\left(\mathbf{Q}(\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle)^{\top}\right)\right\rangle+\mathbb{E}\langle\operatorname{Tr}(\mathbf{Q}(\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle))\rangle \\
& -\sum_{\ell=1}^{K} \mathbb{E}\left\langle Q_{\ell \ell}\left(Q_{\ell \ell}-\mathbb{E}\left\langle Q_{\ell \ell}\right\rangle\right)\right\rangle \\
= & \mathbb{E}\left\langle\|\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle\|^{2}\right\rangle+\mathbb{E}\langle\operatorname{Tr}(\mathbf{Q}(\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle))\rangle-\sum_{\ell=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell}-\mathbb{E}\left\langle Q_{\ell \ell}\right\rangle\right)^{2}\right\rangle, \\
B= & \mathbb{E}\left\langle\operatorname{Tr}\left(\mathbf{Q}\left(\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}-\mathbb{E}\langle\mathbf{Q}\rangle\right)\right)\right\rangle-\frac{1}{2} \sum_{\ell=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell \ell}\right\rangle\left(\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}-\mathbb{E}\langle\mathbf{Q}\rangle\right)_{\ell \ell}\right] .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
B-A= & -\mathbb{E}\left\langle\|\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle\|^{2}\right\rangle-\mathbb{E}\left\langle\operatorname{Tr}\left(\mathbf{Q}\left(\mathbf{Q}-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right)\right\rangle \\
& +\sum_{\ell=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell}-\mathbb{E}\left\langle Q_{\ell \ell}\right\rangle\right)^{2}\right\rangle-\frac{1}{2} \sum_{\ell=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell \ell}\right\rangle\left(\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}-\mathbb{E}\langle\mathbf{Q}\rangle\right)_{\ell \ell}\right] \\
= & -\mathbb{E}\left\langle\|\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle\|^{2}\right\rangle-\mathbb{E}\left\langle\operatorname{Tr}\left(\mathbf{Q}\left(\mathbf{Q}-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right)\right\rangle \\
& +\frac{1}{2} \sum_{\ell=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell}-\mathbb{E}\left\langle Q_{\ell \ell}\right\rangle\right)^{2}\right\rangle+\frac{1}{2} \sum_{\ell=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell \ell}\left(\mathbf{Q}-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)_{\ell \ell}\right\rangle\right] \\
= & -\mathbb{E}\left\langle\|\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle\|^{2}\right\rangle-\mathbb{E}\left\langle\operatorname{Tr}\left(\mathbf{Q}-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)^{2}\right\rangle \\
& +\frac{1}{2} \sum_{\ell=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell}-\mathbb{E}\left\langle Q_{\ell \ell}\right\rangle\right)^{2}\right\rangle+\frac{1}{2} \sum_{\ell=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell}-\left.\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right|_{\ell \ell}\right)^{2}\right\rangle
\end{aligned}
$$

Plugging this last identity back in (4.103) gives

$$
\begin{align*}
\mathbb{E}\langle\| \mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle & \left.\|^{2}\right\rangle-\frac{1}{2} \sum_{\ell=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell}-\mathbb{E}\left\langle Q_{\ell \ell}\right\rangle\right)^{2}\right\rangle \\
= & -\operatorname{Tr} \mathbb{E}\langle\mathbf{Q}(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle)\rangle-\mathbb{E}\left\langle\operatorname{Tr}\left(\mathbf{Q}-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)^{2}\right\rangle \\
& +\frac{1}{2} \sum_{\ell=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell}-\left.\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right|_{\ell \ell}\right)^{2}\right\rangle \\
& +\sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell^{\prime} \ell}\right\rangle \operatorname{Tr}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left(\langle\mathbf{Q}\rangle-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right)\right] . \tag{4.104}
\end{align*}
$$

On one hand,

$$
\begin{equation*}
\frac{1}{2} \mathbb{E}\left\langle\|\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle\|^{2}\right\rangle \leq \mathbb{E}\left\langle\|\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle\|^{2}\right\rangle-\frac{1}{2} \sum_{\ell=1}^{K} \mathbb{E}\left\langle\left(Q_{\ell \ell}-\mathbb{E}\left\langle Q_{\ell \ell}\right\rangle\right)^{2}\right\rangle \tag{4.105}
\end{equation*}
$$

On the other hand, by the Cauchy-Schwarz inequality we have

$$
\begin{align*}
-\operatorname{Tr} \mathbb{E}\langle\mathbf{Q}(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle)\rangle & \leq \sqrt{\mathbb{E}\left\langle\|\mathbf{Q}\|^{2}\right\rangle \mathbb{E}\left\langle\|\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle\|^{2}\right\rangle}  \tag{4.106}\\
-\mathbb{E}\left\langle\operatorname{Tr}\left(\mathbf{Q}-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)^{2}\right\rangle & \leq \mathbb{E}\left\langle\left\|\mathbf{Q}-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right\|^{2}\right\rangle \tag{4.107}
\end{align*}
$$

Let $M:=\sqrt{4 K^{2}+K^{1 / 2} \operatorname{Tr}\left(\Sigma_{X}\right)^{p-1}}$. Note that

$$
\begin{equation*}
\|\sqrt{R}\|=\sqrt{\operatorname{Tr} R}=\sqrt{\operatorname{Tr}\left(R \cdot I_{K}\right)} \leq \sqrt{\|R\|\left\|I_{K}\right\|}=\sqrt{K\|R\|} \leq M, \tag{4.108}
\end{equation*}
$$

where the last inequality is due to point (i) of Lemma 4.8. This upper bound together with (4.55) (see Lemma 4.8 yields $\left\|\frac{\partial \sqrt{R}}{\partial R_{\ell^{\prime}}} \sqrt{R}\right\| \leq\left\|\frac{\partial \sqrt{R}}{\partial R_{\ell^{\prime}}}\right\|\|\sqrt{R}\| \leq M / \sqrt{2 s_{n}}$. Then, by Cauchy-Schwarz inequality and Jensen's inequality,

$$
\begin{align*}
\sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\langle Q _ { \ell ^ { \prime } \ell } \rangle \operatorname { T r } \left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\right.\right. & \left.\left.\sqrt{R}\left(\langle\mathbf{Q}\rangle-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right)\right] \\
& \left.\leq \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\left|\left\langle Q_{\ell^{\prime}}\right\rangle\right\rangle\| \| \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\| \|\left\|\langle\mathbf{Q}\rangle-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right\|\right]\right] \\
& \leq \frac{M}{\sqrt{2 s_{n}}} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\langle | Q_{\ell^{\prime} \ell}| \rangle\left\|\langle\mathbf{Q}\rangle-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right\|\right] \\
& \leq \frac{M K}{\sqrt{2 s_{n}}}\left(\mathbb{E}\left\langle\|\mathbf{Q}\|^{2}\right\rangle \mathbb{E}\left\|\langle\mathbf{Q}\rangle-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right\|^{2}\right)^{1 / 2} \\
& \leq \frac{B K}{\sqrt{2 s_{n}}} \sqrt{\mathbb{E}\left\langle\|\mathbf{Q}\|^{2}\right\rangle \mathbb{E}\left\langle\|\mathbf{Q}-\langle\mathbf{Q}\rangle\|^{2}\right\rangle} \tag{4.109}
\end{align*}
$$

The last inequality is due to Jensen's inequality and the Nishimori identity,

$$
\mathbb{E}\left\|\langle\mathbf{Q}\rangle-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right\|^{2} \leq \mathbb{E}\left\langle\left\|\mathbf{Q}-\frac{\mathbf{x}^{\top}\langle\mathbf{x}\rangle}{n}\right\|^{2}\right\rangle=\mathbb{E}\left\langle\left\|\mathbf{Q}^{\top}-\langle\mathbf{Q}\rangle^{\top}\right\|^{2}\right\rangle=\mathbb{E}\left\langle\|\mathbf{Q}-\langle\mathbf{Q}\rangle\|^{2}\right\rangle .
$$

Combining the identity (4.104) together with the inequalities 4.105), 4.106), (4.107), and (4.109), gives

$$
\begin{align*}
\frac{\mathbb{E}\left\langle\|\mathbf{Q}-\mathbb{E}\langle\mathbf{Q}\rangle\|^{2}\right\rangle}{2} \leq C\left(\sqrt{\mathbb{E}\left\langle\|\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle\|^{2}\right\rangle}\right. & +\mathbb{E}\left\langle\left\|\mathbf{Q}-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right\|^{2}\right\rangle \\
& \left.+\sqrt{\frac{\mathbb{E}\left\langle\|\mathbf{Q}-\langle\mathbf{Q}\rangle\|^{2}\right\rangle}{s_{n}}}\right) \tag{4.110}
\end{align*}
$$

where $C$ is a positive constant depending only on $P_{X}, K$ and $p$. To end the proof of Theorem 4.7, we simply need to integrate both sides of 4.110) over $C\left(\mathcal{R}_{n, t}\right)$ and apply Propositions 4.9, 4.11, 4.12.

### 4.7 Conclusion and discussion for odd-order tensors

In this work, we have proved the conjectured replica formula for even-order symmetric tensors. It would be desirable to extend both Theorem 4.2 and Theorem 4.3 to the odd-order case. For the case $K=1$ we refer to 45]. For $K>1$, this is still an open problem and we now briefly discuss where our proofs fall short in this case.

Ideally, to extend Theorem 4.2 to an odd order $p$, we would show that the integral on the r.h.s. of 4.18), that is, $\int_{0}^{1} d t \sum_{\ell, \ell^{\prime}} \mathbb{E}\left\langle h_{p}\left(S_{\ell \ell^{\prime}}, Q_{\ell \ell^{\prime}}\right)\right\rangle_{t, 0}$ where $h_{p}(r, q):=q^{p}-p q r^{p-1}+(p-1) r^{p}$, is nonnegative. However, when $p$ is odd, $h_{p}$ is not nonnegative on its whole domain of definition. To say something about the integral, we have to take a Gibbs average $\langle-\rangle_{t, 0}$ of $Q_{\ell \ell^{\prime}}$ before applying $h_{p}$. To do so, we split the integral in two as follows:

$$
\begin{align*}
& \int_{0}^{1} d t \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle h_{p}\left(S_{\ell \ell^{\prime}}, Q_{\ell \ell^{\prime}}\right)\right\rangle_{t, 0} \\
&=\int_{0}^{1} d t \mathbb{E}\left\langle\operatorname{Tr} \mathbf{Q}^{\top}\left(\mathbf{Q}^{\circ(p-1)}-\left(\frac{\langle\mathbf{x}\rangle_{t, 0}^{\top}\langle\mathbf{x}\rangle_{t, 0}}{n}\right)^{\circ(p-1)}\right)\right\rangle_{t, 0} \\
&+\int_{0}^{1} d t \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E} h_{p}\left(S_{\ell \ell^{\prime}},\left.\frac{\langle\mathbf{x}\rangle_{t, 0}^{\top}\langle\mathbf{x}\rangle_{t, 0}}{n}\right|_{\ell \ell^{\prime}}\right) . \tag{4.111}
\end{align*}
$$

When $K=1$, both $\langle\mathbf{x}\rangle_{t, 0}^{\top}\langle\mathbf{x}\rangle_{t, 0}$ and $S:=S_{11}$ are nonnegative real numbers. The nonnegativity of $h_{p}(r, q)$ for $r, q \geq 0$ thus ensures that the second integral on the r.h.s. of (4.111) is nonnegative. Besides, we can cancel the first integral by
introducing a small perturbation $\epsilon$ on which we integrate (as was done in the proof of Theorem 4.3). This is how the lower bound is proved in 45. When $K>1$, we only know that $\langle\mathbf{x}\rangle_{t, 0}^{\top}\langle\mathbf{x}\rangle_{t, 0}$ and $S$ are symmetric positive semidefinite matrices; a priori nothing can be said on the sign of their individual entries. The problem remains if instead we write

$$
\left.\begin{array}{rl}
\int_{0}^{1} d t \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left\langle h_{p}\left(S_{\ell \ell^{\prime}}, Q_{\ell \ell^{\prime}}\right)\right\rangle_{t, 0}=\int_{0}^{1} & d t \mathbb{E}\langle
\end{array} \operatorname{Tr}\left(\mathbf{Q}^{\top}\left(\mathbf{Q}^{\circ(p-1)}-\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, 0}\right]^{\circ(p-1)}\right)\right)\right\rangle_{t, 0}, ~(4.112) ~ ? ~+\int_{0}^{1} d t \sum_{\ell, \ell^{\prime}=1}^{K} h_{p}\left(S_{\ell \ell^{\prime}}, \mathbb{E}\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, 0}\right) .
$$

While $\mathbb{E}\langle\mathbf{Q}\rangle_{t, 0}$ and $S$ are positive semidefinite, nothing can be said on the sign of their individual entries. Most probably, we need to consider the full sum over $\left(\ell, \ell^{\prime}\right)$ in order to determine the sign of the second integral on the r.h.s. of (4.112). Indeed, using $A \succcurlyeq B \succcurlyeq 0 \Rightarrow \forall k \in \mathbb{N}: A^{\circ k} \succcurlyeq B^{\circ k} \succcurlyeq 0$, we can show that $\sum_{\ell, \ell^{\prime}=1}^{K} h_{p}\left(S_{\ell \ell^{\prime}}, \mathbb{E}\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, 0}\right)$ is nonnegative if $S \succcurlyeq \mathbb{E}\langle\mathbf{Q}\rangle_{t, 0}$ or $\mathbb{E}\langle\mathbf{Q}\rangle_{t, 0} \succcurlyeq S$. As far as we can tell, it is not clear why such partial ordering between $S$ and $\mathbb{E}\langle\mathbf{Q}\rangle_{t, 0}$ (which itself depends on $S$ ) holds.

Regarding Theorem 4.3, the whole proof directly applies to $p$ odd if we can show that the divergence (4.25) is nonnegative. However, this is more difficult than for $p$ even. Indeed, while the $\Delta_{\ell \ell}$ 's are still nonnegative, it is not necessarily the case of $\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, R}\right]^{p-2}$ when $p-2$ is odd.

## Appendices

## 4.A Divergence of the function $G_{n}$

Remember that we denote by the angular brackets $\langle-\rangle_{t, R}$ the expectation with respect to the posterior distribution

$$
d P\left(\mathbf{x} \mid \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right):=\frac{1}{\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right)} e^{-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)} \prod_{j=1}^{n} d P_{X}\left(x_{j}\right)
$$

where $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)$ are defined in 4.20) and

$$
\begin{aligned}
\mathcal{H}_{t, R}(\mathbf{x} ; \mathbf{y}, \widetilde{\mathbf{y}}):=\sum_{\underline{i} \in \mathcal{I}} & \frac{(1-t)(p-1)!}{2 n^{p-1}}\left(\sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k}\right)^{2}-\sqrt{\frac{(1-t)(p-1)!}{n^{p-1}}} y_{\underline{i}} \sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k} \\
& +\sum_{j=1}^{n} \frac{x_{j}^{\top} R x_{j}}{2}-\widetilde{y}_{j}^{\top} \sqrt{R} x_{j} .
\end{aligned}
$$

In this appendix we prove a formula for the divergence of the function

$$
G_{n}: \begin{array}{cc}
{[0,1] \times S_{K}^{+}} & \rightarrow S_{K}^{+} \\
(t, R) & \mapsto \mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]^{\circ(p-1)}
\end{array}
$$

Lemma 4.13 (Divergence of $G_{n}$ ). Let $\delta_{\ell \ell^{\prime}}=0$ if $\ell \neq \ell^{\prime}, 1$ otherwise. For all pair $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$, we have

$$
\begin{align*}
\left.\frac{\partial\left(G_{n}\right)_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R}= & \frac{n(p-1)}{1+\delta_{\ell \ell^{\prime}}} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle_{t, R}\right]^{p-2}\left(\left.\mathbb{E}\left[\left\langle\mathbf{Q} \circ\left(\mathbf{Q}+\mathbf{Q}^{\top}-\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle_{t, R}\right)\right\rangle_{t, R}\right]\right|_{\ell \ell^{\prime}}\right. \\
& \left.-\left.\mathbb{E}\left[\left\langle\mathbf{Q}^{\top}\right\rangle_{t, R} \circ\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle_{t, R}-2 \frac{\langle\mathbf{x}\rangle_{t, R}^{\top}\langle\mathbf{x}\rangle_{t, R}}{n}\right)\right]\right|_{\ell \ell^{\prime}}\right) \tag{4.113}
\end{align*}
$$

Besides, the divergence of $G_{n}$ is

$$
\left.\sum_{1 \leq \ell \leq \ell^{\prime} \leq K} \frac{\partial\left(G_{n}\right)_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R}=n(p-1) \operatorname{Tr}\left(\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]^{\circ(p-2)} \boldsymbol{\Delta}\right)
$$

where

$$
\Delta:=\mathbb{E}\left[\left\langle\left(\frac{\mathbf{Q}+\mathbf{Q}^{\top}}{2}-\left\langle\frac{\mathbf{Q}+\mathbf{Q}^{\top}}{2}\right\rangle_{t, R}\right)^{\circ 2}\right\rangle_{t, R}-\left(\left\langle\frac{\mathbf{Q}+\mathbf{Q}^{\top}}{2}\right\rangle_{t, R}-\frac{\langle\mathbf{x}\rangle_{t, R}^{\top}\langle\mathbf{x}\rangle_{t, R}}{n}\right)^{\circ 2}\right]
$$

Proof. To lighten notations, we omit the subscripts of the angular brackets $\langle-\rangle_{t, R}$. Let $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$. The partial derivative of $R \mapsto\left(G_{n}(t, R)\right)_{\ell \ell^{\prime}}$ with respect to $R_{\ell \ell^{\prime}}$ reads

$$
\begin{align*}
\left.\frac{\partial\left(G_{n}\right)_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R} & =\left.\frac{\partial \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\right]^{p-1}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R} \\
& =(p-1) \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\right]^{p-2} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle-\left\langle Q_{\ell \ell^{\prime}} \frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle\right], \tag{4.114}
\end{align*}
$$

where

$$
\begin{equation*}
\frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}=\sum_{j=1}^{n} \frac{1}{2} x_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}-X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}-\widetilde{Z}_{j}^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j} . \tag{4.115}
\end{equation*}
$$

After we plug the r.h.s. of (4.115) in (4.114, two expectations involving the standard Gaussian randon vectors $\widetilde{Z}_{j}, j \in\{1, \ldots, n\}$, appear. A Gaussian integration by parts gives

$$
\begin{aligned}
\mathbb{E} & {\left[\left\langle Q_{\ell \ell^{\prime}} \widetilde{Z}_{j}^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle\right]=\sum_{k=1}^{K} \mathbb{E}\left[\widetilde{Z}_{j k}\left\langle Q_{\ell^{\prime}}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right)_{k}\right\rangle\right] } \\
& =\sum_{k=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\left(\sqrt{R} x_{j}\right)_{k}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right)_{k}\right\rangle\right]-\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right)_{k}\right\rangle\left\langle\left(\sqrt{R} x_{j}\right)_{k}\right\rangle\right] \\
& =\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}} x_{j}^{\top} \sqrt{R} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle\right]-\mathbb{E}\left[\left\langle Q_{\ell^{\prime}} x_{j}^{\top}\right\rangle \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left\langle x_{j}\right\rangle\right] \\
& =\frac{1}{2} \mathbb{E}\left[\left\langle Q_{\ell^{\prime}} x_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle\right]-\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}} x_{j}^{\top}\right\rangle \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\left\langle x_{j}\right\rangle\right],
\end{aligned}
$$

and

$$
\begin{aligned}
& \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle\widetilde{Z}_{j}^{\top} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle\right]=\sum_{k=1}^{K} \mathbb{E}\left[\widetilde{Z}_{j k}\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right)_{k}\right\rangle\right] \\
&=\sum_{k=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle\left(\sqrt{R} x_{j}\right)_{k}\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right)_{k}\right\rangle\right]-2 \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle\left(\sqrt{R} x_{j}\right)_{k}\right\rangle\left\langle\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right)_{k}\right\rangle\right] \\
&+\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\left(\sqrt{R} x_{j}\right)_{k}\right\rangle\left\langle\left(\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right)_{k}\right\rangle\right] \\
&=\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle x_{j}^{\top} \sqrt{R} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle\right]-2 \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle x_{j}\right\rangle^{\top} \sqrt{R} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right] \\
&+\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}} x_{j}^{\top}\right\rangle \sqrt{R} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right] \\
&=\frac{1}{2} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle x_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle\right]-\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle x_{j}\right\rangle^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right] \\
&+\mathbb{E}\left[\left\langle Q_{\ell^{\prime}} x_{j}^{\top}\right\rangle \sqrt{R} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right] .
\end{aligned}
$$

In both chains of equalities, the last one follows from an identity similar to 4.45,

$$
\begin{equation*}
\forall v \in \mathbb{R}^{K}: v^{\top} \sqrt{R} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} v=\frac{1}{2} v^{\top}\left(\sqrt{R} \frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}}+\frac{\partial \sqrt{R}}{\partial R_{\ell \ell^{\prime}}} \sqrt{R}\right) v=\frac{1}{2} v^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} v . \tag{4.116}
\end{equation*}
$$

Using (4.116) and the two identities yielded by the integration by parts, we get

$$
\begin{align*}
\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle-\right. & \left.\left\langle Q_{\ell \ell^{\prime}} \frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle\right] \\
=\sum_{j=1}^{n} \mathbb{E} & {\left[\left\langle Q_{\ell \ell^{\prime}} X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}} x_{j}\right\rangle-\left\langle Q_{\ell \ell^{\prime}}\right\rangle X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right] } \\
& +\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle x_{j}\right\rangle^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle-\left\langle Q_{\ell \ell^{\prime}} x_{j}^{\top}\right\rangle \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right] . \tag{4.117}
\end{align*}
$$

By the Nishimori identity, we have

$$
\begin{aligned}
\mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle x_{j}\right\rangle^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle-\right. & \left.\left\langle Q_{\ell \ell^{\prime}} x_{j}^{\top}\right\rangle \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right] \\
& =\mathbb{E}\left[\left\langle Q_{\ell^{\prime} \ell}\right\rangle\left\langle x_{j}\right\rangle^{\top} \frac{\partial R}{\partial R_{\ell^{\prime}}}\left\langle x_{j}\right\rangle-\left\langle Q_{\ell^{\prime} \ell}\right\rangle X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right] .
\end{aligned}
$$

Hence, (4.117) further simplifies,

$$
\begin{align*}
& \mathbb{E} {\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle-\left\langle Q_{\ell \ell^{\prime}} \frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle\right] } \\
&= \sum_{j=1}^{n} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}} X_{j}^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left(x_{j}-\left\langle x_{j}\right\rangle\right)\right\rangle\right]-\mathbb{E}\left[\left\langle Q_{\ell^{\prime} \ell}\right\rangle\left(X_{j}-\left\langle x_{j}\right\rangle\right)^{\top} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left\langle x_{j}\right\rangle\right] \\
&= \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}} \operatorname{Tr}\left(\mathbf{X} \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}(\mathbf{x}-\langle\mathbf{x}\rangle)^{\top}\right)\right\rangle\right]-\mathbb{E}\left[\left\langle Q_{\ell^{\prime} \ell}\right\rangle \operatorname{Tr}\left((\mathbf{X}-\langle\mathbf{x}\rangle) \frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\langle\mathbf{x}\rangle^{\top}\right)\right] \\
&=n \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}} \operatorname{Tr}\left(\frac{\partial R}{\partial R_{\ell \ell^{\prime}}}(\mathbf{Q}-\langle\mathbf{Q}\rangle)\right)\right\rangle\right] \\
&-n \mathbb{E}\left[\left\langle Q_{\ell^{\prime} \ell}\right\rangle \operatorname{Tr}\left(\frac{\partial R}{\partial R_{\ell \ell^{\prime}}}\left(\langle\mathbf{Q}\rangle-\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right)\right], \tag{4.118}
\end{align*}
$$

where in the last equality we use the cyclic property of the trace. Now consider the case $\ell \neq \ell^{\prime}$. All the entries of $\partial R / \partial R_{\ell \ell^{\prime}}$ are zeros save for the entries $\left(\ell, \ell^{\prime}\right)$ and $\left(\ell^{\prime}, \ell\right)$ which are both one. Then, equation (4.118) reads

$$
\begin{aligned}
& \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle-\left\langle Q_{\ell \ell^{\prime}} \frac{\partial \mathcal{H}_{t, R}}{\partial R_{\ell \ell^{\prime}}}\right\rangle\right] \\
= & n \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\left(\mathbf{Q}+\mathbf{Q}^{\top}-\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle\right)_{\ell \ell^{\prime}}\right\rangle\right]-n \mathbb{E}\left[\left\langle Q_{\ell^{\prime} \ell}\right\rangle\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle-2 \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)_{\ell \ell^{\prime}}\right] .
\end{aligned}
$$

Combining this last identity with 4.114 gives 4.113) for $\ell \neq \ell^{\prime}$. The case $\ell=\ell^{\prime}$ is obtained in a similar way except that now the entries of $\partial R / \partial R_{\ell \ell}$ are zeros save for the entry $(\ell, \ell)$ which is one.

We now prove the identity for the divergence of $G_{n}$. The divergence is

$$
\mathcal{D}:=\left.\sum_{\ell \leq \ell^{\prime}} \frac{\partial\left(G_{n}\right)_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R}=\left.\sum_{\ell \leq \ell^{\prime}} \frac{\partial\left(G_{n}\right)_{\ell^{\prime} \ell}}{\partial R_{\ell^{\prime} \ell}}\right|_{t, R}=\left.\frac{1}{2} \sum_{\ell \leq \ell^{\prime}} \frac{\partial\left(G_{n}\right)_{\ell \ell^{\prime}}}{\partial R_{\ell \ell^{\prime}}}\right|_{t, R}+\left.\frac{1}{2} \sum_{\ell \leq \ell^{\prime}} \frac{\partial\left(G_{n}\right)_{\ell^{\prime} \ell}}{\partial R_{\ell^{\prime} \ell}}\right|_{t, R}
$$

Replacing the summands by their formula (4.113) in the second equality yields

$$
\begin{align*}
\mathcal{D}= & \frac{n(p-1)}{2} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\left\langle Q_{\ell \ell^{\prime}}\right\rangle\right]^{p-2}\left(\left.\mathbb{E}\left[\left\langle\mathbf{Q} \circ\left(\mathbf{Q}+\mathbf{Q}^{\top}-\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle\right)\right\rangle\right]\right|_{\ell \ell^{\prime}}\right. \\
& \left.-\left.\mathbb{E}\left[\left\langle\mathbf{Q}^{\top}\right\rangle \circ\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle-2 \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right]\right|_{\ell \ell^{\prime}}\right) \\
= & \frac{n(p-1)}{2} \operatorname{Tr}\left(\mathbb{E}[\langle\mathbf{Q}\rangle]^{\circ(p-2)} \mathbb{E}\left[\left\langle\mathbf{Q}^{\top} \circ\left(\mathbf{Q}+\mathbf{Q}^{\top}-\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle\right)\right\rangle\right]\right) \\
- & \frac{n(p-1)}{2} \operatorname{Tr}\left(\mathbb{E}[\langle\mathbf{Q}\rangle]^{\circ(p-2)} \mathbb{E}\left[\langle\mathbf{Q}\rangle \circ\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle-2 \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right]\right) . \tag{4.119}
\end{align*}
$$

Remember that $\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]$, and therefore $\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]^{\circ(p-2)}$, are symmetric. Using that the trace is invariant by transposition and cyclic permutation, the two traces in (4.119) satisfy

$$
\begin{aligned}
& \operatorname{Tr}\left(\mathbb{E}[\langle\mathbf{Q}\rangle]^{\circ(p-2)} \mathbb{E}\left[\left\langle\mathbf{Q}^{\top} \circ\left(\mathbf{Q}+\mathbf{Q}^{\top}-\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle\right)\right\rangle\right]\right) \\
& \quad=\frac{1}{2} \operatorname{Tr}\left(\mathbb{E}[\langle\mathbf{Q}\rangle]^{\circ(p-2)} \mathbb{E}\left[\left\langle\left(\mathbf{Q}+\mathbf{Q}^{\top}\right) \circ\left(\mathbf{Q}+\mathbf{Q}^{\top}-\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle\right)\right\rangle\right]\right) \\
& \operatorname{Tr}\left(\mathbb{E}[\langle\mathbf{Q}\rangle]^{\circ(p-2)} \mathbb{E}\left[\langle\mathbf{Q}\rangle \circ\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle-2 \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right]\right) \\
& \quad=\frac{1}{2} \operatorname{Tr}\left(\mathbb{E}[\langle\mathbf{Q}\rangle]^{\circ(p-2)} \mathbb{E}\left[\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle \circ\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle-2 \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right]\right)
\end{aligned}
$$

Clearly,

$$
\mathbb{E}\left\langle\left(\mathbf{Q}+\mathbf{Q}^{\top}\right) \circ\left(\mathbf{Q}+\mathbf{Q}^{\top}-\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle\right)\right\rangle=\mathbb{E}\left\langle\left(\mathbf{Q}+\mathbf{Q}^{\top}-\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle\right)^{\circ 2}\right\rangle
$$

Similarly, we have

$$
\mathbb{E}\left[\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle \circ\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle-2 \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right]=\mathbb{E}\left[\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle-2 \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)^{\circ 2}\right]
$$

For this last equality, we can complete the square because

$$
\begin{aligned}
\mathbb{E}\left[\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n} \circ\right. & \left.\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle-2 \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)\right] \\
& =\mathbb{E}\left[\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n} \circ \frac{\langle\mathbf{x}\rangle^{\top} \mathbf{X}+\mathbf{X}^{\top}\langle\mathbf{x}\rangle}{n}\right]-2 \mathbb{E}\left[\left(\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)^{\circ 2}\right] \\
& =\mathbb{E}\left[\left\langle\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n} \circ \frac{\langle\mathbf{x}\rangle^{\top} \mathbf{x}+\mathbf{x}^{\top}\langle\mathbf{x}\rangle}{n}\right\rangle\right]-2 \mathbb{E}\left[\left(\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)^{\circ 2}\right] \\
& =\mathbb{E}\left[\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n} \circ \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle+\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right]-2 \mathbb{E}\left[\left(\frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)^{\circ 2}\right]=0
\end{aligned}
$$

where the second equality is due to the Nishimori identity. Plugging the two identities for the traces back in 4.119), and using the subsequent identities where we complete the square, we finally obtain

$$
\begin{aligned}
\mathcal{D}=\frac{n(p-1)}{4} & \operatorname{Tr}\left(\mathbb{E}\left[\langle\mathbf{Q}\rangle_{t, R}\right]^{\circ(p-2)}\right. \\
& \left.\cdot \mathbb{E}\left[\left\langle\left(\mathbf{Q}+\mathbf{Q}^{\top}-\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle\right)^{\circ 2}\right\rangle-\left(\left\langle\mathbf{Q}+\mathbf{Q}^{\top}\right\rangle-2 \frac{\langle\mathbf{x}\rangle^{\top}\langle\mathbf{x}\rangle}{n}\right)^{\circ 2}\right]\right)
\end{aligned}
$$

## 4.B Concentration of the free entropy

Once again, we consider the observations $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)$ defined in 4.20). The posterior distribution of $\mathbf{X}$ given $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)$ is

$$
d P\left(\mathbf{x} \mid \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right):=\frac{1}{\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)} e^{-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)} \prod_{j=1}^{n} d P_{X}\left(x_{j}\right)
$$

where

$$
\begin{aligned}
\mathcal{H}_{t, R}(\mathbf{x} ; \mathbf{y}, \widetilde{\mathbf{y}}):=\sum_{\underline{i} \in \mathcal{I}} & \frac{(1-t)(p-1)!}{2 n^{p-1}}\left(\sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k}\right)^{2}-\sqrt{\frac{(1-t)(p-1)!}{n^{p-1}}} y_{\underline{i}} \sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k} \\
& +\sum_{j=1}^{n} \frac{x_{j}^{\top} R x_{j}}{2}-\widetilde{y}_{j}^{\top} \sqrt{R} x_{j} .
\end{aligned}
$$

In this appendix we show that the free entropy

$$
\begin{equation*}
\frac{\ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right)}{n}=\frac{1}{n} \ln \left(\int \prod_{i=1}^{n} d P_{X}\left(x_{i}\right) e^{-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right)}\right) \tag{4.120}
\end{equation*}
$$

concentrates around its expectation. To shorten notations, we omit the arguments and write $\ln \mathcal{Z}_{t, R} / n$.

Theorem 4.14 (Concentration of the free entropy). Assume that $P_{X}$ has finite $(4 p-4)^{\text {th }}$ moments. There exists a positive constant $C$ depending only on $P_{X}, K$, $p$ and $\|R\|$ such that

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{t, R}}{n}-\mathbb{E}\left[\frac{\ln \mathcal{Z}_{t, R}}{n}\right]\right)^{2}\right] \leq \frac{C}{n} \tag{4.121}
\end{equation*}
$$

Proof. We omit the subscripts of the angular brackets $\langle-\rangle_{t, R}$. First, we show that the free entropy concentrates on its conditional expectation given the Gaussian noises $\mathbf{Z}, \widetilde{\mathbf{Z}}$. We see $\ln \mathcal{Z}_{t, R} / n$ as a function of $X_{1}, \ldots, X_{n}$ only and we work
conditionally to $\mathbf{Z}, \widetilde{\mathbf{Z}}$. Let $X_{1}^{\prime}, \ldots, X_{n}^{\prime}$ be i.i.d. samples from $P_{X}$, independent of $\mathbf{X}$. For all $j \in\{1, \ldots, n\}$, we define

$$
\mathcal{Z}_{t, R}^{(j)}:=\int \prod_{i=1}^{n} d P_{X}\left(x_{i}\right) e^{-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(j, t)}, \widetilde{\mathbf{Y}}^{(j, t, R)}\right)}
$$

where $\left(\mathbf{Y}^{(j, t)}, \widetilde{\mathbf{Y}}^{(j, t, R)}\right)$ has the same definition 4.20) than $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)$ except that $X_{j}$ is replaced $X_{j}^{\prime}$. We can consider an inference problem similar to 4.20 for which the observations are $\mathbf{Y}^{(j, t)}, \widetilde{\mathbf{Y}}^{(j, t, R)}$. Then, we denote by the angular brackets $\langle-\rangle_{(j)}$ the expectation with respect to the posterior distribution of $\mathbf{X}$ given $\left(\mathbf{Y}^{(j, t)}, \widetilde{\mathbf{Y}}^{(j, t, R)}\right)$, that is,

$$
\langle g(\mathbf{x})\rangle_{(j)}=\int g(\mathbf{x}) \frac{e^{-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(j, t)}, \tilde{\mathbf{Y}}^{(j, t, R)}\right)}}{\mathcal{Z}_{t, R}^{(j)}} \prod_{i=1}^{n} d P_{X}\left(x_{i}\right)
$$

By the Efron-Stein inequality (see Proposition 2.5),

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{t, R}}{n}-\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{Z}, \tilde{\mathbf{Z}}\right]\right)^{2}\right] \leq \frac{1}{2} \sum_{j=1}^{n} \mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{t, R}}{n}-\frac{\ln \mathcal{Z}_{t, R}^{(j)}}{n}\right)^{2}\right] \tag{4.122}
\end{equation*}
$$

Let $j \in\{1, \ldots, n\}$ be fixed. By Jensen's inequality, we have

$$
\begin{align*}
& \left\langle\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(j, t)}, \widetilde{\mathbf{Y}}^{(j, t, R)}\right)-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)\right\rangle_{(j)} \\
& \quad \leq \ln \mathcal{Z}_{t, R}-\ln \mathcal{Z}_{t, R}^{(j)} \leq\left\langle\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(j, t)}, \widetilde{\mathbf{Y}}^{(j, t, R)}\right)-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)\right\rangle . \tag{4.123}
\end{align*}
$$

Define $\mathcal{I}_{j}:=\left\{\underline{\boldsymbol{i}} \in \mathcal{I}: \exists b \in\{1, \ldots, p\}\right.$ s.t. $\left.i_{b}=j\right\}$ and $\forall \underline{\boldsymbol{i}} \in \mathcal{I}_{j}:$

$$
c(\underline{\boldsymbol{i}}):=\left|\left\{a \in\{1, \ldots, p\}: i_{a}=j\right\}\right| .
$$

The quantity inside the angular brackets $\langle-\rangle$ and $\langle-\rangle_{(j)}$ in 4.123) reads

$$
\begin{aligned}
& \mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(j, t)}, \widetilde{\mathbf{Y}}^{(j, t, R)}\right)-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right) \\
& \quad=\frac{(1-t)(p-1)!}{n^{p-1}} \sum_{\underline{i} \in \mathcal{I}_{j}} \sum_{\ell, \ell^{\prime}=1}^{K}\left(X_{j \ell}^{c(i)}-X_{j \ell}^{\prime c(i)}\right) \prod_{\substack{a=1 \\
i_{a} \neq j}}^{p} X_{i_{a} \ell} \prod_{a=1}^{p} x_{i_{a} \ell^{\prime}}+\left(X_{j}-X_{j}^{\prime}\right)^{\top} R x_{j} .
\end{aligned}
$$

Using $\left(\sum_{i=1}^{m} v_{i}\right)^{2} \leq m \sum_{i=1}^{m} v_{i}^{2}$ and Jensen's inequality, we thus have

$$
\begin{align*}
& \mathbb{E}\left\langle\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(j, t)}, \tilde{\mathbf{Y}}^{(j, t, R)}\right)-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right)\right\rangle^{2} \\
& \leq \frac{2((p-1)!)^{2} K^{2}\left|\mathcal{I}_{j}\right|}{n^{2 p-2}} \sum_{\underline{i} \in \mathcal{I}_{j}} \sum_{\ell, \ell^{\prime}=1}^{K} \mathbb{E}\left[\left(X_{j \ell}^{c(\underline{i})}-X_{j \ell}^{\prime c(\underline{i})}\right)^{2} \prod_{\substack{a=1 \\
i_{a} \neq j}}^{p} X_{i_{a} \ell}^{2}\left\langle\prod_{a=1}^{p} x_{i_{a} \ell^{\prime}}\right\rangle^{2}\right] \\
& \quad+2 \mathbb{E}\left[\left(\left(X_{j}-X_{j}^{\prime}\right)^{\top} R\left\langle x_{j}\right\rangle\right)^{2}\right] . \tag{4.124}
\end{align*}
$$

We bound each summand on the right-hand side of (4.124) separately. For all $\underline{\boldsymbol{i}} \in \mathcal{I}_{j}$ and $\left(\ell, \ell^{\prime}\right) \in\{1, \ldots, K\}^{2}$ :
$\mathbb{E}\left[\left(X_{j \ell}^{c(i)}-X_{j \ell}^{\prime c(i)}\right)^{2} \prod_{\substack{a=1 \\ i_{a} \neq j}}^{p} X_{i_{a} \ell}^{2}\left\langle\prod_{a=1}^{p} x_{i_{a} \ell^{\prime}}\right\rangle^{2}\right]$
$\leq \mathbb{E}\left[\left(X_{j \ell}^{\tau(i)}-X_{j \ell}^{\prime c(i)}\right)^{4} \prod_{\substack{a=1 \\ i_{a} \neq j}}^{p} X_{i_{a} \ell}^{4}\right]^{1 / 2} \mathbb{E}\left[\left\langle\prod_{a=1}^{p} x_{i_{a} \ell^{\prime}}\right\rangle^{4}\right]^{1 / 2}$
$\leq \mathbb{E}\left[\left(X_{j \ell}^{c(i)}-X_{j \ell}^{\prime c(i)}\right)^{4} \prod_{\substack{a=1 \\ i_{a} \neq j}}^{p} X_{i_{a} \ell}^{4}\right]^{1 / 2} \mathbb{E}\left[\left\langle\prod_{a=1}^{p} x_{i_{a} \ell^{\prime}}^{4}\right\rangle\right]^{1 / 2}$
$=\mathbb{E}\left[\left(X_{j \ell}^{c(i)}-X_{j \ell}^{\prime c(i)}\right)^{4} \prod_{\substack{a=1 \\ i_{a} \neq j}}^{p} X_{i_{a} \ell}^{4}\right]^{1 / 2} \mathbb{E}\left[\prod_{a=1}^{p} X_{i_{a} \ell^{\prime}}^{4}\right]^{1 / 2}$
$=\sqrt{\mathbb{E}\left[\left(X_{j \ell}^{c(i)}-X_{j \ell}^{\prime((i)}\right)^{4}\right] \mathbb{E}\left[X_{j \ell^{\prime}}^{4 c(i)}\right] \mathbb{E}\left[\prod_{\substack{a=1 \\ i_{a} \neq j}}^{p} X_{i_{a} \ell}^{4}\right] \mathbb{E}\left[\prod_{\substack{a=1 \\ i_{a} \neq j}}^{p} X_{i_{a} \ell^{\prime}}^{4}\right]}$.
The first inequality follows from the Cauchy-Schwarz inequality, the second one from Jensen's inequality, and the first equality from the Nishimori identity. The final upper bound that we obtain is finite given that $P_{X}$ has finite $(4 p-4)^{\text {th }}$ moments. Hence, there exists a positive constant $C$ depending only on $P_{X}, K$ and $p$ such that the first term on the right-hand side of (4.124) is bounded by

$$
\frac{C\left|\mathcal{I}_{j}\right|^{2}}{n^{2 p-2}} \leq C
$$

where we use that $\left|\mathcal{I}_{j}\right| \leq n^{p-1}$. Regarding the second term on the right-hand side of (4.124), we easily get

$$
\begin{aligned}
\mathbb{E}\left[\left(\left(X_{j}^{\prime}-X_{j}\right)^{\top} R\left\langle x_{j}\right\rangle\right)^{2}\right] & \leq \mathbb{E}\left[\left\|X_{j}^{\prime}-X_{j}\right\|^{2}\|R\|^{2}\left\|\left\langle x_{j}\right\rangle\right\|^{2}\right] \\
& \leq\|R\|^{2} \sqrt{\mathbb{E}\left[\left\|X_{j}^{\prime}-X_{j}\right\|^{4}\right] \mathbb{E}\left[\left\|X_{j}\right\|^{4}\right]}
\end{aligned}
$$

We conclude that there exists a positive constant $C$ depending only on $P_{X}, K, p$ and $\|R\|$ such that $\forall j \in\{1, \ldots, n\}$ :

$$
\begin{equation*}
\mathbb{E}\left[\left\langle\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(j, t)}, \widetilde{\mathbf{Y}}^{(j, t, R)}\right)-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)\right\rangle^{2}\right] \leq C . \tag{4.125}
\end{equation*}
$$

A similar bound holds when the angular brackets $\langle-\rangle$ are replaced by $\langle-\rangle_{(j)}$,

$$
\begin{equation*}
\mathbb{E}\left[\left\langle\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(j, t)}, \tilde{\mathbf{Y}}^{(j, t, R)}\right)-\mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}\right)\right\rangle_{(j)}^{2}\right] \leq C \tag{4.126}
\end{equation*}
$$

Combining (4.123), 4.125) and 4.126) yields

$$
\mathbb{E}\left[\left(\ln \mathcal{Z}_{t, R}-\ln \mathcal{Z}_{t, R}^{(j)}\right)^{2}\right] \leq C
$$

We use this last upper bound in the Efron-Stein inequality (4.122) to finally obtain

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{t, R}}{n}-\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{Z}, \widetilde{\mathbf{Z}}\right]\right)^{2}\right] \leq \frac{C}{2 n} \tag{4.127}
\end{equation*}
$$

where the positive constant $C$ only depends on $P_{X}, K, p$ and $\|R\|$.
The second and final step is to show that the conditional expectation of the free entropy given $\mathbf{Z}, \widetilde{\mathbf{Z}}$ concentrates on its expectation. Let $g(\mathbf{Z}, \widetilde{\mathbf{Z}}):=\mathbb{E}\left[\ln \mathcal{Z}_{t, R} / n \mid \mathbf{Z}, \widetilde{\mathbf{Z}}\right]$. By the Gaussian-Poincaré inequality (see Proposition 2.7),

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{Z}, \widetilde{\mathbf{Z}}\right]-\mathbb{E}\left[\frac{\ln \mathcal{Z}_{t, R}}{n}\right]\right)^{2}\right] \leq \mathbb{E}\|\nabla g(\mathbf{Z}, \widetilde{\mathbf{Z}})\|^{2} \tag{4.128}
\end{equation*}
$$

The squared norm of the gradient of $g$ reads

$$
\|\nabla g\|^{2}=\sum_{\underline{i} \in \mathcal{I}}\left|\frac{\partial g}{\partial Z_{\underline{i}}}\right|^{2}+\sum_{j=1}^{n} \sum_{\ell=1}^{K}\left|\frac{\partial g}{\partial \widetilde{Z}_{j \ell}}\right|^{2}
$$

Each of these partial derivatives takes the form $\frac{\partial g}{\partial x}=-\frac{1}{n} \mathbb{E}\left[\left.\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial x}\right\rangle \right\rvert\, \mathbf{Z}, \widetilde{\mathbf{Z}}\right]$, and

$$
\left|\frac{\partial \mathcal{H}_{t, R}}{\partial Z_{\underline{i}}}\right|=\left|\sqrt{\frac{(1-t)(p-1)!}{n^{p-1}}} \sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k}\right|,\left|\frac{\partial \mathcal{H}_{t, R}}{\partial \widetilde{Z}_{j \ell}}\right|=\left|\left(\sqrt{R} x_{j}\right)_{\ell}\right| .
$$

On one hand, by Jensen's inequality, we have

$$
\begin{align*}
\sum_{\underline{i} \in \mathcal{I}} \mathbb{E}\left|\frac{\partial g}{\partial Z_{\underline{i}}}\right|^{2} & \leq \frac{(p-1)!}{n^{p+1}} \sum_{\underline{i} \in \mathcal{I}} \mathbb{E}\left[\left\langle\left(\sum_{k=1}^{K} \prod_{a=1}^{p} x_{i_{a} k}\right)^{2}\right\rangle\right] \\
& \leq \frac{K(p-1)!}{n^{p+1}} \sum_{\underline{i} \in \mathcal{I}} \sum_{k=1}^{K} \mathbb{E}\left[\left\langle\prod_{a=1}^{p} x_{i_{a} k}^{2}\right\rangle\right] \\
& =\frac{K(p-1)!}{n^{p+1}} \sum_{\underline{i} \in \mathcal{I}} \sum_{k=1}^{K} \mathbb{E}\left[\prod_{a=1}^{p} X_{i_{a} k}^{2}\right] \tag{4.129}
\end{align*}
$$

where the final equality follows from the Nishimori identity. On the other hand, still by Jensen's inequality, we have

$$
\begin{align*}
\sum_{j=1}^{n} \sum_{\ell=1}^{K} \mathbb{E}\left|\frac{\partial g}{\partial \widetilde{Z}_{j \ell}}\right|^{2} & \leq \frac{1}{n^{2}} \sum_{j=1}^{n} \sum_{\ell=1}^{K} \mathbb{E}\left[\left\langle\left(\sqrt{R} x_{j}\right)_{\ell}^{2}\right\rangle\right] \\
& =\frac{1}{n^{2}} \sum_{j=1}^{n} \mathbb{E}\left\|\sqrt{R} X_{j}\right\|^{2}=\frac{\mathbb{E}\left\|\sqrt{R} X_{1}\right\|^{2}}{n} \leq \frac{K\|R\|}{n} \mathbb{E}\left\|X_{1}\right\|^{2} \tag{4.130}
\end{align*}
$$

where the first equality follows from the Nishimori identity and the last inequality from the submultiplicativity of the Frobenius norm and $\|\sqrt{R}\| \leq \sqrt{K\|R\|}$ (see
(4.108). Remember that $|\mathcal{I}| \leq n^{p}$ and that $P_{X}$ has bounded $(4 p-4)^{\text {th }}$ moments. Hence, both 4.129) and (4.130) are $O\left(n^{-1}\right)$. Using 4.129) and 4.130) in the Gaussian-Poincaré inequality (4.128) yields

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{Z}, \widetilde{\mathbf{Z}}\right]-\mathbb{E}\left[\frac{\ln \mathcal{Z}_{t, R}}{n}\right]\right)^{2}\right] \leq \frac{C}{n} \tag{4.131}
\end{equation*}
$$

where $C$ depends only on $P_{X}, K, p$ and $\|R\|$. Combining (4.127) and (4.131) ends the proof of 4.121.

## Tensor estimation with structured priors

### 5.1 Introduction

Natural signals have an underlying structure, an insight that has triggered a paradigm shift in the last fifteen years, and spurred fundamental progress in estimation and inference. Compressive sensing [30, 31 takes sparsity as the model of structure when a signal $\mathbf{X} \in \mathbb{R}^{n}$ has a sparse representation in an appropriate basis, that is, $\mathbf{X}=\Psi \mathbf{Z}$ with $\Psi$ an $n \times n$ change of basis matrix and $\mathbf{Z} \in \mathbb{R}^{n}$ a sparse vector with $p \ll n$ non-zero components. For example, $\mathbf{X}$ can represent a natural image and $\Psi$ a wavelet basis [112]. Despite its success, this model of structure is often too constrained because the appropriate basis may be unknown and, more generally, the linearity of the transformation may be a severe limitation. Deep networks have been proposed as an alternative 113 and, with the advent of generative adversarial networks (GAN) [114 and variational auto-encoders (VAE) [115], such flexible and non-linear "generative models" of structure have been the object of intense interest. Roughly speaking, a generative model can be viewed as a mapping $G: \mathbf{S} \in \mathbb{R}^{p} \mapsto \mathbf{X}=G(\mathbf{S}) \in \mathbb{R}^{n}$ with $p \ll n$ and satisfying certain general regularity assumptions. In other words, the signal X lies on a low $p$-dimensional "manifold" parametrized by S. Such models have been studied in the framework of classical denoising problems from observations $\mathbf{Y}=A \mathbf{X}+\mathbf{Z}$ where $A$ is a sensing matrix and $\mathbf{Z}$ some Gaussian noise [116]. In particular, 116 studies fundamental limits (under minimal Lipshitz conditions on $G$ ) and empirically investigates the problem with learned mappings coming from GAN and VAE. A related line of research uses untrained deep networks (so-called deep decoders or deep image priors) whose parameters are adjusted by an optimization problem over the latent space [117]-[119]. Another kind of untrained generative model takes $G$ equal to a one-layer or multi-layer neural network with fixed weights (i.e., frozen and not learned) drawn from a random matrix ensemble [27], [120]-[123]. Such mappings $G$ are often referred to as generalized linear models (GLMs) and this is the terminology that we adopt here. The simplification of fixed random weights has the virtue of being much
more amenable to mathematical (or at least analytical) analysis. Especially, the mutual information as well as the message passing algorithmic behaviour for classical denoising have been discussed in depth in a Bayesian setting at various levels of rigor [27], [75]. Below we use a one-layer version of GLMs. These originated as generalizations of linear regression models 124 and have many modern applications in communications (e.g., CDMA, sparse regression codes on general channels), signal processing (e.g., one-bit compressive sensing, phase retrieval) as well as machine learning and statistics (e.g., classication tasks). We refer to [29] for a review of this literature and references.

In this chapter we investigate GLMs of structure in the context of estimation of noisy tensors. Tensors representing data have found many modern applications in signal processing, graph analysis, data mining and machine learning [6], [13], [94], with a large part of the literature focusing on tensor decompositions, either in deterministic settings, or in random settings with independent structureless components. Here, we focus on a simple statistical model of noisy symmetric rank-one tensors. A structured signal $\mathbf{X}=\left(X_{1}, \cdots, X_{n}\right) \in \mathbb{R}^{n}$ is generated by a one-layer GLM $X_{i}=\varphi\left((\mathbf{W S})_{i} / \sqrt{p}\right)$. The latent vector $\mathbf{S} \in \mathbb{R}^{p}$ has independent and identically distributed (i.i.d.) entries and $\mathbf{W}$ is a known random matrix with independent standard Gaussian entries. We only observe a noisy version of the rank-one tensor $\mathbf{X}^{\otimes r}(r \geq 2)$ through an additive white Gaussian noise channel, i.e., we observe $\mathbf{Y}:=\frac{\sqrt{\lambda}}{n^{(r-1) / 2}} \mathbf{X}^{\otimes r}+\mathbf{Z}$, where the noise $\mathbf{Z}$ is a symmetric tensor with independent standard Gaussian entries and $\lambda>0$ is the signal-tonoise ratio. We study the high dimensional limit $n, p \rightarrow \infty$ such that $n / p \rightarrow$ $\alpha=\Theta(1)$ and show that, quite remarkably, the asymptotic normalized mutual information $\lim _{n \rightarrow+\infty} I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W}) / n$ is given by a finite-dimensional variational problem (see Theorem 5.1 in Section 5.2.2). We also rigorously deduce the corresponding asymptotic minimum mean-square error (MMSE), which is given by a simple function of the solution to the variational problem (see Theorem 5.2 in Section 5.2.2. For concreteness, and to keep the analysis as simple as possible, we focus on the case $r=3$ and one-layer GLM. However, extensions to any order $r>3$, multi-layer GLM and asymmetric tensors are possible with the techniques used here. An extensive recent study of the matrix case $r=2$ can be found in 98.

The analysis and results presented here go beyond many recent works dealing with i.i.d. components for $\mathbf{X}$, for matrices $r=2$ [34], [36], [46], and tensors $r \geq 3$ [45], [69]. There is a rich phenomenology of phase transitions already for the i.i.d. case which stems from the (simpler) variational formula for the mutual information. In Section 5.3 we discuss the (numerical) solutions to the new variational problem obtained for structured signals for various examples of priors and activation functions, and we illustrate properties of phase transitions. Furthermore, we discuss the similarities and differences between the genuine tensor and matrix cases.

Let us say a few words about the techniques used in this chapter. For structured signals, rigorous proofs of the low-dimensional variational expression for the asymptotic normalized mutual information are virtually non-existent. To

### 5.2. Asymptotic mutual information and MMSE for tensor estimation with a generative prior

the best of our knowledge, besides the treatements in Chapter 3 where $\mathbf{X}$ is uniformly distributed on the sphere (which turns out to be equivalent to an i.i.d. Gaussian prior) and in Chapter 6 where the input to a one-layer GLM is generated by another GLM, there is one recent exception: 98 treats the rank-one matrix case with input coming from a GLM. The latter work uses two different flavors of the interpolation method [35], [125] which do not extend to odd-order tensors nor asymmetric ones. Moreover, certain (reasonable) assumptions are required. In this chapter we rely entirely on the adaptive interpolation method [37], [87]. Our treatment is completely self-contained, leverages on only one method, and can also deal with asymmetric matrices and tensors. We would like to emphasize that the modularity of the adaptive interpolation method plays an important role in this work. It will become clear in Section 5.4 how the normalized mutual information associated with GLMs, itself computed by an interpolation in the high-dimensional limit, appears as a building block of the interpolation for the tensor model. This modular aspect was first emphasized in 69 and is also used in the subsequent Chapter 6. We have here one more example where this modularity finds an application.

In Section 5.2 we formulate the statistical model and present the main theorems on the asymptotic normalized mutual information and MMSE. In Section 5.3 we use our theoretical results to illustrate phase transitions on different examples. In Sections 5.4 and 5.5 we go through the proofs and, in Section 5.6, we give an analysis of the limit $\alpha \rightarrow 0$. The appendices contain technical derivations.

### 5.2 Asymptotic mutual information and MMSE for tensor estimation with a generative prior

We formulate a statistical model for rank-one tensor estimation when the spike is itself generated from another latent vector. We observe a noisy symmetric tensor $\mathbf{Y} \in\left(\mathbb{R}^{n}\right)^{\otimes 3}$ whose entries are

$$
\begin{equation*}
Y_{i j k}:=\frac{\sqrt{\lambda}}{n} X_{i} X_{j} X_{k}+Z_{i j k}, \quad 1 \leq i \leq j \leq k \leq n \tag{5.1}
\end{equation*}
$$

where the positive real number $\lambda$ plays the role of an $\operatorname{SNR}, Z_{i j k} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ for $1 \leq i \leq j \leq k \leq n$ is an additive white Gaussian noise, and $X_{1}, \ldots, X_{n}$ are the entries of the spike $\mathbf{X} \in \mathbb{R}^{n}$. Let $P_{S}$ be a probability distribution on the real numbers. The spike $\mathbf{X}$ is generated by a latent vector $\mathbf{S} \in \mathbb{R}^{p}$, whose entries are i.i.d. with respect to (w.r.t.) $P_{S}$, via the GLM

$$
\begin{equation*}
X_{i}:=\varphi\left(\frac{(\mathbf{W S})_{i}}{\sqrt{p}}\right), i=1, \ldots, n \tag{5.2}
\end{equation*}
$$

The $n \times p$ random matrix $\mathbf{W}$ has entries i.i.d. with respect to $\mathcal{N}(0,1)$. It is often customary to summarize (5.2) by $\mathbf{X}=\varphi(\mathbf{W S} / \sqrt{p})$ where it is understood that the function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ is applied componentwise. A similar statistical model for tensor PCA where the entries of the spike $\mathbf{X}$ are i.i.d was introduced in [19].

### 5.2.1 Scalar Gaussian channels

Before presenting our main results, let us introduce two kinds of scalar Gaussian channels and the mutual informations that they are associated with. These mutual informations are building blocks of the replica symmetric formula that gives the normalized conditional mutual information between $\mathbf{X}$ and $\mathbf{Y}$ given $\mathbf{W}$ in the high-dimensional limit.

The first channel is a scalar linear Gaussian channel. Let $S \sim P_{S}$ and $Z \sim \mathcal{N}(0,1)$ be two independent random variables. Let $r$ be a nonnegative real number akin to a signal-to-noise ratio. Consider the problem of estimating $S$ from the noisy channel observation $\sqrt{r} S+Z$. We denote $I_{P_{S}}(r)$ the mutual information between the input and output of this channel, that is,

$$
I_{P_{S}}(r):=I(S ; \sqrt{r} S+Z)
$$

We list important properties of the function $I_{P_{S}}: r \in[0,+\infty) \mapsto I(S ; \sqrt{r} S+Z)$ in Lemma 2.3 of Chapter 2.

The second channel is still a scalar Gaussian channel but it is in general nonlinear as it involves the function $\varphi$ used to generate our spike $\mathbf{X}$. Let $U, V$ and $Z$ be independent standard Gaussian random variables, that is, $U, V, Z \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. Let $r, \rho$ be nonnegative real numbers and $q \in[0, \rho]$. Consider the problem of estimating $U$ from the noisy observation $\sqrt{r} \varphi(\sqrt{\rho-q} U+\sqrt{q} V)+Z$ while $V$ is perfectly known. We denote $I_{\varphi}(r, q ; \rho)$ the conditional mutual information between $U$ and the output of this channel given $V$, that is,

$$
I_{\varphi}(r, q ; \rho):=I(U ; \sqrt{r} \varphi(\sqrt{\rho-q} U+\sqrt{q} V)+Z \mid V)
$$

Lemma 5.9 in Appendix $5 . \mathrm{A}$ lists important properties of the function

$$
I_{\varphi}(\cdot ; \rho):(r, q) \in[0,+\infty) \times[0, \rho] \mapsto I(U ; \sqrt{r} \varphi(\sqrt{\rho-q} U+\sqrt{q} V)+Z \mid V)
$$

### 5.2.2 Main results

The next two theorems provide a complete information-theoretic characterization of the problem. Theorem 5.1 expresses the normalized mutual information $I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W}) / n$ as an explicit low-dimensional variational problem, in the high-dimensional regime where $n \rightarrow+\infty$ and $n / p=\alpha$ is kept fixed. This variational problem involves an optimization over three parameters and can be solved numerically given the activation function $\varphi$ and prior distribution $P_{S}$.

Theorem 5.1 (Normalized mutual information in the high-dimensional regime). Suppose that the following hypotheses hold:
(H1) There exists $M_{S}>0$ such that the probability distribution $P_{S}$ is supported on $\left[-M_{S}, M_{S}\right]$.
(H2) The function $\varphi$ is bounded and twice differentiable with its first and second derivatives being bounded and continuous. They are denoted $\varphi^{\prime}, \varphi^{\prime \prime}$.

### 5.2. Asymptotic mutual information and MMSE for tensor estimation with a generative prior

Define the second moments $\rho_{s}:=\mathbb{E}\left[S^{2}\right]$ where $S \sim P_{S}$, and $\rho_{x}:=\mathbb{E}\left[\varphi(T)^{2}\right]$ where $T \sim \mathcal{N}\left(0, \rho_{s}\right)$. Define the potential function

$$
\begin{align*}
\psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right):=\frac{I_{P_{S}}\left(r_{s}\right)}{\alpha}+I_{\varphi}\left(\frac{\lambda q_{x}^{2}}{2}, q_{s} ; \rho_{s}\right) & -\frac{r_{s}\left(\rho_{s}-q_{s}\right)}{2 \alpha} \\
& +\frac{\lambda}{12}\left(\rho_{x}-q_{x}\right)^{2}\left(\rho_{x}+2 q_{x}\right) \tag{5.3}
\end{align*}
$$

where $\left(q_{x}, q_{s}, r_{s}\right) \in\left[0, \rho_{x}\right] \times\left[0, \rho_{s}\right] \times[0,+\infty)$. If $n, p$ go to infinity such that $n / p \rightarrow \alpha>0$ then

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n}=\inf _{q_{x} \in\left[0, \rho_{x}\right]} \inf _{\left.q_{s} \in\left[0, \rho_{s}\right]\right]_{s} \geq 0} \sup _{r_{s}} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right), \tag{5.4}
\end{equation*}
$$

where $\mathbf{W} \in \mathbb{R}^{n \times p}$ has independent standard Gaussian entries, $\mathbf{X}$ is defined in (5.2) and $\mathbf{Y}$ in 5.1.

One important quantity to assess the performance of an algorithm designed to recover $\mathbf{X}^{\otimes 3}$ from the knowledge of $\mathbf{Y}$ and $\mathbf{W}$ is the minimum mean-square error (MMSE). The latter serves as a lower bar on the error of any estimator, and as a limit to approach as closely as possible for any algorithm striving to estimate $\mathbf{X}^{\otimes 3}$. It is well-known that the mean square error of an estimator of $\mathbf{X}^{\otimes 3}$ that is a function of $\mathbf{Y}, \mathbf{W}$ only is minimized by the posterior mean $\mathbb{E}\left[\mathbf{X}^{\otimes 3} \mid \mathbf{Y}, \mathbf{W}\right]$. We denote the tensor-MMSE by $\mathrm{MMSE}_{n}\left(\mathbf{X}^{\otimes 3} \mid \mathbf{Y}, \mathbf{W}\right)$, i.e.,

$$
\begin{equation*}
\operatorname{MMSE}_{n}\left(\mathbf{X}^{\otimes 3} \mid \mathbf{Y}, \mathbf{W}\right):=\frac{\mathbb{E}\left\|\mathbf{X}^{\otimes 3}-\mathbb{E}\left[\mathbf{X}^{\otimes 3} \mid \mathbf{Y}, \mathbf{W}\right]\right\|^{2}}{n^{3}} \tag{5.5}
\end{equation*}
$$

It depends on $\lambda$ through the observations Y. Combining Theorem 5.1 with the I-MMSE relationship (see [51])

$$
\begin{equation*}
\frac{\partial}{\partial \lambda}\left(\frac{I(\mathbf{X}, \mathbf{Y} \mid \mathbf{W})}{n}\right)=\frac{1}{12} \mathrm{MMSE}_{n}\left(\mathbf{X}^{\otimes 3} \mid \mathbf{Y}, \mathbf{W}\right)+O\left(n^{-1}\right) \tag{5.6}
\end{equation*}
$$

yields Theorem 5.2. It gives a formula for the tensor-MMSE in the high-dimensional regime that can be calculated from the solution to the variational problem (5.4).
Theorem 5.2 (Tensor-MMSE). Suppose that (H1) and (H2) hold. Define for all $\lambda \in(0,+\infty)$ :

$$
\begin{aligned}
& \mathcal{Q}_{x}^{*}(\lambda) \\
& :=\left\{q_{x}^{*} \in\left[0, \rho_{x}\right]: \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q_{x}^{*}, q_{s}, r_{s}\right)=\inf _{q_{x} \in\left[0, \rho_{x}\right]} \inf _{\left.q_{s} \in\left[0, \rho_{s}\right]\right]_{s} \geq 0} \sup _{r_{s}} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right)\right\} .
\end{aligned}
$$

For every $\lambda>0, \mathcal{Q}_{x}^{*}(\lambda)$ is nonempty and the set

$$
\mathcal{D}:=\left\{\lambda \in(0,+\infty): \mathcal{Q}_{x}^{*}(\lambda) \text { is a singleton }\right\}
$$

is equal to $(0,+\infty)$ minus a countable set. Besides, for every $\lambda \in \mathcal{D}$, we have

$$
\begin{equation*}
\lim _{\substack{n \rightarrow+\infty \\ n / p \rightarrow \alpha}} \operatorname{MMSE}_{n}\left(\mathbf{X}^{\otimes 3} \mid \mathbf{Y}, \mathbf{W}\right)=\rho_{x}^{3}-\left(q_{x}^{*}(\lambda)\right)^{3} \tag{5.7}
\end{equation*}
$$

where $q_{x}^{*}(\lambda)$ is the unique element in $\mathcal{Q}_{x}^{*}(\lambda)$, i.e., $\mathcal{Q}_{x}^{*}(\lambda)=\left\{q_{x}^{*}(\lambda)\right\}$.

We prove Theorem 5.1 in Section 5.4. The proof is based on the adaptive interpolation method [37, [87] whose main difference with the canonical interpolation method [60], [61] is the increased flexibility given to the path followed by the interpolation between its two extremes. The method has been developed separately for symmetric rank-one tensor problems where the spike has i.i.d. components [37], [87], and for one-layer GLMs whose input signal has again i.i.d. components [29]. The problem studied in this contribution combines the two aforementioned models and our proof shows that the two interpolations combine well in a modular way. This modular feature of the adaptive interpolation method has also been used for non-symmetric order-three tensors [69] and two-layer GLMs [75].

The proof of Theorem 5.2 is given in Section 5.5. We rely on the I-MMSE relationship (5.6) and compute the derivative with respect to $\lambda$ of the variational formula (5.4) for the asymptotic mutual information. The computation requires a careful application of an envelope theorem [68, Corollary 4] that eventually shows that, except for a countable set of $\lambda$ 's, it is enough to evaluate the partial derivative with respect to $\lambda$ of the potential (5.3) at the solution of the variational problem.

### 5.2.3 Extensions

Extensions of Theorems 5.1 and 5.2 in various directions are possible with the methods of the present paper, but at the expense of more technical work. First, the analysis for rank-one tensors of any rank $r \geq 3$ is identical. The potential is given by

$$
\begin{aligned}
\psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right):=\frac{I_{P_{S}}\left(r_{s}\right)}{\alpha}+I_{\varphi}\left(\frac{\lambda q_{x}^{r-1}}{(r-1)!}, q_{s} ; \rho_{s}\right) & -\frac{r_{s}\left(\rho_{s}-q_{s}\right)}{2 \alpha} \\
& +\frac{\lambda}{2(r!)}\left(\rho_{x}^{r}+r q_{x}^{r}-r q_{x}^{r-1} \rho_{x}\right)
\end{aligned}
$$

while the asymptotic tensor-MMSE is $\rho_{x}^{r}-\left(q_{x}^{*}(\lambda)\right)^{r}$.
Second, the results can be extended to more general activation function and prior via a limiting process on both sides of (5.4) similar to what is done in [29, Appendices C. 1 and C. 2 of SI]. Assumptions (H1) and (H2) are relaxed to:
(h1) The probability distribution $P_{S}$ has a finite third moment and at least two points in its support.
(h2) The function $\varphi$ is continuous almost everywhere and there exists $\epsilon>0$ such that the sequence $\left(\mathbb{E}\left|\varphi\left((\mathbf{W S})_{1} / \sqrt{\bar{p}}\right)\right|^{2+\epsilon}\right)_{p \geq 1}$ is bounded.
Therefore, the results apply to $\varphi$ being the identity function, the sign function, or the $\operatorname{ReLU} x \mapsto \max (0, x)$. Another direction that should be amenable to analysis with our methods is the case of asymmetric tensors, e.g., $\mathbf{X}^{\otimes 3}$ is replaced by $\mathbf{U} \otimes \mathbf{V} \otimes \mathbf{W}$ where each of the three different vectors is given by a GLM. The structureless case where all three vectors $\mathbf{U}, \mathbf{V}, \mathbf{W}$ have i.i.d. entries is treated in [69], and the variational problem already displays a rich phenomenology in the highly asymmetric case [84].

### 5.3 Examples of phase transitions and their properties

This section illustrates features of the phase transitions found when numerically solving the variational problem (5.4) for $r=3$. We also discuss similarities and differences with the matrix case $r=2$. To find solutions to the variational problem (5.4), we write down the stationary point equations of the potential function (5.3). It yields a fixed point equation for $\left(q_{x}, q_{s}, r_{s}\right)$ that we solve with a fixed-point iteration starting from several different initializations. When multiple fixed points exist, we keep the one corresponding to the smallest potential value as it should be clear from the form of the optimization problem (5.4).

We first focus on the case of odd activation functions $\varphi(-z)=-\varphi(z)$ and centered priors $\mathbb{E}_{S \sim P_{S}}[S]=0$. This implies $\mathbb{E} X_{i}=0$ and, if $\varphi$ is not identically zero, this is a necessary and sufficient condition for the existence of a fixed point $\left(q_{x}, q_{s}, r_{s}\right)$ such that $q_{x}=0$ (in which case we also have $q_{s}=r_{s}=0$ ). The same condition arises in the matrix case [98] but, contrary to what happens there, we find that all eigenvalues of the Jacobian matrix at the all-zero fixed point are zero indicating that it is asymptotically stable for order-3 tensors. Numerically, we observe that there exists a critical value of $\lambda$, denoted $\lambda_{c}(\alpha)$, below which the uninformative fixed point $\left(q_{x}, q_{s}, r_{s}\right)=0$ yields the smallest potential. It means that the asymptotic tensor-MMSE is equal to its maximum $\rho_{x}^{3}$ for $\lambda<\lambda_{c}(\alpha)$; one cannot estimate the signal better than random guessing. When $\lambda>\lambda_{c}(\alpha)$, a fixed point with a lower potential value appears. The asymptotic MMSE has a jump discontinuity at $\lambda=\lambda_{c}(\alpha)$ and decreases for $\lambda>\lambda_{c}(\alpha)$. These features are already observed for the structureless i.i.d. case. In the structured case, we observe that $\lambda_{c}(\alpha)$ has a monotone decrease with increasing $\alpha$. This is illustrated in Figure 1 for a linear activation function and in Figure 2 for a sign activation function

In Section 5.6 we present a non-rigorous calculation which shows that, in the limit $\alpha \rightarrow 0(p \gg n)$, the asymptotic tensor-MMSE - and in particular the threshold $\lambda_{c}(\alpha)$ - is the same than for the tensor denoising problem

$$
\widetilde{Y}_{i j k}:=\frac{\sqrt{\lambda}}{n} \widetilde{X}_{i} \widetilde{X}_{j} \widetilde{X}_{k}+\widetilde{Z}_{i j k}, \quad 1 \leq i \leq j \leq k \leq n,
$$

with $\widetilde{X}_{i}:=\varphi\left(\sqrt{\rho_{s}-\mathbb{E}[S]^{2}} U_{i}+|\mathbb{E} S| V_{i}\right)$ where $U_{1}, \ldots, U_{n} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ are latent variables and $V_{1}, \ldots, V_{n} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ are known. The latter take into account the bias that is present when $\mathbb{E} S \neq 0$. We stress that when $\mathbb{E} S \neq 0$ the asymptotic mutual information of this problem (given by (5.41) in Section 5.6) is not quite the same as the one known in the literature for rank-one tensor problems with i.i.d. $X_{i}$ 's. However, it is not difficult to adapt the proof to account for the side information $\mathbf{V}$ and obtain (5.41). When the prior is centered $(\mathbb{E} S=0)$, the limiting problem is just the usual rank-one tensor denoising problem with spike

[^16]

Figure 1: Asymptotic tensor-MMSE for $r=3$ as a function of $(\lambda, \alpha)$ for a linear activation $\varphi(x)=x$. Left: Gaussian prior $P_{S} \sim \mathcal{N}(0,1)$. Right: Rademacher prior $P_{S}(1)=P_{S}(-1)=\frac{1}{2}$. We observe a unique discontinuity line $\lambda_{c}(\alpha)$ below which the MMSE equals its maximum $\rho_{x}^{3}=1$. Above the line, the MMSE is strictly less than 1 and decreases to zero. For $\alpha$ close to 0 , the threshold $\lambda_{c}(\alpha) \approx 8.73$ is the same threshold than in the i.i.d. case with a Gaussian prior $X_{1}, \ldots, X_{n} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$.
signal $\widetilde{X}_{i} \stackrel{\text { i.i.d. }}{\sim} \varphi\left(\mathcal{N}\left(0, \rho_{s}\right)\right)$. Numerically, we indeed observe in Figure 1 that for both kinds of priors and for $\alpha$ close to 0 the threshold $\lambda_{c}(\alpha) \approx 8.73$ is the same than for a signal $X_{1}, \ldots, X_{n} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. Similarly, in Figure 2 , the curve for $\alpha=10^{-12}$ agrees with the one labelled " $\alpha \rightarrow 0^{+}$" corresponding to the asymptotic tensor-MMSE of the limiting tensor problem and that is computed using the formulas known in the literature.

In the opposite limit $\alpha \rightarrow+\infty(p \ll n)$, corresponding to a very strongly structured prior, the asymptotic MMSE displays on Figure 1 a vanishing phase transition threshold, $\lambda_{c}(\alpha) \rightarrow 0$. This is expected since the dimension of the latent vector is much smaller than that of the feature vector; the problem becomes information-theoretically easier. It is also consistent with a (non-rigorous) inspection of the potential (5.3), whereby the variational problem simplifies and has a solution with maximal overlap $q_{x}^{*}=\rho_{x}$.

We next discuss an example of non-centered latent prior $P_{S}$. In Figure 3 we draw the asymptotic tensor-MMSE for a linear activation function and a Rademacher prior $P_{S}(1)=p, P_{S}(-1)=1-p$ with $p \in\{0.6,0.7\}$. We observe that for a small asymmetry the asymptotic MMSE has a jump discontinuity just as in the centered case, while it becomes continuous once the asymmetry is large enough. Here $\mathbb{E} S=2 p-1$ and the asymptotic MMSE of the predicted limiting problem (5.41) is again in agreement with the one for $\alpha=10^{-12}$ close to 0 .

To conclude this section we wish to briefly discuss the matrix case $r=2$, and point out similarities and differences with genuine tensors $r \geq 3$. In the matrix case, [98] observe for $a$ set of centred priors and odd activations that the asymptotic matrix-MMSE is equal to its maximum $\rho_{x}^{2}$ for $\lambda<\lambda_{c}(\alpha)$ and decreases for $\lambda>\lambda_{c}(\alpha)$ while remaining continuous at $\lambda_{c}(\alpha)$. Again, $\lambda_{c}(\alpha)$ decreases with increasing $\alpha$. We give an example on the left panel of Figure 4 . The continuity of the phase transition is an important qualitative difference with what we observe here for order-3 tensors. Such continuity for Bayesian inference problems is known to go hand in hand with the optimality of approximate message passing


Figure 2: Asymptotic tensor-MMSE for $r=3, P_{S}=\mathcal{N}(0,1)$ and $\varphi(z)=\operatorname{sign}(z)$ as a function of $\lambda$. The location $\lambda_{c}(\alpha)$ of the discontinuity decreases with increasing $\alpha$. For $\alpha=10^{-12}$ the threshold $\lambda_{c}(\alpha) \approx 7.07$ is the same than for the i.i.d. case with Rademacher prior $X_{1}, \ldots, X_{n} \stackrel{\text { i.i.d. }}{\sim} P_{X}( \pm 1)=\frac{1}{2}$ (whose asymptotic MMSE is given by the curve " $\alpha \rightarrow 0^{+"}$ ).


Figure 3: Asymptotic tensor-MMSE for $\varphi(z)=z$ and an asymmetric Rademacher prior $P_{S}(1)=1-P_{S}(-1)=p$.
(AMP) and, as shown in [98], matrix factorization with generative prior is no exception. Because the continuity of the phase transition is observed for all the priors and activations used in [98], it supports the claim that such model of structure makes estimation algorithmically easier. For order-3 tensor estimation, there is an infinite computational-to-statistical gap when the spike $\mathbf{X}$ has i.i.d. entries, i.e., the MMSE becomes nontrivial (lower than its maximum possible value) for $\lambda$ above a critical value $\lambda_{c}=\Theta(1)$ while the MSE of AMP becomes nontrivial above the algorithmic threshold $\lambda_{\text {AMP }}=\Theta(n)$ [19]. In this regard, there are polynomial-time algorithms with an algorithmic threshold in $\Theta(\sqrt{n})$, hence beating $A M P$. These algorithms are based on sum-of-squares relaxation [126], power iteration with spectral initialization applied to tensor matricizations [19], [127], or better approximation of the free energy than the Bethe free energy [128] (note that AMP tries to minimize a high-dimensional approximation of the Bethe free energy). Recently, [127, [129], 130] showed that $\lambda=\Omega(\sqrt{n})$ is an essential requirement for polynomial-time algorithms. Their arguments are based on an average-case reduction to a hypergraphic planted clique (HPC) problem and hold under the hardness hypothesis of HPC detection. Given the observations in [98], we might hope that the generative prior of the spike $\mathbf{X}$ makes the estimation


Figure 4: Asymptotic matrix-MMSE when estimating $\mathbf{X}^{\otimes 2}$ from $\mathbf{Y}=\sqrt{\lambda / n} \mathbf{X}^{\otimes 2}+\mathbf{Z}$. We use a Bernoulli-Rademacher prior $P_{S}(0)=1-\rho, P_{S}( \pm 1)=\rho / 2$ with $\rho=0.05$. Left: generative prior $\mathbf{X}=\mathbf{W S} / \sqrt{p}$ with $\mathbf{S} \stackrel{\text { i.i.d. }}{\sim} P_{S} . \operatorname{Right}: \mathbf{X} \stackrel{\text { i.i.d. }}{\sim} P_{S}$.
of the corresponding tensor of order $r \geq 3$ algorithmically easier. However, the persisting discontinuity of the transition at $\lambda_{c}(\alpha)$ and asymptotic stability of the uninformative fixed point suggest that it is not the case. The observations of 98 should also be nuanced as it is not difficult to come up with a situation where the phase transition is discontinuous. E.g., consider the spiked matrix model with generative prior $\mathbf{X}=\varphi(\mathbf{W s} / \sqrt{p})$ for the odd activation function $\varphi(x)=0$ if $|x| \leq \epsilon$ and $\varphi(x)=\operatorname{sign}(x)$ otherwise, and the centered latent prior $P_{S}=\mathcal{N}(0,1)$. Similarly to what is done in Section 5.6, we can show that when $\alpha$ vanishes the asymptotic matrix-MMSE approaches the one for the spiked matrix model

$$
\widetilde{Y}_{i j}:=\sqrt{\frac{\lambda}{n}} \widetilde{X}_{i} \widetilde{X}_{j}+\widetilde{Z}_{i j}, \quad 1 \leq i \leq j \leq n
$$

where $\widetilde{X}_{1}, \ldots, \widetilde{X}_{n} \stackrel{\text { i.i.d. }}{\sim} \varphi(\mathcal{N}(0,1))$ are i.i.d. Bernoulli-Rademacher random variables. We can make $\mathbb{P}\left(\widetilde{X}_{i}=0\right)=1-2 \mathbb{P}(\mathcal{N}(0,1)<-\epsilon)=1-\rho$ as large as needed by increasing $\epsilon$ (then $\left.\mathbb{P}\left(\widetilde{X}_{i}=1\right)=\mathbb{P}\left(\widetilde{X}_{i}=-1\right)=\rho / 2\right)$. It is known that the asymptotic matrix-MMSE has a jump discontinuity for such prior when the probability of being 0 is large enough, e.g., see the right panel in Figure 4. Therefore, when $\epsilon$ is large enough, the asymptotic matrix-MMSE of the original spiked matrix model with generative prior also has a jump discontinuity, at least for small $\alpha$. An interesting question for future research is whether or not the discontinuity disappears when $\alpha$ is made large enough. If so, it would further support the claim that such generative prior makes estimation algorithmically easier when the ratio $\alpha$ of signal-to-latent space dimensions is large enough. If not, the existence of a jump discontinuity would then merely depend on the choice of activation function and not on the ratio of signal-to-latent space dimensions.

### 5.4 Proof of the variational formula for the mutual information

In this section we present the main steps of the proof of Theorem 5.1. Intermediate results are found in the appendices.

### 5.4.1 Adaptive path interpolation

We introduce a parameter $t \in[0,1]$. The adaptive interpolation interpolates from the original model (5.1) at $t=0$ to a GLM whose asymptotic mutual information is known $\sqrt[29]{ }$ at $t=1$. In between, we follow an interpolation path $R(\cdot, \epsilon):[0,1] \rightarrow(0,+\infty)$ which is a continuously differentiable function of $t$ parametrized by a "small perturbation" $\epsilon \in(0,+\infty)$ and is such that $R(0, \epsilon)=\epsilon$. More precisely, for $t \in[0,1]$, the observations are

$$
\left\{\begin{array}{l}
\mathbf{Y}^{(t)}:=\frac{\sqrt{\lambda(1-t)}}{n} \mathbf{X}^{\otimes 3}+\mathbf{Z}  \tag{5.8}\\
\widetilde{\mathbf{Y}}^{(t, \epsilon)}:=\sqrt{\frac{\lambda R(t, \epsilon)}{2}} \mathbf{X}+\widetilde{\mathbf{Z}}
\end{array},\right.
$$

where $\mathbf{X}:=\varphi(\mathbf{W s} / \sqrt{p}), \widetilde{\mathbf{Z}} \in \mathbb{R}^{n}$ is a standard Gaussian random vector, and $\mathbf{Z} \in\left(\mathbb{R}^{n}\right)^{\otimes 3}$ is a symmetric random tensor with entries $\mathbf{Z}_{\underline{i}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ for triplets $\underline{i}$ in the subset

$$
\mathcal{I}:=\left\{\left(i_{1}, i_{2}, i_{3}\right) \in\{1, \ldots, n\}: i_{1} \leq i_{2} \leq i_{3}\right\} .
$$

Before diving into the proof, we introduce some important quantities and notations. We denote $i_{n}(t, \epsilon)$ the normalized mutual information between $\mathbf{X}$ and $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)$ given $\mathbf{W}$, that is,

$$
\begin{equation*}
i_{n}(t, \epsilon):=\frac{1}{n} I\left(\mathbf{X} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)} \mid \mathbf{W}\right)=\frac{1}{n} I\left(\mathbf{S} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)} \mid \mathbf{W}\right) . \tag{5.9}
\end{equation*}
$$

The last equality holds because $\mathbf{X}$ is a deterministic function of $\mathbf{S}$ when $\mathbf{W}$ is known. We denote $d P_{S}(\mathbf{s}):=\prod_{i=1}^{p} d P_{S}\left(s_{i}\right)$ the prior distribution of $\mathbf{S}$. The Bayes posterior distribution of $\mathbf{S}$ given $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)$ reads

$$
\begin{equation*}
d P\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right):=\frac{d P_{S}(\mathbf{s}) e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)}}{\mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)} \tag{5.10}
\end{equation*}
$$

where

$$
\begin{equation*}
\mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right):=\int d P_{S}(\mathbf{s}) e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)} \tag{5.11}
\end{equation*}
$$

is a normalization factor, and

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right):=\sum_{\underline{i} \in \mathcal{I}} & \left(\frac{\lambda(1-t)}{2 n^{2}} x_{i_{1}}^{2} x_{i_{2}}^{2} x_{i_{3}}^{2}-\frac{\sqrt{\lambda(1-t)}}{n} Y_{\underline{i}}^{(t)} x_{i_{1}} x_{i_{2}} x_{i_{3}}\right) \\
& +\sum_{j=1}^{n}\left(\frac{\lambda R(t, \epsilon)}{4} x_{j}^{2}-\sqrt{\frac{\lambda R(t, \epsilon)}{2}} \widetilde{Y}_{j}^{(t, \epsilon)} x_{j}\right), \tag{5.12}
\end{align*}
$$

with $x_{1}, \ldots, x_{n}$ the entries of $\mathbf{x}:=\varphi\left(\mathrm{W} /{ }_{\sqrt{p}}\right)$. This dependence on $\mathbf{s}$ must be kept in mind each time we use the notation $\mathbf{x}$. It is common to adopt the
statistical mechanics interpretation and call (5.12) a Hamiltonian, (5.11) the partition function, and 5.10 the Gibbs distribution.

To deal with future computations, it is useful to introduce the angular brackets $\langle-\rangle_{t, \epsilon}$ that denote an expectation with respect to the posterior distribution 5.10, that is,

$$
\begin{equation*}
\langle g(\mathbf{s})\rangle_{t, \epsilon}:=\int g(\mathbf{s}) d P\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right) \tag{5.13}
\end{equation*}
$$

for a generic function $g: \mathbb{R}^{p} \rightarrow \mathbb{R}$. Finally, we define the so-called average free entropy

$$
\begin{equation*}
f_{n}(t, \epsilon):=\frac{1}{n} \mathbb{E} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right) \tag{5.14}
\end{equation*}
$$

This is equal to the mutual information $i_{n}(t, \epsilon)$ up to some additive term (see formula (5.43) in Lemma 5.10 in Appendix 5.B). It is often easier to work directly with $f_{n}(t, \epsilon)$ instead of $i_{n}(t, \epsilon)$.

We now focus on the mutual information (5.9) at both extremes of the interpolation path. Letting $t=0$ in (5.8), we see that the observation $\mathbf{Y}^{(0)}$ is exactly (5.1) while $\widetilde{\mathbf{Y}}^{(0, \epsilon)}=\sqrt{\frac{\lambda \epsilon}{2}} \mathbf{X}+\widetilde{\mathbf{Z}}$. The latter channel induces a perturbation of the normalized mutual information associated with the former channel of the order of $\epsilon$ (see Lemma 5.10 in Appendix $5 . B$ for the proof), that is,

$$
\begin{equation*}
i_{n}(0, \epsilon):=\frac{I\left(\mathbf{X} ; \mathbf{Y}^{(0)}, \widetilde{\mathbf{Y}}^{(0, \epsilon)} \mid \mathbf{W}\right)}{n}=\frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n}+O(\epsilon) \tag{5.15}
\end{equation*}
$$

where $|O(\epsilon)| \leq C \epsilon$. At $t=1$, the observation $\mathbf{Y}^{(1)}$ is pure noise while the normalized mutual information between $\mathbf{S}$ and $\widetilde{\mathbf{Y}}^{(1, \epsilon)}=\sqrt{\lambda R(1, \epsilon) / 2} \varphi(\mathrm{WS} / \sqrt{\bar{p}})+\widetilde{\mathbf{Z}}$ is given by a variational formula in the high-dimensional regime $n / p \rightarrow \alpha$ [29]. Let $S \sim P_{S}$ and $U, V, Z, \widetilde{Z} \sim \mathcal{N}(0,1)$ be independent scalar random variables. Define the potential function $\widetilde{\psi}_{\alpha}:[0,+\infty)^{2} \times\left[0, \rho_{s}\right]$ :

$$
\begin{equation*}
\widetilde{\psi}_{\alpha}\left(r, r_{s}, q_{s}\right):=I_{P_{S}}\left(r_{s}\right)+\alpha I_{\varphi}\left(r, q_{s} ; \rho_{s}\right)-\frac{r_{s}\left(\rho_{s}-q_{s}\right)}{2} \tag{5.16}
\end{equation*}
$$

By [29, Corollary 1], we have

$$
\begin{equation*}
i_{n}(1, \epsilon)=\frac{I\left(\mathbf{X} ; \widetilde{\mathbf{Y}}^{(1, \epsilon)} \mid \mathbf{W}\right)}{n}=o_{n}(1)+\frac{1}{\alpha} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(\frac{\lambda R(1, \epsilon)}{2}, r_{s}, q_{s}\right) \tag{5.17}
\end{equation*}
$$

Combining (5.15), (5.17), and the fundamental theorem of calculus $i_{n}(0, \epsilon)=$ $i_{n}(1, \epsilon)-\int_{0}^{1} i_{n}^{\prime}(t, \epsilon) d t$, where $i_{n}^{\prime}(\cdot, \epsilon)$ is the derivative of $i_{n}(\cdot, \epsilon)$, we obtain the sum-rule of the adaptive interpolation.

Proposition 5.3 (Sum-rule). Suppose that (H1) and (H2) hold, and that $R^{\prime}(t, \epsilon)$ is uniformly bounded in $(t, \epsilon) \in[0,1] \times[0,+\infty)$ where $R^{\prime}(\cdot, \epsilon)$ denotes the derivative of $R(\cdot, \epsilon)$ with respect to its first argument. Define the scalar overlap

$$
Q:=\frac{1}{n} \sum_{i=1}^{n} \varphi\left([\mathrm{ws} / \sqrt{\bar{p}}]_{i}\right) \varphi\left([\mathrm{ws} / \sqrt{\bar{p}}]_{i}\right)=\frac{1}{n} \sum_{i=1}^{n} x_{i} X_{i} .
$$

Then,

$$
\begin{align*}
\frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} & =O(\epsilon)+o_{n}(1)+\frac{1}{\alpha} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(\frac{\lambda R(1, \epsilon)}{2}, r_{s}, q_{s}\right) \\
& -\frac{\lambda}{12} \int_{0}^{1}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\rho_{x}^{3}\right) d t-\frac{\lambda}{4} \int_{0}^{1} R^{\prime}(t, \epsilon)\left(\rho_{x}-\mathbb{E}\langle Q\rangle_{t, \epsilon}\right) d t \tag{5.18}
\end{align*}
$$

where $O_{n}(1)$ and $O(\epsilon)$ are independent of $\epsilon$ and $n$, respectively.
Proof. See Lemma 5.11 in Appendix 5.B for the computation of the derivative $i_{n}^{\prime}(t, \epsilon)$.

The sum rule of Proposition 5.3 is valid for the general class of differentiable interpolating paths. By choosing two appropriate interpolation paths we can prove matching upper and lower bounds on the asymptotic normalized mutual information. This is discussed in the next two paragraphs.

### 5.4.2 Upper bound on the asymptotic normalized mutual information

Proposition 5.4. Suppose that (H1) and (H2) hold. Then,

$$
\limsup _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \leq \inf _{q_{x} \in\left[0, \rho_{x}\right]} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right)
$$

Proof. Fix $\epsilon>0$ and pick the linear interpolation path $R(t, \epsilon)=\epsilon+t q^{2}$ where $q \in\left[0, \rho_{x}\right]$. Then, the sum-rule (5.18) in Proposition 5.3 reads

$$
\begin{align*}
\frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n}=O(\epsilon) & +o_{n}(1)+\frac{1}{\alpha} \inf _{q_{s} \in\left[0, \rho_{s}\right] r_{s} \geq 0} \sup _{\alpha} \widetilde{\psi}_{\alpha}\left(\frac{\lambda \epsilon}{2}+\frac{\lambda q^{2}}{2}, r_{s}, q_{s}\right) \\
+ & \frac{\lambda}{12} \rho_{x}^{3}-\frac{\lambda}{4} q^{2} \rho_{x}-\frac{\lambda}{12} \int_{0}^{1}\left(\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{4}\langle Q\rangle_{t, \epsilon}\right]-3 q^{2} \mathbb{E}\langle Q\rangle_{t, \epsilon}\right) d t \\
& -\frac{\lambda}{12} \int_{0}^{1}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{4}\langle Q\rangle_{t, \epsilon}\right]\right) d t \tag{5.19}
\end{align*}
$$

In this last identity, we "artificially" added and subtracted the expectation $\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{4}\langle Q\rangle_{t, \epsilon}\right]$ for reasons that will appear immediately. By the Nishimori identity, we have

$$
\begin{equation*}
\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{4}\langle Q\rangle_{t, \epsilon}\right]=\mathbb{E}\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{6}, \quad \mathbb{E}\langle Q\rangle_{t, \epsilon}=\mathbb{E}\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}, \tag{5.20}
\end{equation*}
$$

and, by convexity of $x \mapsto x^{3}$ on $[0,+\infty)$, we have $\forall a, b \geq 0: a^{3}-3 b^{2} a \geq-2 b^{3}$. Hence, the integrand of the last integral on the right-hand side of (5.19) satisfies

$$
\begin{equation*}
\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{4}\langle Q\rangle_{t, \epsilon}\right]-3 q^{2} \mathbb{E}\langle Q\rangle_{t, \epsilon}=\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{6}-3 q^{2}\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}\right] \geq-2 q^{3} \tag{5.21}
\end{equation*}
$$

Besides, by Lemma 5.9 in Appendix 5.A, $r \mapsto \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \tilde{\psi}_{\alpha}\left(r, r_{s}, q_{s}\right)$ is nondecreasing and ( $\alpha\|\varphi\|_{\infty}^{2} / 2$ )-Lipschitz continuous on $[0,+\infty)$. Therefore,

$$
\begin{equation*}
\inf _{q_{s} \in\left[0, \rho_{s}\right] r_{r_{s} \geq 0}}^{\sup _{r_{2}}} \widetilde{\psi}_{\alpha}\left(\frac{\lambda \epsilon}{2}+\frac{\lambda q^{2}}{2}, r_{s}, q_{s}\right) \leq \frac{\lambda \alpha\|\varphi\|_{\infty}^{2}}{4} \epsilon+\inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(\frac{\lambda q^{2}}{2}, r_{s}, q_{s}\right) . \tag{5.22}
\end{equation*}
$$

Making use of (5.21) and (5.22) to upper bound (5.19) yields

$$
\begin{align*}
\frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \leq & O(\epsilon)+ \\
& +o_{n}(1) \\
& +\inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \frac{1}{\alpha} \tilde{\psi}_{\alpha}\left(\frac{\lambda q^{2}}{2}, r_{s}, q_{s}\right)+\frac{\lambda}{12} \rho_{x}^{3}-\frac{\lambda}{4} q^{2} \rho_{x}+\frac{\lambda}{6} q^{3} \\
& \quad-\frac{\lambda}{12} \int_{0}^{1}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{4}\langle Q\rangle_{t, \epsilon}\right]\right) d t \\
= & O(\epsilon)+o_{n}(1)+\inf _{\left.q_{s} \in\left[0, \rho_{s}\right]\right]_{s} \geq 0} \sup _{r_{, ~}}\left(q, q_{s}, r_{s}\right)  \tag{5.23}\\
& \quad-\frac{\lambda}{12} \int_{0}^{1}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\| \|^{4}\langle Q\rangle_{t, \epsilon}\right]\right) d t
\end{align*}
$$

where the last equality follows from the trivial identity

$$
\begin{equation*}
\psi_{\lambda, \alpha}\left(q, q_{s}, r_{s}\right)=\frac{1}{\alpha} \widetilde{\psi}_{\alpha}\left(\frac{\lambda q^{2}}{2}, r_{s}, q_{s}\right)+\frac{\lambda}{12} \rho_{x}^{3}-\frac{\lambda}{4} q^{2} \rho_{x}+\frac{\lambda}{6} q^{3} . \tag{5.24}
\end{equation*}
$$

It now remains to get rid of the integral on the right-hand side of 5.23). The integrand satisfies

$$
\begin{align*}
& \left|\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{4}\langle Q\rangle_{t, \epsilon}\right]\right|=\left|\mathbb{E}\left\langle Q\left(Q+\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}\right)\left(Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}\right)\right\rangle_{t, \epsilon}\right| \\
& \quad \leq 2\|\varphi\|_{\infty}^{4} \mathbb{E}\langle | Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}| \rangle_{t, \epsilon} \leq 2\|\varphi\|_{\infty}^{4} \sqrt{\mathbb{E}\left\langle\left(Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}\right)^{2}\right\rangle_{t, \epsilon}} . \tag{5.25}
\end{align*}
$$

We see that if the overlap $Q:=\mathbf{x}^{\top} \mathbf{x} / n$ would concentrate on $\langle\mathbf{x}\rangle_{t, \epsilon}^{\top}\langle\mathbf{x}\rangle_{t, \epsilon} / n$ then the remaining integral in (5.23) would be negligible. However, such a concentration property only holds when we average on a well-chosen set of perturbations $\epsilon$. In essence, the average over $\epsilon$ smoothens the phase transitions that might appear for particular choices of $\epsilon$ when $n$ goes to infinity. We now take $\epsilon \in\left[s_{n}, 2 s_{n}\right]$ where $s_{n}:=n^{-\eta}, \eta>0$, and average both sides of (5.23) over $\epsilon$. We get

$$
\begin{aligned}
\frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n}= & \int_{s_{n}}^{2 s_{n}} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \frac{d \epsilon}{s_{n}} \\
\leq & O\left(s_{n}\right)+o_{n}(1)+\inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q, q_{s}, r_{s}\right) \\
& \quad-\frac{\lambda}{12} \int_{0}^{1} d t \int_{s_{n}}^{2 s_{n}}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\mathbb{E}\left[\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{4}\langle Q\rangle_{t, \epsilon}\right]\right) \frac{d \epsilon}{s_{n}}
\end{aligned}
$$

$$
\begin{align*}
& \leq O\left(s_{n}\right)+o_{n}(1)+\inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q, q_{s}, r_{s}\right) \\
& \quad+\frac{\lambda\|\varphi\|_{\infty}^{4}}{6} \int_{0}^{1} d t \int_{s_{n}}^{2 s_{n}} \sqrt{\mathbb{E}\left\langle\left(Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}\right)^{2}\right\rangle_{t, \epsilon}} \frac{d \epsilon}{s_{n}} . \tag{5.26}
\end{align*}
$$

Let us focus on the integral over $\epsilon$ in the last term on the right-hand side of (5.26). Let $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}\right)$ be observations defined exactly as $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right.$ ) in (5.8) except that the nonnegative real number $R$ replaces $R(t, \epsilon)$. We denote by $\langle-\rangle_{t, R}$ the expectation with respect to a sample $\mathbf{s}$ drawn from the posterior distribution of $\mathbf{S}$ given $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)$ (see (5.29) in Subsection 5.4.3). Since $R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $\left[s_{n}, 2 s_{n}\right]$ to its image $R\left(t,\left[s_{n}, 2 s_{n}\right]\right)=\left[s_{n}+t q^{2}, 2 s_{n}+t q^{2}\right]$, we make the change of variables $\epsilon \rightarrow R=R(t, \epsilon)$ and obtain for all $t \in[0,1]$ :

$$
\begin{align*}
\int_{s_{n}}^{2 s_{n}} \sqrt{\mathbb{E}\left\langle\left(Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}\right)^{2}\right\rangle_{t, \epsilon} \frac{d \epsilon}{s_{n}}} & \leq \sqrt{\int_{s_{n}}^{2 s_{n}} \mathbb{E}\left\langle\left(Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}\right)^{2}\right\rangle_{t, \epsilon} \frac{d \epsilon}{s_{n}}} \\
& =\sqrt{\int_{s_{n}+t q^{2}}^{2 s_{n}+t q^{2}} \mathbb{E}\left\langle\left(Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, R}}{\sqrt{n}}\right\|^{2}\right)^{2}\right\rangle_{t, R} \frac{d R}{s_{n}}}, \tag{5.27}
\end{align*}
$$

where the first inequality is due to Cauchy-Schwarz inequality. By Proposition 5.14 in Appendix 5.C and the inequality (5.27), we get (remember that $s_{n}:=n^{-\eta}$ )

$$
\begin{aligned}
\frac{\lambda\|\varphi\|_{\infty}^{4}}{6} \int_{s_{n}}^{2 s_{n}} \sqrt{\mathbb{E}\left\langle\left(Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, \epsilon}}{\sqrt{n}}\right\|^{2}\right)^{2}\right\rangle_{t, \epsilon}} \frac{d \epsilon}{s_{n}} & \leq \frac{\sqrt{\lambda}\|\varphi\|_{\infty}^{4}}{3} \sqrt{\frac{\|\varphi\|_{\infty}^{3}}{s_{n}} \sqrt{\frac{\lambda s_{n}}{2 n}}} \\
& =\frac{\lambda^{\frac{3}{4}}\|\varphi\|_{\infty}^{11 / 2}}{3 \cdot 2^{1 / 4}} n^{\frac{\eta-1}{4}}
\end{aligned}
$$

Therefore, if we pick $\eta=1 / 5$, the remaining integral on the right-hand side of (5.26) vanishes as $O\left(n^{-1 / 5}\right)=O\left(s_{n}\right)$ and (5.26) simply reads

$$
\frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \leq O\left(s_{n}\right)+o_{n}(1)+\inf _{q_{s} \in\left[0, \rho_{s}\right]_{r_{s} \geq 0}}^{\sup _{\lambda}} \psi_{\lambda, \alpha}\left(q, q_{s}, r_{s}\right)
$$

Finally, passing to the limit superior on both sides of the latter inequality yields

$$
\limsup _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \leq \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q, q_{s}, r_{s}\right)
$$

This inequality is true for all $q \in\left[0, \rho_{x}\right]$ and Proposition 5.4 follows directly.

### 5.4.3 Matching lower bound on the asymptotic normalized mutual information

We now prove a matching lower bound by considering a different choice for $R(\cdot, \epsilon)$ in the sum-rule (5.18). The interpolation path $R(\cdot, \epsilon)$ is the solution to a first-order ordinary differential equations (ODE). We first describe this ODE and then derive the lower bound.

## An ordinary differential equation

Let $\mathbf{S}$ be a random vector with entries $S_{1}, \ldots, S_{p} \stackrel{\text { i.i.d. }}{\sim} P_{S}$. For $t \in[0,1]$ and $R \in[0,+\infty)$, consider the problem of estimating $\mathbf{S}$ from the observations

$$
\begin{cases}\mathbf{Y}^{(t)} & :=\frac{\sqrt{\lambda(1-t)}}{n} \mathbf{X}^{\otimes 3}+\mathbf{Z}  \tag{5.28}\\ \widetilde{\mathbf{Y}}^{(t, R)} & :=\sqrt{\frac{\lambda R}{2}} \mathbf{X}+\widetilde{\mathbf{Z}}\end{cases}
$$

where $\mathbf{X}:=\varphi(\mathbf{W S} / \sqrt{p}), \widetilde{\mathbf{Z}} \in \mathbb{R}^{n}$ is a standard Gaussian random vector, and $\mathbf{Z} \in\left(\mathbb{R}^{n}\right)^{\otimes 3}$ is a symmetric noise tensor with entries $\mathbf{Z}_{\underline{i}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ for triplets $\underline{\boldsymbol{i}} \in \mathcal{I}$. The Bayes posterior distribution of $\mathbf{S}$ given $\left(\mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)$ is

$$
\begin{equation*}
d P\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right):=\frac{d P_{S}(\mathbf{s}) e^{-\mathcal{H}_{t, R}\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)}}{\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)} \tag{5.29}
\end{equation*}
$$

where $\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right):=\int d P_{S}(\mathbf{s}) e^{-\mathcal{H}_{t, R}\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)}$ and

$$
\begin{align*}
& \mathcal{H}_{t, R}\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right):=\sum_{\underline{i} \in \mathcal{I}} \frac{\lambda(1-t)}{2 n^{2}} x_{i_{1}}^{2} x_{i_{2}}^{2} x_{i_{3}}^{2}-\frac{\sqrt{\lambda(1-t)}}{n} Y_{\underline{i}}^{(t)} x_{i_{1}} x_{i_{2}} x_{i_{3}} \\
&+\sum_{j=1}^{n} \frac{\lambda R}{4} x_{j}^{2}-\sqrt{\frac{\lambda R}{2}} \widetilde{Y}_{j}^{(t, R)} x_{j} . \tag{5.30}
\end{align*}
$$

Again, (5.30) has the interpretation of a Hamiltonian and (5.29) of a Gibbs distribution. The angular brackets $\langle-\rangle_{t, R}$ denote the expectation with respect to the posterior (5.29). Finally, we define the following function used to formulate the first-order ODE satisfied by the interpolation path,

$$
\begin{equation*}
G(t, R):=\left(\mathbb{E}\langle Q\rangle_{t, R}\right)^{2} . \tag{5.31}
\end{equation*}
$$

Lemma 5.5. Assume $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ is continuous and bounded. For all $\epsilon \in[0,+\infty)$, there exists a unique global solution, denoted $R(\cdot, \epsilon):[0,1] \rightarrow[0,+\infty)$, to the initial value problem

$$
\begin{equation*}
R^{\prime}=G(t, R), \quad R(0)=\epsilon . \tag{5.32}
\end{equation*}
$$

$R(\cdot, \epsilon)$ is continuously differentiable with bounded derivative $R^{\prime}(\cdot, \epsilon)$ and, for any $\delta>0, R^{\prime}([0,1], \epsilon) \subseteq\left[0,\left(\rho_{x}+\delta\right)^{2}\right]$ for $n$ large enough independent of $\epsilon$. Besides, for all $t \in[0,1], R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $[0,+\infty)$ into its image whose derivative w.r.t. $\epsilon$ is greater than or equal to one, i.e.,

$$
\forall \epsilon \in[0,+\infty):\left.\frac{\partial R}{\partial \epsilon}\right|_{t, \epsilon} \geq 1
$$

Remark. Lemma 5.5 guarantees a unique global solution $R_{n}(t, \epsilon)$ for each finite $n$. Slightly abusively, we do not explicitly indicate the dependence on $n$ and simply write $R(t, \epsilon)$ for the solution.

Proof. The function $G:(t, R) \in[0,1] \times[0,+\infty) \mapsto G(t, R)$ is continuous in $t$ and uniformly Lipschitz continuous in $R$ (meaning that the Lipschitz constant is independent of $t$ ). The latter is readily checked by showing that the partial derivative $\partial G / \partial R$ is uniformly bounded in $(t, R)$; we have

$$
\begin{equation*}
\left.\frac{\partial G}{\partial R}\right|_{t, R}=\frac{\lambda \mathbb{E}\langle Q\rangle_{t, R}}{n} \sum_{i, j=1}^{n} \mathbb{E}\left[\left(\left\langle x_{i} x_{j}\right\rangle_{t, R}-\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right)^{2}\right] \tag{5.33}
\end{equation*}
$$

so ${ }^{\partial G} /\left.\partial R\right|_{t, R} \in\left[0,4 \lambda\|\varphi\|_{\infty}^{6} n\right]$. Therefore, by the Picard-Lindelöf theorem 104 , Theorem 1.1], for all $\epsilon \geq 0$ there exists a unique solution $R(\cdot, \epsilon):[0, \gamma] \rightarrow[0,+\infty)$ to the initial value problem (5.32). Here $\gamma \in[0,1]$ is such that $[0, \gamma]$ is the maximal interval of existence of the solution. By the Cauchy-Schwarz inequality and Nishimori identity, we have

$$
\begin{aligned}
\mathbb{E}\langle Q\rangle_{t, R} \leq \frac{\mathbb{E}\langle\|\mathbf{x}\|\|\mathbf{X}\|\rangle_{t, R}}{n} & \leq \frac{1}{n} \sqrt{\mathbb{E}\left\langle\|\mathbf{x}\|^{2}\right\rangle_{t, R} \mathbb{E}\|\mathbf{X}\|^{2}} \\
& =\frac{\mathbb{E}\|\mathbf{X}\|^{2}}{n}=\mathbb{E}\left[\varphi\left(\frac{\mathbf{W}_{1, \cdot} \mathbf{S}}{\sqrt{p}}\right)^{2}\right] \underset{n \rightarrow+\infty}{ } \rho_{x}
\end{aligned}
$$

where in the last equality we denote by $\mathbf{W}_{1, \text {. }}$ the first row of $\mathbf{W}$. See 75, Lemma 3 of Supplementary material] for a proof of the latter limit. Besides, by the Nishimori identity, $\mathbb{E}\langle Q\rangle_{t, R}=n^{-1} \mathbb{E}\left\|\langle\mathbf{x}\rangle_{t, R}\right\|^{2}$ is nonnegative. Hence, for any $\delta>0$, $G$ has its image in $\left[0,\left(\rho_{x}+\delta\right)^{2}\right]$ and $R([0, \gamma], \epsilon) \subseteq\left[\epsilon, \epsilon+\gamma\left(\rho_{x}+\delta\right)^{2}\right]$ as long as $n$ is large enough. It implies that $\gamma=1$ (the solution never leaves the domain of definition of $G$ ).

Each initial condition $\epsilon \in[0,+\infty)$ is tied to a unique solution $R(\cdot, \epsilon)$. This implies that the function $\epsilon \mapsto R(t, \epsilon)$ is injective. Its derivative is given by Liouville's formula 104

$$
\left.\frac{\partial R}{\partial \epsilon}\right|_{t, \epsilon}=\exp \left\{\left.\int_{0}^{t} d s \frac{\partial G}{\partial R}\right|_{s, R(s, \epsilon)}\right\}
$$

and is greater than, or equal to one, by nonnegativity of $\frac{\partial G}{\partial R}$ (see formula (5.33) above). The fact that this partial derivative is bounded away from 0 uniformly in $\epsilon$ implies, by the inverse function theorem, that the injective function $\epsilon \mapsto R(t, \epsilon)$ is a $C^{1}$-diffeomorphism from $[0,+\infty)$ onto its image.

## Derivation of the lower bound

Proposition 5.6. Suppose that (H1) and (H2) hold. Then,

$$
\liminf _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \geq \inf _{q_{x} \in\left[0, \rho_{x}\right]} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right)
$$

Proof. For all $\epsilon \in[0,+\infty)$, the interpolation path is the unique solution $R(\cdot, \epsilon)$ to the initial value problem (5.32). Fix $\nu>0$. Let $n$ be large enough so
that $\forall \epsilon \in[0,+\infty): R^{\prime}(\cdot, \epsilon) \subseteq\left[0,\left(\rho_{x}+\nu\right)^{2}\right]$. The interpolation path satisfies $R^{\prime}(t, \epsilon)=\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{2}$ so the sum-rule of Proposition 5.3 reads

$$
\begin{align*}
& \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n}=O(\epsilon)+ o_{n}(1) \\
&+ \frac{1}{\alpha} \inf _{q_{s} \in\left[0, \rho_{s}\right] r_{s} \geq 0} \sup _{\psi_{\alpha}} \widetilde{\psi}_{\alpha}\left(\frac{\lambda \epsilon}{2}+\int_{0}^{1} \frac{\lambda R^{\prime}(t, \epsilon)}{2} d t, r_{s}, q_{s}\right) \\
&+\int_{0}^{1}\left(\frac{\lambda}{12} \rho_{x}^{3}+\frac{\lambda}{6}\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{3}-\frac{\lambda}{4}\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{2} \rho_{x}\right) d t  \tag{5.34}\\
&-\frac{\lambda}{12} \int_{0}^{1}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{3}\right) d t
\end{align*}
$$

By Lemma 5.9 in Appendix 5.A, the function $r \mapsto \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(r, r_{s}, q_{s}\right)$ is nondecreasing and concave. Therefore,

$$
\begin{align*}
& \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(\frac{\lambda \epsilon}{2}+\int_{0}^{1} \frac{\lambda R^{\prime}(t, \epsilon)}{2} d t, r_{s}, q_{s}\right) \\
& \geq \int_{0}^{1} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(\frac{\lambda R^{\prime}(t, \epsilon)}{2}, r_{s}, q_{s}\right) d t \tag{5.35}
\end{align*}
$$

Combining the identity (5.34) with (5.35) yields

$$
\begin{align*}
& \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \geq O(\epsilon)+o_{n}(1)+\int_{0}^{1}\left\{\begin{array}{l}
\inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \frac{1}{\alpha} \widetilde{\psi}_{\alpha}\left(\frac{\lambda\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{2}}{2}, r_{s}, q_{s}\right) \\
\\
\\
\left.\quad+\frac{\lambda \rho_{x}^{3}}{12}+\frac{\lambda\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{3}}{6}-\frac{\lambda\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{2} \rho_{x}}{4}\right\} d t \\
\\
\\
\quad-\frac{\lambda}{12} \int_{0}^{1}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{3}\right) d t
\end{array}\right. \\
& \geq O(\epsilon)+o_{n}(1)+\inf _{q_{x} \in\left[0, \rho_{x}+\nu\right] q_{s} \in\left[0, \rho_{s}\right]} \inf _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right) \\
&-\frac{\lambda}{12} \int_{0}^{1}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{3}\right) d t
\end{align*}
$$

The second inequality follows from identity (5.24) and $\mathbb{E}\langle Q\rangle_{t, \epsilon} \in\left[0, \rho_{x}+\nu\right]$.
The result of the proposition will follow if we can get rid of the integral term on the right-hand side of (5.36) This is achieved by proceeding exactly as in the proof of the upper bound in Section 5.4.2, that is, we integrate (5.36) over $\epsilon \in\left[s_{n}, 2 s_{n}\right]$ where $s_{n}=n^{-\eta}, \eta>0$. Then,

$$
\begin{align*}
\frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n}= & \int_{s_{n}}^{2 s_{n}} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \frac{d \epsilon}{s_{n}} \\
\geq & O\left(s_{n}\right)+o_{n}(1)+\inf _{q_{x} \in\left[0, \rho_{x}+\nu\right]} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right) \\
& \quad-\frac{\lambda}{12} \int_{0}^{1} d t \int_{s_{n}}^{2 s_{n}} \frac{d \epsilon}{s_{n}}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{3}\right) \\
\geq & O\left(s_{n}\right)+o_{n}(1)+\inf _{q_{x} \in\left[0, \rho_{x}+\nu\right]} \inf _{q_{s} \in\left[0, \rho_{s}\right] r_{s} \geq 0} \sup _{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right) \\
& \quad-\frac{\lambda\|\varphi\|_{\infty}^{4}}{6} \int_{0}^{1} d t \int_{s_{n}}^{2 s_{n}} \frac{d \epsilon}{s_{n}} \sqrt{\mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{2}\right\rangle_{t, \epsilon}} \tag{5.37}
\end{align*}
$$

The last inequality is simply due to

$$
\begin{aligned}
\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\left(\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{3} & =\mathbb{E}\left\langle Q\left(Q+\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)\left(Q-\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)\right\rangle_{t, \epsilon} \\
& \leq 2\|\varphi\|_{\infty}^{4} \sqrt{\mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{2}\right\rangle_{t, \epsilon}}
\end{aligned}
$$

We can bound the remaining integral on the right-hand side of (5.37) in a way similar to (5.27),

$$
\begin{aligned}
\int_{s_{n}}^{2 s_{n}} \sqrt{\mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{2}\right\rangle_{t, \epsilon}} \frac{d \epsilon}{s_{n}} & \leq \sqrt{\int_{s_{n}}^{2 s_{n}} \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{2}\right\rangle_{t, \epsilon} \frac{d \epsilon}{s_{n}}}, \\
& \leq \sqrt{\int_{s_{n}}^{2\left(s_{n}+\rho_{x}^{2}\right)} \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \frac{d R}{s_{n}}} .
\end{aligned}
$$

To obtain the second inequality, we make the change of variables $\epsilon \rightarrow R=R(t, \epsilon)$ (by Lemma 5.5, $R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $[0,+\infty)$ onto its image) and subsequently use $\partial R(t, \epsilon) / \partial \epsilon \geq 1, R\left(t,\left[s_{n}, 2 s_{n}\right]\right) \subseteq\left[s_{n}, 2\left(s_{n}+\rho_{x}^{2}\right)\right]$ for $n$ large enough (again by Lemma 5.5). For $n$ large enough, we apply Proposition 5.12 in Appendix 5.C with $M=2\left(1+\rho_{x}^{2}\right), a=s_{n}, b=2\left(s_{n}+\rho_{x}^{2}\right)$ and $\delta=s_{n} n^{\frac{2 \eta-1}{3}}$ in order to further bound the right-hand side of the last inequality and obtain

$$
\left|\frac{\lambda\|\varphi\|_{\infty}^{4}}{6} \int_{s_{n}}^{2 s_{n}} \sqrt{\mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)^{2}\right\rangle_{t, \epsilon}} \frac{d \epsilon}{s_{n}}\right| \leq C n^{\frac{5 \eta-1}{6}},
$$

where $C$ is a positive constant that does not depend on $t$ and $n$. Thus, the remaining term on the right-hand side of (5.37) vanishes when $n$ goes to infinity as long as $\eta<1 / 5$. Passing to the limit inferior on both sides of the inequality (5.37) yields

$$
\liminf _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \geq \inf _{q_{x} \in\left[0, \rho_{x}+\nu\right]} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right)
$$

This is true for all $\nu>0$ and Proposition 5.6 follows directly.

### 5.5 Derivation of the asymptotic Tensor-MMSE

The derivation of the asymptotic Tensor-MMSE rests on the following preliminary proposition.

Proposition 5.7. Suppose that (H1) and (H2) hold. Define for all $\lambda \in(0,+\infty)$ :

$$
\begin{aligned}
h(\lambda) & :=\inf _{q_{x} \in\left[0, \rho_{x}\right]} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right) ; \\
\mathcal{Q}_{x}^{*}(\lambda) & :=\left\{q_{x}^{*} \in\left[0, \rho_{x}\right]: \inf _{\left.q_{s} \in\left[0, \rho_{s}\right]\right]_{s} \geq 0} \sup _{r_{s}} \psi_{\lambda \alpha}\left(q_{x}^{*}, q_{s}, r_{s}\right)=h(\lambda)\right\} .
\end{aligned}
$$

For every $\lambda>0, \mathcal{Q}_{x}^{*}(\lambda)$ is nonempty. The function $h$ is differentiable at $\lambda$ if, and only if, the set $\mathcal{Q}_{x}^{*}(\lambda)$ is a singleton. In this case, letting $\mathcal{Q}_{x}^{*}(\lambda)=\left\{q_{x}^{*}(\lambda)\right\}$, the derivative of $h$ at $\lambda$ satisfies

$$
\begin{equation*}
h^{\prime}(\lambda)=\frac{1}{12}\left(\rho_{x}^{3}-\left(q_{x}^{*}(\lambda)\right)^{3}\right) \tag{5.38}
\end{equation*}
$$

We give the proof of this result in Appendix 5.E. We now prove Theorem 5.2 . Proof of Theorem 5.2. Let $n \in \mathbb{N}^{*}$. The angular brackets $\langle-\rangle_{n, \lambda}$ denote the expectation with respect to the posterior distribution of $\mathbf{S}$ given $(\mathbf{Y}, \mathbf{W})$. Define $h_{n}: \lambda \in(0,+\infty) \mapsto \frac{I(\mathbf{X}, \mathbf{Y} \mid \mathbf{W})}{n}$ (the mutual information depends on $\lambda$ through the observation $\mathbf{Y})$. We have for all $\lambda \in(0,+\infty)$ :

$$
\begin{aligned}
h_{n}(\lambda)= & \frac{\lambda}{2 n^{3}} \sum_{\underline{i} \in \mathcal{I}} \mathbb{E}\left[X_{i_{1}}^{2} X_{i_{2}}^{2} X_{i_{3}}^{2}\right] \\
& \left.-\frac{1}{n} \mathbb{E} \ln \int d P_{S}(\mathbf{s}) e^{\sum_{\underline{i} \in \mathcal{I}} x_{i_{1}} x_{i_{2}} x_{i_{3}}\left(-\frac{\lambda}{2 n^{2}} x_{i_{1}} x_{i_{2}} x_{i_{3}}+\frac{\lambda}{n^{2}} X_{i_{1}} X_{i_{2}} X_{i_{3}}+\frac{\sqrt{\lambda}}{n} Z_{\underline{i}}\right.}\right), \\
h_{n}^{\prime}(\lambda)= & \frac{1}{2 n^{3}} \sum_{\underline{i} \in \mathcal{I}} \mathbb{E}\left[\left(X_{i_{1}}^{2} X_{i_{2}}^{2} X_{i_{3}}-\left\langle x_{i_{1}} x_{i_{2}} x_{i_{3}}\right\rangle_{n, \lambda}\right)^{2}\right] \\
= & \frac{\operatorname{MMSE}_{n}\left(\mathbf{X}^{\otimes 3} \mid \mathbf{Y}, \mathbf{W}\right)}{12}+O\left(n^{-1}\right), \\
h_{n}^{\prime \prime}(\lambda)= & -\frac{1}{2 n^{5}} \sum_{\left(\underline{i}, \underline{i}^{\prime}\right) \in \mathcal{I}^{2}} \mathbb{E}\left[\left(\left\langle x_{i_{1}} x_{i_{2}} x_{i_{3}} x_{i_{1}^{\prime}}^{\prime} x_{i_{2}^{\prime}} x_{i_{3}^{\prime}}\right\rangle_{n, \lambda}-\left\langle x_{i_{1}} x_{i_{2}} x_{i_{3}}\right\rangle_{n, \lambda}\left\langle x_{i_{1}^{\prime}} x_{i_{2}^{\prime}} x_{i_{3}^{\prime}}\right\rangle_{n, \lambda}\right)^{2}\right] .
\end{aligned}
$$

The differentiations under the integral sign that yield the first and second derivatives of $h_{n}$ are justified by the domination properties implied by (H1), (H2). The second derivative $h_{n}^{\prime \prime}$ is nonpositive so $h_{n}$ is concave on $(0,+\infty)$. By Theorem 5.1, the sequence of continuously differentiable concave functions $\left(h_{n}\right)_{n \in \mathbb{N}^{*}}$ converges pointwise on $(0,+\infty)$ to

$$
h: \lambda \mapsto \inf _{q_{x} \in\left[0, \rho_{x}\right]} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right) .
$$

Hence, $h$ is concave and thus differentiable on $(0,+\infty)$ minus a countable set. By Griffiths' lemma [52, Appendix A], for every $\lambda$ where $h$ is differentiable, we have

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} h_{n}^{\prime}(\lambda)=h^{\prime}(\lambda)=\frac{\rho_{x}^{3}-\left(q_{x}^{*}(\lambda)\right)^{3}}{12} \tag{5.39}
\end{equation*}
$$

where the last equality is due to Proposition 5.7 and $q_{x}^{*}(\lambda)$ is the unique element of $\mathcal{Q}_{x}^{*}(\lambda)$. Combining the I-MMSE relationship $h_{n}^{\prime}(\lambda)=\frac{\operatorname{MMSE}_{n}\left(\mathbf{X}{ }^{\otimes 3} \mid \mathbf{Y}, \mathbf{W}\right)}{12}+O\left(n^{-1}\right)$ with (5.39) yields the theorem.

### 5.6 Limit of vanishing $\alpha$

In this section, we give a non-rigorous derivation of the limit of the asymptotic normalized mutual information when $\alpha$ goes to 0 .

Let $\lambda>0$ be fixed. We define the function

$$
\Psi_{*}: \alpha \mapsto \inf _{q_{x} \in\left[0, \rho_{x}\right]} \inf _{q_{s} \in\left[0, \rho_{s}\right]} \sup _{r_{s} \geq 0} \Psi\left(q_{x}, q_{s}, r_{s}, \alpha\right),
$$

where $\Psi\left(q_{x}, q_{s}, r_{s}, \alpha\right):=\alpha \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right)$ and $\psi_{\lambda, \alpha}$ is defined in (5.3). The function $\Psi_{*}$ is concave on $[0,+\infty)$, thus continuous on $[0,+\infty)$ and differentiable almost everywhere on $(0,+\infty)$. Note that

$$
\left.\frac{\partial \Psi}{\partial \alpha}\right|_{\left(q_{x}, q_{s}, r_{s}, \alpha\right)}=I_{\varphi}\left(\frac{\lambda q_{x}^{2}}{2}, q_{s} ; \rho_{s}\right)+\frac{\lambda}{12}\left(\rho_{x}-q_{x}\right)^{2}\left(\rho_{x}+2 q_{x}\right) .
$$

Assuming that we can apply an envelope theorem [68, Corollary 4] as in Appendix 5.E, it comes

$$
\Psi_{*}^{\prime}(\alpha)=I_{\varphi}\left(\frac{\lambda q_{x}^{*}(\alpha)^{2}}{2}, q_{s}^{*}(\alpha) ; \rho_{s}\right)+\frac{\lambda}{12}\left(\rho_{x}-q_{x}^{*}(\alpha)\right)^{2}\left(\rho_{x}+2 q_{x}^{*}(\alpha)\right),
$$

whenever $\left(q_{x}^{*}(\alpha), q_{s}^{*}(\alpha), r_{s}^{*}(\alpha)\right)$ is the unique triplet satisfying

$$
\Psi_{*}(\alpha)=\Psi\left(q_{x}^{*}(\alpha), q_{s}^{*}(\alpha), r_{s}^{*}(\alpha), \alpha\right) .
$$

At $\alpha=0, \Psi\left(q_{x}, q_{s}, r_{s}, \alpha\right)=I_{P_{S}}\left(r_{s}\right)-\frac{r_{s}\left(\rho_{s}-q_{s}\right)}{2}$ so $\Psi_{*}(0)=\Psi\left(q_{x}, q_{s}^{*}(0), r_{s}^{*}(0), \alpha\right)=0$, where $q_{s}^{*}(0)=m_{S}^{2}$ with $m_{S}:=\mathbb{E}_{S \sim P_{S}}[S]$ and $r_{s}^{*}(0)=0$. By Theorem 5.1.

$$
\lim _{\substack{n \rightarrow+\infty \\ n / p \rightarrow \alpha}} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n}=\frac{\Psi_{*}(\alpha)}{\alpha} .
$$

Using L'Hôpital's rule, it follows that

$$
\begin{equation*}
\lim _{\substack{\alpha \rightarrow 0^{+} \\
\lim _{\begin{subarray}{c}{n \rightarrow+\infty \\
n \rightarrow p} }}}\end{subarray}} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n}=\lim _{\alpha \rightarrow 0^{+}} \Psi_{*}^{\prime}(\alpha), \tag{5.40}
\end{equation*}
$$

provided that the limit on the right-hand side exists. Let us assume that $\lim _{\alpha \rightarrow 0^{+}}\left(q_{s}^{*}(\alpha), r_{s}^{*}(\alpha)\right)=\left(q_{s}^{*}(0), r_{s}^{*}(0)\right)=\left(m_{s}^{2}, 0\right)$. On one hand, we have

$$
\begin{aligned}
& \lim _{\alpha \rightarrow 0^{+}} \psi_{\lambda, \alpha}\left(q_{x}^{*}(\alpha), q_{s}^{*}(\alpha), r_{s}^{*}(\alpha)\right)=\lim _{\alpha \rightarrow 0^{+}} \psi_{\lambda, \alpha}\left(q_{x}^{*}(\alpha), m_{s}^{2}, 0\right) \\
& \quad=\lim _{\alpha \rightarrow 0+} I_{\varphi}\left(\frac{\lambda q_{x}^{*}(\alpha)^{2}}{2}, m_{s}^{2} ; \rho_{s}\right)+\frac{\lambda}{12}\left(\rho_{x}-q_{x}^{*}(\alpha)\right)^{2}\left(\rho_{x}+2 q_{x}^{*}(\alpha)\right)=\lim _{\alpha \rightarrow 0+} \Psi_{*}^{\prime}(\alpha) .
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
& \lim _{\alpha \rightarrow 0^{+}} \psi_{\lambda, \alpha}\left(q_{x}^{*}(\alpha), q_{s}^{*}(\alpha), r_{s}^{*}(\alpha)\right)=\lim _{\alpha \rightarrow 0^{+}} \inf _{q_{x} \in\left[0, \rho_{x}\right]} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}^{*}(\alpha), r_{s}^{*}(\alpha)\right) \\
& \quad=\inf _{q_{x} \in\left[0, \rho_{x}\right]} \psi_{\lambda, \alpha}\left(q_{x}, m_{S}^{2}, 0\right)=\inf _{q_{x} \in\left[0, \rho_{x}\right]} I_{\varphi}\left(\frac{\lambda q_{x}^{2}}{2}, m_{S}^{2} ; \rho_{s}\right)+\frac{\lambda}{12}\left(\rho_{x}-q_{x}\right)^{2}\left(\rho_{x}+2 q_{x}\right) .
\end{aligned}
$$

Combining both chains of equalities together with (5.40) gives

Thus, we conjecture that the asymptotic normalized multual information converges when $\alpha \rightarrow 0^{+}$to the asymptotic normalized mutual information associated with the channel

$$
\widetilde{Y}_{i j k}:=\frac{\sqrt{\lambda}}{n} \widetilde{X}_{i} \widetilde{X}_{j} \widetilde{X}_{k}+\widetilde{Z}_{i j k}, \quad 1 \leq i \leq j \leq k \leq n
$$

where $\widetilde{X}_{i}:=\varphi\left(\sqrt{\rho_{s}-m_{S}^{2}} U_{i}+\left|m_{S}\right| V_{i}\right)$ with $U_{1}, \ldots, U_{n}, V_{1}, \ldots, V_{n} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ and $\mathbf{V}$ is known. Proofs in the literature can be easily adapted to show that $\lim _{n \rightarrow+\infty} \frac{I(\widetilde{\mathbf{X}} ; \tilde{\mathbf{Y}} \mid \mathbf{V})}{n}$ is equal to the right-hand side of (5.41).

## Appendices

## 5.A Auxiliary lemmas

Lemma 5.8. Let $P_{S}$ be a probability distribution on $\mathbb{R}$ with finite second moment. Let $S \sim P_{S}$ and $Z \sim \mathcal{N}(0,1)$ be independent random variables. Define the functions

$$
\begin{aligned}
I_{P_{S}}:[0,+\infty) & \longrightarrow[0,+\infty) \\
r_{s} & \longmapsto I\left(S ; \sqrt{r_{s}} S+Z\right)
\end{aligned}
$$

and

$$
I_{P_{S}}^{*}: \begin{array}{lll}
\mathbb{R} & \longrightarrow[0,+\infty] \\
x & \longmapsto \sup _{r_{s} \geq 0} I_{P_{S}}\left(r_{s}\right)+x r_{s}
\end{array}
$$

Then, $I_{P_{S}}$ is twice-differentiable, nondecreasing, concave and $\frac{\mathbb{E} S^{2}}{2}$-Lipschitz continuous. Besides, $I_{P_{S}}^{*}$ is nondecreasing, convex, finite on $(-\infty, 0)$, equal to $+\infty$ on $(0,+\infty)$ and

$$
I_{P_{S}}^{*}(0)=\lim _{r_{s} \rightarrow+\infty} I_{P_{S}}\left(r_{s}\right) \in[0,+\infty]
$$

Proof. The properties of $I_{P_{S}}$ correspond to Lemma 2.3 that we state and prove in Chapter 2.

If $\operatorname{Var} S=0$ then $I_{P_{S}}$ is zero on its whole domain of definition so $I_{P_{S}}^{*}(x)=0$ if $x \leq 0,=+\infty$ otherwise. The properties of $I_{P_{S}}^{*}$ stated in the lemma directly follow. Suppose that $\operatorname{Var} S>0$. The function $I_{P_{S}}^{*}$ is the Legendre transform of the convex function $-I_{P_{S}}$, hence it is well-defined and convex. Besides, $I_{P_{S}}^{*}$ is defined as the supremum of nondecreasing affine functions of $x$ so it is nondecreasing. The trivial lower bound $I_{P_{S}}^{*}(x) \geq \sup _{r_{s} \geq 0} x r_{s}$ shows that $I_{P_{S}}^{*}$ is nonnegative and is equal to $+\infty$ on $(0,+\infty)$. For all $r_{s} \geq 0, I_{P_{S}}\left(r_{s}\right)=h\left(\sqrt{r_{s}} S+Z\right)-h(Z) \leq \ln \left(1+r_{s} \operatorname{Var} S\right) / 2$ where $h(\cdot)$ denotes the differential entropy. Hence, for all $x \in(-\infty, 0)$ :

$$
0 \leq I_{P_{S}}^{*}(x) \leq \sup _{r_{s} \geq 0} \frac{\ln \left(1+r_{s} \operatorname{Var} S\right)}{2}+x r_{s}=\frac{1}{2} \ln \left(\frac{\operatorname{Var} S}{2 e|x|}\right)+\frac{|x|}{\operatorname{Var} S}<+\infty
$$

Finally, $I_{P_{S}}^{*}(0):=\sup _{r_{s} \geq 0} I_{P_{S}}\left(r_{s}\right)=\lim _{r_{s} \rightarrow+\infty} I_{P_{S}}\left(r_{s}\right)$.
Lemma 5.9. Assume that $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ is a bounded continuous function. Let $U, V$ and $Z$ be independent standard Gaussian random variables. For every $(r, \rho) \in[0,+\infty)^{2}$ and $q \in[0, \rho]$, define

$$
I_{\varphi}(r, q ; \rho):=I(U ; \sqrt{r} \varphi(\sqrt{\rho-q} U+\sqrt{q} V)+Z \mid V) .
$$

Let $\rho \in[0,+\infty)$ be fixed. Then, $I_{\varphi}(\cdot, \cdot ; \rho)$ is continuous on $[0,+\infty) \times[0, \rho]$. Besides, for all $q \in[0, \rho], I_{\varphi}(\cdot, q ; \rho)$ is twice-differentiable, nondecreasing, concave and $\frac{\|\varphi\|_{\infty}^{2}}{2}$-Lipschitz continuous on $[0,+\infty)$.

Let $P_{S}$ be a probability distribution on $\mathbb{R}$ and $I_{P_{S}}$ the function defined in Lemma 5.8. For fixed $\alpha, \rho \geq 0$, define

$$
\begin{aligned}
\tilde{\psi}_{\alpha}: \begin{aligned}
{[0,+\infty)^{2} \times[0, \rho] } & \longrightarrow[0,+\infty) \\
\left(r, r_{s}, q\right) & \longmapsto I_{P_{S}}\left(r_{s}\right)+\alpha I_{\varphi}(r, q ; \rho)-\frac{r_{s}(\rho-q)}{2}
\end{aligned} . . . ~
\end{aligned}
$$

Then, the functions $r \mapsto \sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(r, r_{s}, q\right)$ and $r \mapsto \inf _{q \in[0, \rho]} \sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(r, r_{s}, q\right)$ are both nondecreasing, concave and $\frac{\alpha\|\varphi\|_{\infty}^{2}}{2}$-Lipschitz on $[0,+\infty)$.
Proof. Let $U, V, Z \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. Fix $\rho \geq 0$. For $(r, q) \in[0,+\infty) \times[0, \rho]$, define $Y^{(r, q)}:=\sqrt{r} \varphi(\sqrt{\rho-q} U+\sqrt{q} V)+Z$. Then,

$$
\begin{equation*}
I_{\varphi}(r, q ; \rho)=I\left(U ; Y^{(r, q)} \mid V\right)=-\mathbb{E} \ln \int \frac{d u}{\sqrt{2 \pi}} e^{-\frac{u^{2}}{2}-\mathcal{H}(r, q ; U, V, Z)} \tag{5.42}
\end{equation*}
$$

where

$$
\begin{aligned}
\mathcal{H}(r, q ; U, V, Z):=\frac{r}{2} & (\varphi(\sqrt{\rho-q} U+\sqrt{q} V)-\varphi(\sqrt{\rho-q} u+\sqrt{q} V))^{2} \\
& +\sqrt{r}(\varphi(\sqrt{\rho-q} U+\sqrt{q} V)-\varphi(\sqrt{\rho-q} u+\sqrt{q} V)) Z
\end{aligned}
$$

We denote by $\langle-\rangle_{r, q}$ the expectation with respect to the posterior distribution of $U$ given $\left(Y^{(r, q)}, V\right)$. The assumptions on $\varphi$ imply domination assumptions that justify the continuity of $I_{\varphi}(\cdot, \cdot ; \rho)$ and the twice differentiability of $r \mapsto I_{\varphi}(r, q ; \rho)$. Differentiating w.r.t. $r$ the right-hand side of (5.42) yields

$$
\begin{aligned}
\frac{\partial I_{\varphi}}{\partial r}(r, q ; \rho)=\frac{1}{2} \mathbb{E}\langle(\varphi(\sqrt{\rho-q} U & \left.+\sqrt{q} V)-\varphi(\sqrt{\rho-q} u+\sqrt{q} V))^{2}\right\rangle_{r, q} \\
& -\frac{1}{2 \sqrt{r}} \mathbb{E}\left[\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V)\rangle_{r, q} Z\right] \\
=\frac{1}{2} \mathbb{E}\left[\varphi^{2}(\sqrt{\rho-q} U\right. & \left.+\sqrt{q} V)-\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V))\rangle_{r, q}^{2}\right]
\end{aligned}
$$

where the last equality is obtained thanks to a Gaussian integration by parts w.r.t. $Z$ and the Nishimori identity

$$
\mathbb{E}\left[\varphi(\sqrt{\rho-q} U+\sqrt{q} V)\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V)\rangle_{r, q}\right]=\mathbb{E}\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V)\rangle_{r, q}^{2}
$$

By Jensen's inequality and the Nishimori identity, we have
$\mathbb{E}\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V)\rangle_{r, q}^{2} \leq \mathbb{E}\left\langle\varphi^{2}(\sqrt{\rho-q} u+\sqrt{q} V)\right\rangle_{r, q}=\mathbb{E} \varphi^{2}(\sqrt{\rho-q} U+\sqrt{q} V)$,
hence

$$
\left.\frac{\partial I_{\varphi}}{\partial r}(r, q ; \rho)=\frac{1}{2} \mathbb{E}\left[\varphi^{2}(\sqrt{\rho-q} U+\sqrt{q} V)-\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V))\right\rangle_{r, q}^{2}\right] \geq 0
$$

It follows from the nonnegativity of the first-order partial derivative that $I_{\varphi}(\cdot, q ; \rho)$ is nondecreasing. Further differentiating, and using integration by parts w.r.t. $Z$ and the Nishimori identity where necessary, we obtain

$$
\frac{\partial^{2} I_{\varphi}}{\partial r^{2}}(r, q ; \rho)=-\frac{1}{2} \mathbb{E}\left\langle\left(\varphi(\sqrt{\rho-q} u+\sqrt{q} V)-\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V)\rangle_{r, q}\right)^{2}\right\rangle_{r, q}^{2} .
$$

Thus, the first-order partial derivative is nonpositive and $I_{\varphi}(\cdot, q ; \rho)$ is concave. The Lipschitzness follows simply from

$$
0 \leq \frac{\partial I_{\varphi}}{\partial r}(r, q ; \rho) \leq \frac{1}{2} \mathbb{E}\left[\varphi^{2}(\sqrt{\rho-q} U+\sqrt{q} V)\right] \leq \frac{\|\varphi\|_{\infty}^{2}}{2} .
$$

The properties of $r \mapsto \sup _{r_{s} \geq 0} \tilde{\psi}_{\alpha}\left(r, r_{s}, q\right)$ follow directly from the ones of $I_{\varphi}(\cdot, q ; \rho)$ as

$$
\sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(r, r_{s}, q\right)=\alpha I_{\varphi}(r, q ; \rho)+\sup _{r_{s} \geq 0} I_{P_{S}}\left(r_{s}\right)-\frac{r_{s}(\rho-q)}{2} .
$$

Finally, $r \mapsto \inf _{q \in[0, \rho]} \sup _{r_{s} \geq 0} \widetilde{\psi}_{\alpha}\left(r, r_{s}, q\right)$ is the infimum of nondecreasing, concave, $\alpha\|\varphi\|_{\infty}^{2} / 2$-Lipschitz functions, hence its properties.

## 5.B Establishing the sum-rule

Lemma 5.10 (Link between average free entropy and normalized mutual information). Suppose that (H1) and (H2) hold, and that $R^{\prime}(t, \epsilon)$ is uniformly bounded in $(t, \epsilon) \in[0,1] \times[0,+\infty)$ where $R^{\prime}(\cdot, \epsilon)$ denotes the derivative of $R(\cdot, \epsilon)$. The normalized mutual information (5.9) and its partial derivative with respect to $t$, which we denote $i_{n}^{\prime}(t, \epsilon)$, satisfy

$$
\begin{align*}
& i_{n}(t, \epsilon)=-f_{n}(t, \epsilon)+\frac{\lambda R(t, \epsilon)}{4} \rho_{x}+\frac{\lambda(1-t)}{12} \rho_{x}^{3}+(1-t) o_{n}(1),  \tag{5.43}\\
& i_{n}^{\prime}(t, \epsilon)=-f_{n}^{\prime}(t, \epsilon)+\frac{\lambda R^{\prime}(t, \epsilon)}{4} \rho_{x}-\frac{\lambda}{12} \rho_{x}^{3}+o_{n}(1) . \tag{5.44}
\end{align*}
$$

The quantity $O_{n}(1)$ does not depend on $(t, \epsilon)$ and vanishes when $n$ goes to infinity. Besides, at $t=0$, for all $\epsilon \in[0,+\infty)$ :

$$
\begin{equation*}
\left|i_{n}(0, \epsilon)-\frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n}\right| \leq \frac{\lambda\|\varphi\|_{\infty}^{2}}{2} \epsilon . \tag{5.45}
\end{equation*}
$$

Proof. By definition of the normalized mutual information (5.9), we have

$$
\begin{align*}
i_{n}(t, \epsilon)= & \frac{1}{n} H\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)} \mid \mathbf{W}\right)-\frac{1}{n} H\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)} \mid \mathbf{S}, \mathbf{W}\right) \\
= & -\frac{1}{n} \mathbb{E} \ln \left(\mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right) e^{-\frac{\sum_{\underline{i} \in \mathcal{I}} Y_{i}^{2}+\|\tilde{\mathbf{Y}}\|^{2}}{2}}\right)+\frac{1}{n} \mathbb{E}\left[\ln e^{\left.-\frac{\sum_{\underline{i} \in \mathcal{I} \mathcal{I}_{i}^{2}+\|\tilde{\mathbf{Z}}\|^{2}}^{2}}{}\right]}=\right. \\
= & -f_{n}(t, \epsilon)+\frac{\lambda R(t, \epsilon)}{4 n} \sum_{j=1}^{n} \mathbb{E}\left[X_{j}^{2}\right]+\frac{\lambda(1-t)}{2 n^{3}} \sum_{\underline{i} \in \mathcal{I}} \mathbb{E}\left[X_{i_{1}}^{2} X_{i_{2}}^{2} X_{i_{3}}^{2}\right] \\
= & -f_{n}(t, \epsilon)+\frac{\lambda R(t, \epsilon)}{4} \mathbb{E}\left[X_{1}^{2}\right] \\
& +\frac{\lambda(1-t)}{2 n^{3}}\left(\binom{n}{3} \mathbb{E}\left[X_{1}^{2} X_{2}^{2} X_{3}^{2}\right]+n(n-1) \mathbb{E}\left[X_{1}^{2} X_{2}^{4}\right]+n \mathbb{E}\left[X_{1}^{6}\right]\right) \\
= & -f_{n}(t, \epsilon)+\frac{\lambda R(t, \epsilon)}{4} \mathbb{E}\left[X_{1}^{2}\right]+\frac{\lambda(1-t)}{12} \mathbb{E}\left[X_{1}^{2} X_{2}^{2} X_{3}^{2}\right]+\lambda(1-t) O\left(n^{-1}\right) . \tag{5.46}
\end{align*}
$$

The quantity $O\left(n^{-1}\right)$ appearing in the last equality does not depend on $(t, \epsilon, \lambda)$ and is such that $\left|O\left(n^{-1}\right)\right| \leq C / n$ with $C:=\|\varphi\|_{\infty}^{6} / 2$. It directly follows that

$$
\begin{equation*}
i_{n}^{\prime}(t, \epsilon)=-f_{n}^{\prime}(t, \epsilon)+\frac{\lambda R^{\prime}(t, \epsilon)}{4} \mathbb{E}\left[X_{1}^{2}\right]-\frac{\lambda}{12} \mathbb{E}\left[X_{1}^{2} X_{2}^{2} X_{3}^{2}\right]-\lambda O\left(n^{-1}\right) \tag{5.47}
\end{equation*}
$$

where the quantity $O\left(n^{-1}\right)$ on the right-hand side of (5.47) is the same as the one appearing on the right-hand side of (5.46). Note that $\mathbb{E}\left[X_{1}^{2} X_{2}^{2} X_{3}^{2}\right]=\mathbb{E}\left[\mathbb{E}\left[X_{1}^{2} \mid \mathbf{S}\right]^{3}\right]$ converges to $\rho_{x}^{3}$ as $n$ goes to infinity (the proof of this limit is similar to 775 , Lemma 3 of Supplementary material]). This limit together with (5.46) and (5.47) yields (5.43) and (5.44).

At $t=0$, we can use (5.46) to obtain (remember that $R(0, \epsilon)=\epsilon$ )

$$
\begin{equation*}
\left|i_{n}(0, \epsilon)-i_{n}(0,0)\right| \leq\left|f_{n}(0, \epsilon)-f_{n}(0,0)\right|+\frac{\lambda \epsilon}{4} \mathbb{E}\left[X_{1}^{2}\right] \tag{5.48}
\end{equation*}
$$

It is clear that $i_{n}(0,0)=I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W}) / n$ where $\mathbf{Y}, \mathbf{X}$ are defined in (5.1), (5.2). At $t=0$, the average free entropy (5.14) reads

$$
\begin{equation*}
f_{n}(0, \epsilon):=\frac{1}{n} \mathbb{E} \ln \int d P_{S}(\mathbf{s}) e^{-\mathcal{H}_{0, \epsilon}(\mathbf{s} ; \mathbf{Z}, \widetilde{\mathbf{z}}, \mathbf{X}, \mathbf{W})} \tag{5.49}
\end{equation*}
$$

where (remember that $x_{1}, \ldots, x_{n}$ are the entries of $\left.\mathbf{x}:=\varphi(\mathbf{W s} / \sqrt{p})\right)$

$$
\begin{gather*}
\mathcal{H}_{0, \epsilon}(\mathbf{s} ; \mathbf{Z}, \widetilde{\mathbf{Z}}, \mathbf{X}, \mathbf{W}):=\sum_{\underline{i} \in \mathcal{I}} \frac{\lambda}{2 n^{2}} x_{i_{1}}^{2} x_{i_{2}}^{2} x_{i_{3}}^{2}-\frac{\lambda}{n^{2}} X_{i_{1}} X_{i_{2}} X_{i_{3}} x_{i_{1}} x_{i_{2}} x_{i_{3}}-\frac{\sqrt{\lambda}}{n} Z_{\underline{i}} x_{i_{1}} x_{i_{2}} x_{i_{3}} \\
 \tag{5.50}\\
+\sum_{j=1}^{n} \frac{\lambda \epsilon}{4} x_{j}^{2}-\frac{\lambda \epsilon}{2} X_{j} x_{j}-\sqrt{\frac{\lambda \epsilon}{2}} \widetilde{Z}_{j} x_{j}
\end{gather*}
$$

Differentiating (5.49) under the integral sign yields

$$
\left.\frac{\partial f_{n}}{\partial \epsilon}\right|_{0, \epsilon}=-\mathbb{E}\left\langle\frac{\partial \mathcal{H}_{0, \epsilon}}{\partial \epsilon}\right\rangle_{0, \epsilon}=-\mathbb{E}\langle\mathcal{L}\rangle_{0, \epsilon}
$$

where

$$
\mathcal{L}:=\frac{1}{n} \sum_{j=1}^{n} \frac{\lambda}{4} x_{j}^{2}-\frac{\lambda}{2} X_{j} x_{j}-\frac{1}{2} \sqrt{\frac{\lambda}{2 \epsilon}} \widetilde{Z}_{j} x_{j} .
$$

By a Gaussian integration by parts we show in Lemma 5.13 of Appendix 5.C that

$$
\mathbb{E}\langle\mathcal{L}\rangle_{0, \epsilon}=-\frac{\lambda}{4} \mathbb{E}\langle Q\rangle_{0, \epsilon},
$$

where $Q:=\mathbf{x}^{\top} \mathbf{x} / n$ is the overlap. Therefore, $\left|\partial f_{n} / \partial \epsilon\right|_{0, \epsilon} \mid \leq \lambda\|\varphi\|_{\infty}^{2} / 4$. By the mean value theorem, it follows that $\left|f_{n}(0, \epsilon)-f_{n}(0,0)\right| \leq \lambda\|\varphi\|_{\infty}^{2} \epsilon / 4$. Making use of this upper bound in (5.48) yields (5.45).

Lemma 5.11 (Derivative of the normalized mutual information). Suppose that (H1) and (H2) hold, and that $R^{\prime}(t, \epsilon)$ is uniformly bounded in $(t, \epsilon) \in[0,1] \times[0,+\infty)$ where $R^{\prime}(\cdot, \epsilon)$ denotes the derivative of $R(\cdot, \epsilon)$. Define the overlap

$$
Q:=\frac{1}{n} \sum_{i=1}^{n} \varphi\left(\left[\frac{\mathbf{W} \mathbf{s}}{\sqrt{p}}\right]_{i}\right) \varphi\left(\left[\frac{\mathbf{W S}}{\sqrt{p}}\right]_{i}\right)=\frac{1}{n} \sum_{i=1}^{n} x_{i} X_{i} .
$$

We denote by $i_{n}^{\prime}(\cdot, \epsilon)$ the derivative of the normalized mutual information $i_{n}(\cdot, \epsilon)$ defined in (5.9), and $\forall(t, \epsilon) \in[0,1] \times(0,+\infty)$ :

$$
\begin{equation*}
i_{n}^{\prime}(t, \epsilon)=\frac{\lambda}{12}\left(\mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}-\rho_{x}^{3}\right)+\frac{\lambda R^{\prime}(t, \epsilon)}{4}\left(\rho_{x}-\mathbb{E}\langle Q\rangle_{t, \epsilon}\right)+o_{n}(1), \tag{5.51}
\end{equation*}
$$

where $o_{n}(1)$ vanishes uniformly in $(t, \epsilon)$ as $n$ goes to infinity.
Proof. The average interpolating free entropy satisfies

$$
f_{n}(t, \epsilon)=\frac{1}{n} \mathbb{E}_{\mathbf{S}, \mathbf{W}}\left[\int d \mathbf{y} d \widetilde{\mathbf{y}} \frac{e^{-\frac{1}{2}\left(\sum_{\underline{i} \epsilon \mathcal{I}} y_{\underline{2}}^{2}+\|\tilde{\mathbf{y}}\|^{2}\right)}}{\sqrt{2 \pi}^{n+|\mathcal{I}|}} e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)} \ln \mathcal{Z}_{t, \epsilon}(\mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W})\right] .
$$

Differentiating the right-hand side of this identity under the integral sign yields

$$
\begin{gather*}
f_{n}^{\prime}(t, \epsilon)=-\frac{1}{n} \mathbb{E}\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{S} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right) \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)\right] \\
-\frac{1}{n} \mathbb{E}\left[\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)\right\rangle_{t, \epsilon}\right] \tag{5.52}
\end{gather*}
$$

where

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}^{\prime}(\mathbf{s} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}):=\sum_{\underline{i} \in \mathcal{I}} & -\frac{\lambda}{2 n^{2}} x_{i_{1}}^{2} x_{i_{2}}^{2} x_{i_{3}}^{2}+\frac{1}{2 n} \sqrt{\frac{\lambda}{1-t}} y_{\underline{i}} x_{i_{1}} x_{i_{2}} x_{i_{3}} \\
& +\sum_{j=1}^{n} \frac{\lambda R^{\prime}(t, \epsilon)}{4} x_{i}^{2}-\frac{R^{\prime}(t, \epsilon)}{2} \sqrt{\frac{\lambda}{2 R(t, \epsilon)}} \widetilde{y}_{j} x_{j} . \tag{5.53}
\end{align*}
$$

Equation (5.53) comes from differentiating with respect to $t$ the interpolating Hamiltonian (5.12). Evaluating (5.53) at $(\mathbf{s}, \mathbf{y}, \widetilde{\mathbf{y}})=\left(\mathbf{S}, \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)$ gives

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{S} ; \mathbf{Y}^{(t)},\right. & \left.\widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right) \\
& =\sum_{\underline{i} \in \mathcal{I}} \frac{1}{2 n} \sqrt{\frac{\lambda}{1-t}} Z_{\underline{i}} X_{i_{1}} X_{i_{2}} X_{i_{3}}-\sum_{j=1}^{n} \frac{R^{\prime}(t, \epsilon)}{2} \sqrt{\frac{\lambda}{2 R(t, \epsilon)}} \widetilde{Z}_{j} X_{j} . \tag{5.54}
\end{align*}
$$

The second expectation on the right-hand side of 5.52 is easily shown to be zero thanks to the Nishimori identity,

$$
\mathbb{E}\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{s} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)\right\rangle_{t, \epsilon}=\mathbb{E} \mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{S} ; \mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)=0
$$

Therefore, the identity (5.52) simplifies to

$$
\begin{align*}
f_{n}^{\prime}(t, \epsilon)=-\frac{1}{2 n^{2}} & \sqrt{\frac{\lambda}{1-t}} \sum_{\underline{i} \in \mathcal{I}} \mathbb{E}\left[Z_{\underline{i}} X_{i_{1}} X_{i_{2}} X_{i_{3}} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)\right] \\
& +\frac{R^{\prime}(t, \epsilon)}{2 n} \sqrt{\frac{\lambda}{2 R(t, \epsilon)}} \sum_{j=1}^{n} \mathbb{E}\left[\widetilde{Z}_{j} X_{j} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)\right] \tag{5.55}
\end{align*}
$$

The two expectations appearing on the right-hand side of (5.55) are now simplified by a Gaussian integration by parts w.r.t. $Z_{\underline{i}}$ and $\widetilde{Z}_{j}$ (see Lemma 2.2.,

$$
\begin{aligned}
\mathbb{E}\left[Z_{\underline{i}} X_{i_{1}} X_{i_{2}} X_{i_{3}} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)\right] & =\frac{\sqrt{\lambda(1-t)}}{n} \mathbb{E}\left\langle x_{i_{1}} X_{i_{1}} x_{i_{2}} X_{i_{2}} x_{i_{3}} X_{i_{3}}\right\rangle_{t, \epsilon} \\
\mathbb{E}\left[\widetilde{Z}_{j} X_{j} \ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}\right)\right] & =\sqrt{\frac{\lambda R(t, \epsilon)}{2}} \mathbb{E}\left\langle x_{j} X_{j}\right\rangle_{t, \epsilon}
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
f_{n}^{\prime}(t, \epsilon) & =-\frac{\lambda}{2 n^{3}} \sum_{\underline{i} \in \mathcal{I}} \mathbb{E}\left\langle x_{i_{1}} X_{i_{1}} x_{i_{2}} X_{i_{2}} x_{i_{3}} X_{i_{3}}\right\rangle_{t, \epsilon}+\frac{\lambda R^{\prime}(t, \epsilon)}{4 n} \sum_{j=1}^{n} \mathbb{E}\left\langle x_{j} X_{j}\right\rangle_{t, \epsilon} \\
& =-\frac{\lambda}{12} \mathbb{E}\left\langle Q^{3}\right\rangle_{t, \epsilon}+\frac{\lambda R^{\prime}(t, \epsilon)}{4} \mathbb{E}\langle Q\rangle_{t, \epsilon}+\frac{\lambda}{2} O\left(n^{-1}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
\left|O\left(n^{-1}\right)\right| & =\frac{1}{n^{3}}\left|\sum_{\underline{i} \in \mathcal{I}} \mathbb{E}\left\langle x_{i_{1}} X_{i_{1}} x_{i_{2}} X_{i_{2}} x_{i_{3}} X_{i_{3}}\right\rangle_{t, \epsilon}-\frac{1}{6} \sum_{i_{1}, i_{2}, i_{3}=1}^{n} \mathbb{E}\left\langle x_{i_{1}} X_{i_{1}} x_{i_{2}} X_{i_{2}} x_{i_{3}} X_{i_{3}}\right\rangle_{t, \epsilon}\right| \\
& \leq \frac{\|\varphi\|_{\infty}^{6}}{n} .
\end{aligned}
$$

## 5.C Concentration of the overlap

One important result in order to prove Propositions 5.4 and 5.6 is the concentration of the overlap $Q:=\mathbf{x}^{\top} \mathbf{x} / n$ around its expectation $\mathbb{E}\langle Q\rangle_{t, R}$ as long as we integrate over $R$ in a bounded subset of $(0,+\infty)$. Remember that the angular brackets $\langle-\rangle_{t, R}$ denote the expectation with respect to the posterior distribution (5.29).

Proposition 5.12 (Concentration of the overlap around its expectation). Suppose that (H1) and (H2) hold. Let $M$ be a positive real number. For $n$ large enough, there exists a constant $C$ that depends only on $\|\varphi\|_{\infty},\left\|\varphi^{\prime}\right\|_{\infty},\left\|\varphi^{\prime \prime}\right\|_{\infty}, M_{S}, \lambda, M$, and such that $\forall b \in(0, M), \forall a \in(0, \min \{1, b\}), \forall \delta \in(0, a), \forall t \in[0,1]:$

$$
\int_{a}^{b} \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R \leq C\left(\frac{1}{\delta^{2} n}-\frac{\ln (a)}{n}+\frac{\delta}{a-\delta}\right)
$$

The concentration of the scalar overlap around its expectation follows from the concentration of the quantity

$$
\begin{equation*}
\mathcal{L}:=\frac{1}{n} \sum_{j=1}^{n} \frac{\lambda}{4} x_{j}^{2}-\frac{\lambda}{2} x_{j} X_{j}-\frac{1}{2} \sqrt{\frac{\lambda}{2 R}} x_{j} \widetilde{Z}_{j} \tag{5.56}
\end{equation*}
$$

Lemma 5.13 (Link between the fluctuations of $\mathcal{L}$ and $Q$ ). Assume that the function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ is continuous and bounded. For all $(t, R) \in[0,1] \times(0,+\infty)$ :

$$
\begin{align*}
\mathbb{E}\langle\mathcal{L}\rangle_{t, R} & =-\frac{\lambda}{4} \mathbb{E}\langle Q\rangle_{t, R}  \tag{5.57}\\
\frac{\lambda}{4} \mathbb{E}\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R} & \leq \frac{\|\varphi\|_{\infty}^{2}}{\sqrt{2}} \sqrt{\mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\frac{1}{n} \mathbb{E}\left\langle\frac{\partial \mathcal{L}}{\partial R}\right\rangle_{t, R}}  \tag{5.58}\\
\frac{\lambda^{2}}{16} \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R} & \leq \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} . \tag{5.59}
\end{align*}
$$

Proof. Fix $(t, R) \in[0,1] \times(0,+\infty)$. By the definition (5.56) of $\mathcal{L}$, we have

$$
\begin{align*}
\mathbb{E}\langle\mathcal{L}\rangle_{t, R} & =\frac{1}{n} \sum_{j=1}^{n} \frac{\lambda}{4} \mathbb{E}\left\langle x_{j}^{2}\right\rangle_{t, R}-\frac{\lambda}{2} \mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R} X_{j}\right]-\frac{1}{2} \sqrt{\frac{\lambda}{2 R}} \mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R} \widetilde{Z}_{j}\right]  \tag{5.60}\\
\mathbb{E}\langle Q \mathcal{L}\rangle_{t, R} & =\frac{1}{n} \sum_{j=1}^{n} \frac{\lambda}{4} \mathbb{E}\left\langle Q x_{j}^{2}\right\rangle_{t, R}-\frac{\lambda}{2} \mathbb{E}\left[\left\langle Q x_{j}\right\rangle_{t, R} X_{j}\right]-\frac{1}{2} \sqrt{\frac{\lambda}{2 R}} \mathbb{E}\left[\left\langle Q x_{j}\right\rangle_{t, R} \widetilde{Z}_{j}\right] . \tag{5.61}
\end{align*}
$$

After integrating by parts with respect to the standard Gaussian random variable $\widetilde{Z}_{j}$, the last expectation on the right-hand side of each of (5.60) and (5.61) reads

$$
\begin{align*}
\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R} \widetilde{Z}_{j}\right] & =\sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\left\langle x_{j}^{2}\right\rangle_{t, R}\right]-\sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{2}\right] ;  \tag{5.62}\\
\mathbb{E}\left[\left\langle Q x_{j}\right\rangle_{t, R} \widetilde{Z}_{j}\right] & =\sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\left\langle Q x_{j}^{2}\right\rangle_{t, R}\right]-\sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\left\langle Q x_{j}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right] . \tag{5.63}
\end{align*}
$$

Plugging (5.62) in (5.60) yields

$$
\begin{aligned}
\mathbb{E}\langle\mathcal{L}\rangle_{t, R} & =\frac{\lambda}{2 n} \sum_{j=1}^{n} \frac{1}{2} \mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{2}\right]-\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R} X_{j}\right] \\
& =-\frac{\lambda}{4 n} \sum_{j=1}^{n} \frac{\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R} X_{j}\right]}{n}=-\frac{\lambda}{4} \mathbb{E}\langle Q\rangle_{t, R},
\end{aligned}
$$

where the second equality is due to the Nishimori identity $\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R}^{2}\right]=\mathbb{E}\left[\left\langle x_{j}\right\rangle_{t, R} X_{j}\right]$. This ends the proof of (5.57). Plugging (5.63) in (5.61), it comes

$$
\begin{align*}
\mathbb{E}\langle Q \mathcal{L}\rangle_{t, R} & =\frac{\lambda}{2 n} \sum_{j=1}^{n_{v}} \frac{1}{2} \mathbb{E}\left[\left\langle Q x_{j}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right]-\mathbb{E}\left[\left\langle Q x_{j}\right\rangle_{t, R} X_{j}\right] \\
& =\frac{\lambda}{2 n} \sum_{j=1}^{n_{v}} \frac{1}{2} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j} X_{j}\right\rangle_{t, R}\right]-\mathbb{E}\left[\left\langle Q x_{j}\right\rangle_{t, R} X_{j}\right] \\
& =\frac{\lambda}{2}\left(\frac{1}{2} \mathbb{E}\left[\langle Q\rangle_{t, R}^{2}\right]-\mathbb{E}\left[\left\langle Q^{2}\right\rangle_{t, R}\right]\right), \tag{5.64}
\end{align*}
$$

where the second equality again follows again from the Nishimori identity. We now combine (5.64) and (5.57) to obtain

$$
\begin{aligned}
& \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, R}\right)\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)\right\rangle_{t, R}=\mathbb{E}\langle Q \mathcal{L}\rangle_{t, R}-\mathbb{E}\langle Q\rangle_{t, R} \mathbb{E}\langle\mathcal{L}\rangle_{t, R} \\
&=\frac{\lambda}{4}\left(\mathbb{E}\left[\langle Q\rangle_{t, R}^{2}\right]-2 \mathbb{E}\left[\left\langle Q^{2}\right\rangle_{t, R}\right]+\left(\mathbb{E}\langle Q\rangle_{t, R}\right)^{2}\right) \\
&=-\frac{\lambda}{4}\left(\mathbb{E}\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R}+\mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R}\right)
\end{aligned}
$$

From this last identity, it follows that

$$
\begin{aligned}
\frac{\lambda}{4} \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R} & \leq-\mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, R}\right)\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)\right\rangle_{t, R} \\
& \leq \sqrt{\mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R} \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}}
\end{aligned}
$$

where the second inequality is due to Cauchy-Schwarz inequality. This upperbound directly implies (5.59).

The proof of the inequality (5.58) is more involved. We have the two useful identities (just replace $Q$ by its definition)

$$
\begin{align*}
\mathbb{E}\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R}= & \frac{1}{n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[X_{i} X_{j}\left(\left\langle x_{i} x_{j}\right\rangle_{t, R}-\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right)\right]  \tag{5.65}\\
\mathbb{E}\left[\left(\langle Q\rangle_{t, R}-\left\|\frac{\langle\mathbf{x}\rangle_{t, R}}{\sqrt{n}}\right\|^{2}\right)^{2}\right]= & \frac{1}{n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[X_{i} X_{j}\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right] \\
& -2 \mathbb{E}\left[X_{i}\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}^{2}\right]+\mathbb{E}\left[\left\langle x_{i}\right\rangle_{t, R}^{2}\left\langle x_{j}\right\rangle_{t, R}^{2}\right] \\
= & \frac{1}{n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[X_{i} X_{j}\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right]-\mathbb{E}\left[\left\langle x_{i}\right\rangle_{t, R}^{2}\left\langle x_{j}\right\rangle_{t, R}^{2}\right] . \tag{5.66}
\end{align*}
$$

Differentiating with respect to $R$ on both sides of (5.57), and then dividing by $-n$, yields

$$
\begin{equation*}
\mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\frac{1}{n} \mathbb{E}\left\langle\frac{\partial \mathcal{L}}{\partial R}\right\rangle_{t, R}=-\frac{\lambda}{4}\left(\mathbb{E}\langle Q \mathcal{L}\rangle_{t, R}-\mathbb{E}\langle Q\rangle_{t, R}\langle\mathcal{L}\rangle_{t, R}\right) . \tag{5.67}
\end{equation*}
$$

Let us simplify the right-hand side of (5.67). By definition, we have

$$
\begin{align*}
\mathbb{E}\langle Q\rangle_{t, R}\langle\mathcal{L}\rangle_{t, R}=\frac{1}{n} \sum_{j=1}^{n} \frac{\lambda}{4} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j}^{2}\right\rangle_{t, R}\right] & -\frac{\lambda}{2} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R} X_{j}\right] \\
& -\frac{1}{2} \sqrt{\frac{\lambda}{2 R}} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R} \widetilde{Z}_{j}\right] \tag{5.68}
\end{align*}
$$

After a Gaussian integration by parts with respect to $\widetilde{Z}_{j}$, the third expectation in the summand of (5.68) reads

$$
\begin{aligned}
\mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R} \widetilde{Z}_{j}\right]=\sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\left\langle Q x_{j}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right] & +\sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j}^{2}\right\rangle_{t, R}\right] \\
& -2 \sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}^{2}\right] \\
=\sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j} X_{j}\right\rangle_{t, R}\right] & +\sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j}^{2}\right\rangle_{t, R}\right] \\
& -2 \sqrt{\frac{\lambda R}{2}} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}^{2}\right]
\end{aligned}
$$

Plugging this result back in (5.68) gives

$$
\begin{align*}
\mathbb{E}\langle Q\rangle_{t, R}\langle\mathcal{L}\rangle_{t, R} & =\frac{\lambda}{2 n} \sum_{j=1}^{n} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}^{2}\right]-\frac{3}{2} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\langle x_{j} X_{j}\right\rangle_{t, R}\right] \\
& =\frac{\lambda}{2} \mathbb{E}\left[\langle Q\rangle_{t, R}\left\|\frac{\langle\mathbf{x}\rangle_{t, R}}{\sqrt{n}}\right\|^{2}\right]-\frac{3 \lambda}{4} \mathbb{E}\left[\langle Q\rangle_{t, R}^{2}\right] \tag{5.69}
\end{align*}
$$

Finally, combining (5.64) and (5.69) yields the following expression for the righthand side of (5.67),

$$
\begin{align*}
-\frac{\lambda}{4}(\mathbb{E} & \left.\langle Q \mathcal{L}\rangle_{t, R}-\mathbb{E}\langle Q\rangle_{t, R}\langle\mathcal{L}\rangle_{t, R}\right) \\
& =\frac{\lambda^{2}}{8}\left(\mathbb{E}\left[\left\langle Q^{2}\right\rangle_{t, R}\right]-\mathbb{E}\left[\langle Q\rangle_{t, R}^{2}\right]+\mathbb{E}\left[\langle Q\rangle_{t, R}\left\|\frac{\langle\mathbf{x}\rangle_{t, R}}{\sqrt{n}}\right\|^{2}\right]-\mathbb{E}\left[\langle Q\rangle_{t, R}^{2}\right]\right) \\
& =\frac{\lambda^{2}}{8}\left(\mathbb{E}\left[\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R}\right]-\mathbb{E}\left[\left(\langle Q\rangle_{t, R}-\left\|\frac{\langle\mathbf{x}\rangle_{t, R}}{\sqrt{n}}\right\|^{2}\right)^{2}\right]\right) \\
& =\frac{\lambda^{2}}{8 n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[X_{i} X_{j}\left\langle x_{i} x_{j}\right\rangle_{t, R}\right]-2 \mathbb{E}\left[X_{i} X_{j}\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right]+\mathbb{E}\left[\left\langle x_{i}\right\rangle_{t, R}^{2}\left\langle x_{j}\right\rangle_{t, R}^{2}\right] \\
& =\frac{\lambda^{2}}{8 n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[\left(\left\langle x_{i} x_{j}\right\rangle_{t, R}-\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right)^{2}\right] . \tag{5.70}
\end{align*}
$$

The second-to-last equality follows from (5.65) and (5.66) while the factorization of the last equality appears after applying the Nishimori identity
$\mathbb{E}\left[X_{i} X_{j}\left\langle x_{i} x_{j}\right\rangle_{t, R}\right]=\mathbb{E}\left\langle x_{i} x_{j}\right\rangle_{t, R}^{2}, \mathbb{E}\left[X_{i} X_{j}\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right]=\mathbb{E}\left[\left\langle x_{i} x_{j}\right\rangle_{t, R}\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right]$.
We now come back to the identity (5.65) and apply Jensen's inequality to its right-hand side. We get

$$
\begin{aligned}
\frac{\lambda}{4} \mathbb{E}\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R} & \leq \frac{\lambda\|\varphi\|_{\infty}^{2}}{4 n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[\left|\left\langle x_{i} x_{j}\right\rangle_{t, R}-\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right|\right] \\
& \leq \frac{\|\varphi\|_{\infty}^{2}}{\sqrt{2}} \sqrt{\frac{\lambda^{2}}{8 n^{2}} \sum_{i, j=1}^{n} \mathbb{E}\left[\left(\left\langle x_{i} x_{j}\right\rangle_{t, R}-\left\langle x_{i}\right\rangle_{t, R}\left\langle x_{j}\right\rangle_{t, R}\right)^{2}\right]} \\
& =\frac{\|\varphi\|_{\infty}^{2}}{\sqrt{2}} \sqrt{-\frac{\lambda}{4}\left(\mathbb{E}\langle Q \mathcal{L}\rangle_{t, R}-\mathbb{E}\langle Q\rangle_{t, R}\langle\mathcal{L}\rangle_{t, R}\right)},
\end{aligned}
$$

where the equality follows from (5.70). We finally obtain (5.58) by combining the latter upper bound with (5.67).

## 5.C. 1 Concentration of $\mathcal{L}$ around its expectation

To prove concentration results on $\mathcal{L}$, it is useful to work with the free entropy $\ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right) / n$ where $\mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)$ is the normalization factor of the posterior distribution (5.29). In Appendix 5.D, we prove that this free entropy concentrates around its expectation when $n \rightarrow+\infty$. We define

$$
\begin{aligned}
F_{n}(t, R) & :=\frac{1}{n} \ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right) \\
f_{n}(t, R) & :=\frac{1}{n} \mathbb{E}\left[\ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)\right]=\mathbb{E} F_{n}(t, R)
\end{aligned}
$$

Proposition 5.14 (Thermal fluctuations of $\mathcal{L}$ and $Q$ ). Assume that $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ is continuous and bounded. For all positive real numbers $a<b$ and $t \in[0,1]$ :

$$
\begin{aligned}
\int_{a}^{b} \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R & \leq \frac{\lambda\|\varphi\|_{\infty}^{2}}{4 n}\left(\frac{\ln (b / a)}{2}+1\right) \\
\frac{\lambda}{4} \int_{a}^{b} \mathbb{E}\left\langle\left(Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, R}}{\sqrt{n}}\right\|^{2}\right)^{2}\right\rangle_{t, R} d R & \leq\|\varphi\|_{\infty}^{3} \sqrt{\frac{\lambda(b-a)}{2 n}}
\end{aligned}
$$

Proof. Fix $(n, t) \in \mathbb{N}^{*} \times[0,1]$. Note that $\forall R \in(0,+\infty)$ :

$$
\begin{equation*}
\left.\frac{\partial f_{n}}{\partial R}\right|_{t, R}=-\frac{1}{n} \mathbb{E}\left[\left\langle\frac{\partial \mathcal{H}_{t, R}\left(\mathbf{x} ; \mathbf{Y}^{(t)}, \tilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)}{\partial R}\right\rangle_{t, R}\right]=-\mathbb{E}\langle\mathcal{L}\rangle_{t, R} \tag{5.71}
\end{equation*}
$$

Further differentiating, we obtain

$$
\begin{align*}
\left.\frac{\partial^{2} f_{n}}{\partial R^{2}}\right|_{t, R} & =\mathbb{E}\left[\left\langle\mathcal{L} \frac{\partial \mathcal{H}_{t, R}}{\partial R}\right\rangle_{t, R}\right]-\mathbb{E}\left[\langle\mathcal{L}\rangle_{t, R}\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial R}\right\rangle_{t, R}\right]-\mathbb{E}\left\langle\frac{\partial \mathcal{L}}{\partial R}\right\rangle_{t, R} \\
& =n \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\mathbb{E}\left\langle\frac{\partial \mathcal{L}}{\partial R}\right\rangle_{t, R} \tag{5.72}
\end{align*}
$$

It follows directly from (5.72) and the definition (5.56) of $\mathcal{L}$ that

$$
\begin{equation*}
\mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}=\left.\frac{1}{n} \frac{\partial^{2} f_{n}}{\partial R^{2}}\right|_{t, R}+\frac{1}{4 R} \sqrt{\frac{\lambda}{2 R}} \frac{\mathbb{E}\left[\langle\mathbf{x}\rangle_{t, R}^{\top} \widetilde{\mathbf{Z}}\right]}{n^{2}} \tag{5.73}
\end{equation*}
$$

We start by upper bounding the integral over the second summand on the righthand side of (5.73). Thanks to a Gaussian integration by parts with respect to $\widetilde{Z}_{j}, j \in\{1, \ldots, n\}$, it comes:

$$
\begin{equation*}
\frac{1}{4 R} \sqrt{\frac{\lambda}{2 R}} \frac{\mathbb{E}\left[\langle\mathbf{x}\rangle_{t, R}^{\top} \widetilde{\mathbf{Z}}\right]}{n^{2}}=\frac{\lambda}{8 R} \frac{\mathbb{E}\left[\left\langle\|\mathbf{x}\|^{2}\right\rangle_{t, R}-\left\|\langle\mathbf{x}\rangle_{t, R}\right\|^{2}\right]}{n^{2}} \leq \frac{\lambda\|\varphi\|_{\infty}^{2}}{8 R n} . \tag{5.74}
\end{equation*}
$$

Therefore,

$$
\begin{equation*}
\int_{a}^{b} \frac{d R}{4 R} \sqrt{\frac{\lambda}{2 R}} \frac{\mathbb{E}\left[\langle\mathbf{x}\rangle_{t, R}^{\top} \widetilde{\mathbf{Z}}\right]}{n^{2}} \leq \frac{\lambda\|\varphi\|_{\infty}^{2}}{8 n} \ln (b / a) . \tag{5.75}
\end{equation*}
$$

It remains to upper bound $\left.\int_{a}^{b} \frac{d R}{n} \frac{\partial^{2} f_{n}}{\partial R^{2}}\right|_{t, R}=\left.\frac{1}{n} \frac{\partial f_{n}}{\partial R}\right|_{t, R=b}-\left.\frac{1}{n} \frac{\partial f_{n}}{\partial R}\right|_{t, R=a}$. Note that $\forall R \in(0,+\infty)$ :

$$
\begin{equation*}
0 \leq\left.\frac{\partial f_{n}}{\partial R}\right|_{t, R}=-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}=\frac{\lambda}{4} \mathbb{E}\langle Q\rangle_{t, R}=\frac{\lambda}{4 n} \mathbb{E}\left\|\langle\mathbf{x}\rangle_{t, R}\right\|^{2} \leq \frac{\lambda}{4}\|\varphi\|_{\infty}^{2} \tag{5.76}
\end{equation*}
$$

where the first equality follows from (5.71), the second from (5.57) in Lemma 5.13 , and the third from the Nishimori identity. Integrating both sides of (5.73) over $R \in(a, b)$ and using (5.75), 5.76) yields the first inequality in the proposition,

$$
\int_{a}^{b} \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R \leq \frac{\lambda\|\varphi\|_{\infty}^{2}}{4 n}\left(\frac{\ln (b / a)}{2}+1\right)
$$

To prove the second inequality, we first integrate both sides of the inequality (5.58) with respect to $R$ and then use Cauchy-Schwarz inequality. We obtain

$$
\begin{align*}
\frac{\lambda}{4} \int_{a}^{b} \mathbb{E}\langle(Q- & \left.\left.\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R \\
& \leq\|\varphi\|_{\infty}^{2} \sqrt{\frac{b-a}{2} \int_{a}^{b}\left(\mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\frac{1}{n} \mathbb{E}\left\langle\frac{\partial \mathcal{L}}{\partial R}\right\rangle_{t, R}\right) d R} \\
& =\|\varphi\|_{\infty}^{2} \sqrt{\left.\frac{b-a}{2} \int_{a}^{b} \frac{d R}{n} \frac{\partial^{2} f_{n}}{\partial R^{2}}\right|_{t, R}} \\
& \leq\|\varphi\|_{\infty}^{3} \sqrt{\frac{\lambda(b-a)}{8 n}} \tag{5.77}
\end{align*}
$$

Finally, note that

$$
\begin{align*}
\mathbb{E}\left\langle\left(Q-\left\|\frac{\langle\mathbf{x}\rangle_{t, R}}{\sqrt{n}}\right\|^{2}\right)^{2}\right\rangle_{t, R} & =\mathbb{E}\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R}+\mathbb{E}\left[\left(\langle Q\rangle_{t, R}-\left\|\frac{\langle\mathbf{x}\rangle_{t, R}}{\sqrt{n}}\right\|^{2}\right)^{2}\right] \\
& =\mathbb{E}\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R}+\mathbb{E}\left[\left\langle Q-\frac{\langle\mathbf{x}\rangle_{t, R}^{\top} \mathbf{x}}{n}\right\rangle_{t, R}^{2}\right] \\
& \leq \mathbb{E}\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R}+\mathbb{E}\left[\left\langle\left(Q-\frac{\langle\mathbf{x}\rangle_{t, R}^{\top} \mathbf{x}}{n}\right)^{2}\right\rangle_{t, R}\right] \\
& =\mathbb{E}\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R}+\mathbb{E}\left[\left\langle\left(Q-\frac{\langle\mathbf{x}\rangle_{t, R}^{\top} \mathbf{X}}{n}\right)^{2}\right\rangle_{t, R}\right] \\
& =2 \mathbb{E}\left\langle\left(Q-\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R}, \tag{5.78}
\end{align*}
$$

where the inequality follows from Jensen's inequality and the subsequent equality from the Nishimori identity. Combining (5.77) and (5.78) gives the second inequality in the proposition.
Proposition 5.15 (Quenched fluctuations of $\mathcal{L}$ ). Suppose that (H1) and (H2) hold. Let $M$ be a positive real number. For n large enough, there exists a constant $C$ that depends only on $\|\varphi\|_{\infty},\left\|\varphi^{\prime}\right\|_{\infty},\left\|\varphi^{\prime \prime}\right\|_{\infty}, M_{S}, \lambda, M$, and such that $\forall b \in(0, M), \forall a \in(0, \min \{1, b\}), \forall \delta \in(0, a), \forall t \in[0,1]:$

$$
\begin{equation*}
\int_{a}^{b} \mathbb{E}\left[\left(\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right] d R \leq C\left(\frac{1}{\delta^{2} n}-\frac{\ln (a)}{n}+\frac{\delta}{a-\delta}\right) . \tag{5.79}
\end{equation*}
$$

Proof. Fix $(n, t) \in \mathbb{N}^{*} \times[0,1]$. For all $R \in(0,+\infty)$, we have

$$
\begin{align*}
\left.\frac{\partial F_{n}}{\partial R}\right|_{t, R} & =-\langle\mathcal{L}\rangle_{t, R}  \tag{5.80}\\
\left.\frac{\partial^{2} F_{n}}{\partial R^{2}}\right|_{t, R} & =n\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\frac{1}{4 R} \sqrt{\frac{\lambda}{2 R}} \frac{\langle\mathbf{x}\rangle_{t, R}^{\top} \widetilde{\mathbf{Z}}}{n}  \tag{5.81}\\
\left.\frac{\partial f_{n}}{\partial R}\right|_{t, R} & =-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}  \tag{5.82}\\
\left.\frac{\partial^{2} f_{n}}{\partial R^{2}}\right|_{t, R} & =n \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R}-\frac{1}{4 R} \sqrt{\frac{\lambda}{2 R}} \frac{\mathbb{E}\left[\langle\mathbf{x}\rangle_{t, R}^{\top} \widetilde{\mathbf{Z}}\right]}{n} . \tag{5.83}
\end{align*}
$$

The second term on the right-hand side of (5.81) can be upper bounded with Cauchy-Schwarz inequality,

$$
\begin{equation*}
\left|\frac{1}{4 R} \sqrt{\frac{\lambda}{2 R}} \frac{\langle\mathbf{x}\rangle_{t, R}^{\top} \widetilde{\mathbf{Z}}}{n}\right| \leq \frac{1}{4 R} \sqrt{\frac{\lambda}{2 R}} \frac{\left\|\langle\mathbf{x}\rangle_{t, R}\right\|\|\widetilde{\mathbf{Z}}\|}{n} \leq \frac{\|\varphi\|_{\infty}}{4 R} \sqrt{\frac{\lambda}{2 R}} \frac{\|\widetilde{\mathbf{Z}}\|}{\sqrt{n}} . \tag{5.84}
\end{equation*}
$$

We now define for all $R \in(0,+\infty)$ :

$$
\begin{align*}
F(R) & :=F_{n}(t, R)-\|\varphi\|_{\infty} \sqrt{\frac{\lambda R}{2}} \frac{\|\widetilde{\mathbf{Z}}\|}{\sqrt{n}}  \tag{5.85}\\
f(R) & :=f_{n}(t, R)-\|\varphi\|_{\infty} \sqrt{\frac{\lambda R}{2}} \frac{\mathbb{E}\|\widetilde{\mathbf{Z}}\|}{\sqrt{n}} . \tag{5.86}
\end{align*}
$$

$F$ is convex on $(0,+\infty)$ as it is twice differentiable with a nonnegative second derivative by (5.81) and (5.84). The same holds for $f$. Note that for all $R \in$ $(0,+\infty)$ :

$$
\begin{aligned}
F(R)-f(R) & =F_{n}(t, R)-f_{n}(t, R)-\|\varphi\|_{\infty} \sqrt{\frac{\lambda R}{2}} \frac{\|\widetilde{\mathbf{Z}}\|-\mathbb{E}\|\widetilde{\mathbf{Z}}\|}{\sqrt{n}} \\
F^{\prime}(R)-f^{\prime}(R) & =-\left(\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)-\frac{\|\varphi\|_{\infty}}{2} \sqrt{\frac{\lambda}{2 R}} \frac{\|\widetilde{\mathbf{Z}}\|-\mathbb{E}\|\widetilde{\mathbf{Z}}\|}{\sqrt{n}}
\end{aligned}
$$

It follows from Lemma 2.8 (applied to the convex functions $G=F, g=f$ ) and these last two identities that $\forall R \in(0,+\infty), \forall \delta \in(0, R)$ :

$$
\begin{aligned}
&\left|\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right| \leq \frac{\|\varphi\|_{\infty}}{2} \sqrt{\frac{\lambda}{2 R}} \frac{|\|\widetilde{\mathbf{Z}}\|-\mathbb{E}\|\widetilde{\mathbf{Z}}\||}{\sqrt{n}}+C_{\delta}(R) \\
& \quad+\frac{1}{\delta} \sum_{x \in\{-\delta, 0, \delta\}}|F(R+x)-f(R+x)| \\
& \leq\|\varphi\|_{\infty} \sqrt{\frac{\lambda}{2}}\left(\frac{1}{2 \sqrt{R}}+3 \sqrt{R}\right) \frac{|\|\widetilde{\mathbf{Z}}\|-\mathbb{E}\|\widetilde{\mathbf{Z}}\||}{\sqrt{n}}+C_{\delta}(R) \\
& \left.+\frac{1}{\delta} \sum_{x \in\{-\delta, 0, \delta\}} \right\rvert\, F_{n}(t, R+x)-f_{n}(t,(R+x) \mid,
\end{aligned}
$$

where $C_{\delta}(r):=f^{\prime}(r+\delta)-f^{\prime}(r-\delta)$ is nonnegative by convexity of $f$. We now use the inequality $\left(\sum_{i=1}^{5} v_{i}\right)^{2} \leq 5 \sum_{i=1}^{5} v_{i}^{2}$ to obtain $\forall R \in(0,+\infty), \forall \delta \in(0, R)$ :

$$
\begin{align*}
\mathbb{E}\left[\left(\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right] \leq & 5\|\varphi\|_{\infty}^{2} \frac{\lambda}{2}\left(\frac{1}{4 R}+3+9 R\right) \frac{\operatorname{Var}\|\widetilde{\mathbf{Z}}\|}{n}+5 C_{\delta}(R)^{2} \\
& +\frac{5}{\delta^{2}} \sum_{x \in\{-\delta, 0, \delta\}} \mathbb{E}\left[\left(F_{n}(t, R+x)-f_{n}(t, R+x)\right)^{2}\right] . \tag{5.87}
\end{align*}
$$

The next step is to bound the integral of the three summands on the right-hand side of (5.87). By 99, Theorem 3.1.1], there exists $C_{1}$ such that $\mathbb{V a r}\|\widetilde{\mathbf{Z}}\| \leq C_{1}$ independently of the dimension $n$. Then,

$$
\begin{equation*}
\int_{a}^{b} d R\left(\frac{1}{4 R}+3+9 R\right) \frac{\operatorname{Var}\|\widetilde{\mathbf{Z}}\|}{n} \leq\left(\frac{\ln (b / a)}{4}+3 b+\frac{9}{2} b^{2}\right) \frac{C_{1}}{n} . \tag{5.88}
\end{equation*}
$$

Note that $C_{\delta}(R)=\left|C_{\delta}(R)\right| \leq\left|f^{\prime}(R+\delta)\right|+\left|f^{\prime}(R-\delta)\right|$ and for all $R \in(0,+\infty)$ :

$$
\left|f^{\prime}(R)\right| \leq\left|\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right|+\frac{\|\varphi\|_{\infty}}{2} \sqrt{\frac{\lambda}{2 R}} \frac{\mathbb{E}\|\widetilde{\mathbf{Z}}\|}{\sqrt{n}} \leq \frac{1}{2} \sqrt{\frac{\lambda\|\varphi\|_{\infty}^{2}}{2}}\left(\sqrt{\frac{\lambda}{2}}\|\varphi\|_{\infty}+\frac{1}{\sqrt{R}}\right)
$$

where the second inequality is due to the upper bounds $\left|\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right| \leq \lambda\|\varphi\|_{\infty}^{2} / 4$ (see
(5.76) and $\mathbb{E}\|\widetilde{\mathbf{Z}}\| \leq \sqrt{n}$. Thus, for the second summand, we obtain $\forall \delta \in(0, a)$ :

$$
\begin{align*}
& \int_{a}^{b} d R C_{\delta}(R)^{2} \leq \sqrt{\frac{\lambda\|\varphi\|_{\infty}^{2}}{2}}\left(\sqrt{\frac{\lambda}{2}}\|\varphi\|_{\infty}+\frac{1}{\sqrt{a-\delta}}\right) \int_{a}^{b} d R C_{\delta}(R) \\
&=\sqrt{\frac{\lambda\|\varphi\|_{\infty}^{2}}{2}}\left(\sqrt{\frac{\lambda}{2}}\|\varphi\|_{\infty}+\frac{1}{\sqrt{a-\delta}}\right) \\
& \cdot[f(b+\delta)-f(b-\delta)-(f(a+\delta)-f(a-\delta))] \\
& \leq \lambda\|\varphi\|_{\infty}^{2}\left(\sqrt{\frac{\lambda}{2}}\|\varphi\|_{\infty}+\frac{1}{\sqrt{a-\delta}}\right)^{2} \delta \tag{5.89}
\end{align*}
$$

The last inequality is a simple application of the mean value theorem. We turn to the third and last summand. By Proposition 5.16 in Appendix 5.D, there exists a positive constant $C_{2}$ depending only on $a, b,\|\varphi\|_{\infty},\left\|\varphi^{\prime}\right\|_{\infty},\left\|\varphi^{\prime \prime}\right\|_{\infty}, M_{S}$ and $\lambda$ such that $\forall(t, R) \in[0,1] \times(0, b+a)$ :

$$
\begin{equation*}
\mathbb{E}\left[\left(F_{n}(t, R)-f_{n}(t, R)\right)^{2}\right] \leq \frac{C_{2}}{n} \tag{5.90}
\end{equation*}
$$

Using (5.90), we see that the integral of the third summand satisfies $\forall \delta \in(0, a)$ :

$$
\begin{equation*}
\int_{a}^{b} d R \frac{5}{\delta^{2}} \sum_{x \in\{-\delta, 0, \delta\}} \mathbb{E}\left[\left(F_{n}(t, R+x)-f_{n}(t, R+x)\right)^{2}\right] \leq \frac{15 C_{2}}{\delta^{2} n} b \tag{5.91}
\end{equation*}
$$

To end the proof, we integrate over $R \in(a, b)$ on both sides of (5.87) and use the three upper bounds (5.88), (5.89), 5.91).

## 5.C. 2 Concentration of $Q$ around its expectation

Proof of Proposition 5.12. Using the upper bound (5.59), it directly comes

$$
\frac{\lambda^{2}}{16} \int_{a}^{b} \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R \leq \int_{a}^{b} \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R
$$

We then use Propositions 5.14 and 5.15 to upper bound

$$
\begin{aligned}
& \int_{a}^{b} \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R \\
&=\int_{a}^{b} \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right\rangle_{t, R} d R+\int_{a}^{b} \mathbb{E}\left[\left(\langle\mathcal{L}\rangle_{t, R}-\mathbb{E}\langle\mathcal{L}\rangle_{t, R}\right)^{2}\right] d R
\end{aligned}
$$

ending the proof.

## 5.D Concentration of the free entropy

Consider the inference problem (5.28). Once both observations $\mathbf{Y}^{(t)}$ and $\widetilde{\mathbf{Y}}^{(t, R)}$ have been replaced by their definitions, the Hamiltonian associated with the
posterior distribution of $\mathbf{S}$ given $\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)$ reads

$$
\begin{aligned}
& \mathcal{H}_{t, R}(\mathbf{s} ; \mathbf{S}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \mathbf{W}):=\sum_{j=1}^{n} \frac{\lambda R}{4} x_{j}^{2}-\frac{\lambda R}{2} X_{j} x_{j}-\sqrt{\frac{\lambda R}{2}} \widetilde{Z}_{j} x_{j} \\
& \quad+\sum_{\underline{i} \in \mathcal{I}} \frac{\lambda(1-t)}{2 n^{2}} x_{i_{1}}^{2} x_{i_{2}}^{2} x_{i_{3}}^{2}-\frac{\lambda(1-t)}{n^{2}} X_{i_{1}} X_{i_{2}} X_{i_{3}} x_{i_{1}} x_{i_{2}} x_{i_{3}}-\frac{\sqrt{\lambda(1-t)}}{n} Z_{\underline{i}} x_{i_{1}} x_{i_{2}} x_{i_{3}} .
\end{aligned}
$$

In this section, we show that the free entropy

$$
\begin{equation*}
\frac{1}{n} \ln \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, R)}, \mathbf{W}\right)=\frac{1}{n} \ln \int d P_{s}(\mathbf{s}) e^{-\mathcal{H}_{t, R}(\mathbf{s} ; \mathbf{S}, \mathbf{Z}, \tilde{\mathbf{Z}}, \mathbf{W})} \tag{5.92}
\end{equation*}
$$

concentrates around its expectation. To shorten notations we write $\frac{\ln \mathcal{Z}_{t, R}}{n}$, omitting the arguments.

Proposition 5.16 (Concentration of the free entropy). Suppose that (H1), (H2) hold. There exists a polynomial $C\left(\|\varphi\|_{\infty},\left\|\varphi^{\prime}\right\|_{\infty},\left\|\varphi^{\prime \prime}\right\|_{\infty}, M_{S}, \lambda, R\right)$ with positive coefficients such that $\forall t \in[0,1]$ :

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{t, R}}{n}-\mathbb{E}\left[\frac{\ln \mathcal{Z}_{t, R}}{n}\right]\right)^{2}\right] \leq \frac{C\left(\|\varphi\|_{\infty},\left\|\varphi^{\prime}\right\|_{\infty},\left\|\varphi^{\prime \prime}\right\|_{\infty}, M_{S}, \lambda, R\right)}{n} \tag{5.93}
\end{equation*}
$$

Proof. First, we show that the free entropy concentrates on its conditional expectation given ( $\mathbf{W}, \mathbf{S}$ ). We see $\ln \mathcal{Z}_{t, R} / n$ as a function of the Gaussian random variables $\mathbf{Z}, \widetilde{\mathbf{Z}}$ and we work conditionally to $(\mathbf{W}, \mathbf{S})$. Let $g(\mathbf{Z}, \widetilde{\mathbf{Z}}):=\ln \mathcal{Z}_{t, R} / n$. By the Gaussian-Poincaré inequality (see Proposition 2.7.

$$
\mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{t, R}}{n}-\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{S}, \mathbf{W}\right]\right)^{2}\right] \leq \mathbb{E}\left[\|\nabla g(\mathbf{Z}, \widetilde{\mathbf{Z}})\|^{2}\right]
$$

The squared norm of the gradient of $g$ reads $\|\nabla g\|^{2}=\sum_{\underline{i} \in \mathcal{I}}\left|\partial g / \partial z_{\underline{i}}\right|^{2}+\sum_{j}\left|\partial g / \partial \tilde{z}_{j}\right|^{2}$. Each of these partial derivatives takes the form $\partial g / \partial x=-n^{-1}\left\langle\partial \mathcal{H}_{t, R} / \partial x\right\rangle$. More precisely,

$$
\frac{\partial g}{\partial Z_{\underline{i}}}=\frac{\sqrt{\lambda(1-t)}}{n^{2}}\left\langle x_{i_{1}} x_{i_{2}} x_{i_{3}}\right\rangle_{t, R} \quad ; \quad \frac{\partial g}{\partial \widetilde{Z}_{j}}=\frac{1}{n} \sqrt{\frac{\lambda R}{2}}\left\langle x_{j}\right\rangle_{t, R} .
$$

We see that $\left|\partial g / \partial Z_{\underline{\underline{1}}}\right| \leq \frac{\sqrt{\lambda}}{n^{2}}\|\varphi\|_{\infty}^{3}$ and $\left|\frac{\partial g}{\partial \widetilde{Z}_{j}}\right| \leq \frac{1}{n} \sqrt{\frac{\lambda R}{2}}\|\varphi\|_{\infty}$. Therefore,

$$
\|\nabla g(\mathbf{Z}, \widetilde{\mathbf{Z}})\|^{2} \leq \frac{\lambda}{2 n}\|\varphi\|_{\infty}^{6}+\frac{\lambda R}{2 n}\|\varphi\|_{\infty}^{2} .
$$

Making use of the Gaussian-Poincaré inequality, we obtain

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{t, R}}{n}-\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{S}, \mathbf{W}\right]\right)^{2}\right] \leq \frac{\lambda\|\varphi\|_{\infty}^{2}}{2 n}\left(\|\varphi\|_{\infty}^{4}+R\right) \tag{5.94}
\end{equation*}
$$

Next we show that $\mathbb{E}\left[\ln \mathcal{Z}_{t, R} / n \mid \mathbf{S}, \mathbf{W}\right]$ concentrates on its conditional expectation given $\mathbf{S}$. We see $\ln \mathcal{Z}_{t, R} / n$ as a function of the standard Gaussian random variables $\mathbf{W}$ and we work conditionally to $\mathbf{S}$. Let $g(\mathbf{W}):=\mathbb{E}\left[\ln \mathcal{Z}_{t, R} / n \mid \mathbf{W}, \mathbf{S}\right]$. We again rely on the Gaussian-Poincaré inequality. To lighten the equations, we drop the subscripts of the angular brackets $\langle-\rangle_{t, R}$, introduce the notation $\widetilde{\mathbb{E}}[\cdot]:=\mathbb{E}[\cdot \mid \mathbf{S}, \mathbf{W}]$, and define

$$
\mathbf{X}^{\prime}:=\varphi^{\prime}\left(\frac{\mathbf{W S}}{\sqrt{p}}\right) \quad ; \quad \mathbf{x}^{\prime}:=\varphi^{\prime}\left(\frac{\mathbf{W S}}{\sqrt{p}}\right)
$$

The squared norm of the gradient of $g$ reads $\|\nabla g\|^{2}=\sum_{i, j}\left|\partial g / \partial W_{i j}\right|^{2}$ where $\forall(i, j) \in\{1, \ldots, n\} \times\{1, \ldots, p\}$ :

$$
\begin{aligned}
\frac{\partial g}{\partial W_{i j}}= & O\left(n^{-5 / 2}\right) \\
+ & \frac{1}{2 n} \sum_{\substack{k=1 \\
k \neq i}}^{n} \sum_{\substack{\ell=1 \\
\ell \neq k, i}}^{n}\left(-\frac{\lambda(1-t)}{n^{2} \sqrt{p}} \widetilde{\mathbb{E}}\left\langle x_{i} x_{i}^{\prime} s_{j} x_{k}^{2} x_{\ell}^{2}\right\rangle+\frac{\lambda(1-t)}{n^{2} \sqrt{p}} S_{j} X_{i}^{\prime} X_{k} X_{\ell} \widetilde{\mathbb{E}}\left\langle x_{i} x_{k} x_{\ell}\right\rangle\right. \\
& \left.+\frac{\lambda(1-t)}{n^{2} \sqrt{p}} X_{i} X_{k} X_{\ell} \widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime} x_{k} x_{\ell}\right\rangle+\frac{\sqrt{\lambda(1-t)}}{n \sqrt{p}} \widetilde{\mathbb{E}} Z_{i k \ell}\left\langle s_{j} x_{i}^{\prime} x_{k} x_{\ell}\right\rangle\right) \\
+ & \frac{1}{n} \sqrt{\frac{\lambda R}{2 p}}\left(-\sqrt{\frac{\lambda R}{2}} \widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime} x_{i}\right\rangle+\sqrt{\frac{\lambda R}{2}} S_{j} X_{i}^{\prime} \widetilde{\mathbb{E}}\left\langle x_{i}\right\rangle\right. \\
& \left.+\sqrt{\frac{\lambda R}{2}} X_{i} \widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime}\right\rangle+\widetilde{\mathbb{E}} \widetilde{Z}_{i}\left\langle s_{j} x_{i}^{\prime}\right\rangle\right)
\end{aligned}
$$

The quantity $O\left(n^{-5 / 2}\right)$ corresponds to the terms associated with triplets $\underline{i} \in$ $\mathcal{I}$ whose elements are non unique. In order to further simplify these partial derivatives, we do a Gaussian integration by parts with respect to $\mathbf{Z}$ and $\widetilde{\mathbf{Z}}$. It yields

$$
\begin{aligned}
\widetilde{\mathbb{E}} Z_{i k \ell}\left\langle s_{j} x_{i}^{\prime} x_{k} x_{\ell}\right\rangle & =\frac{\sqrt{\lambda(1-t)}}{n} \widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime} x_{i} x_{k}^{2} x_{\ell}^{2}\right\rangle-\frac{\sqrt{\lambda(1-t)}}{n} \widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime} x_{k} x_{\ell}\right\rangle\left\langle x_{i} x_{k} x_{\ell}\right\rangle \\
\widetilde{\mathbb{E}} \widetilde{Z}_{i}\left\langle s_{j} x_{i}^{\prime}\right\rangle & =\sqrt{\frac{\lambda R}{2}} \widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime} x_{i}\right\rangle-\sqrt{\frac{\lambda R}{2}} \widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime}\right\rangle\left\langle x_{i}\right\rangle
\end{aligned}
$$

Therefore, $\forall(i, j) \in\{1, \ldots, n\} \times\{1, \ldots, p\}$ :

$$
\begin{gathered}
\frac{\partial g}{\partial W_{i j}}=O\left(n^{-5 / 2}\right)+\frac{\lambda R}{2 n \sqrt{p}}\left(S_{j} X_{i}^{\prime} \widetilde{\mathbb{E}}\left\langle x_{i}\right\rangle+X_{i} \widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime}\right\rangle-\widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime}\right\rangle\left\langle x_{i}\right\rangle\right) \\
+\frac{\lambda(1-t)}{2 n^{3} \sqrt{p}} \sum_{\substack{k=1 \\
k \neq i}}^{n} \sum_{\substack{\ell=1 \\
\ell \neq k, i}}^{n}\left(S_{j} X_{i}^{\prime} X_{k} X_{\ell} \widetilde{\mathbb{E}}\left\langle x_{i} x_{k} x_{\ell}\right\rangle+X_{i} X_{k} X_{\ell} \widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime} x_{k} x_{\ell}\right\rangle\right. \\
\left.-\widetilde{\mathbb{E}}\left\langle s_{j} x_{i}^{\prime} x_{k} x_{\ell}\right\rangle\left\langle x_{i} x_{k} x_{\ell}\right\rangle\right)
\end{gathered}
$$

Making use of the boundedness assumptions for $\varphi$ and its derivative, we obtain the following uniform bound on the partial derivatives,

$$
\left|\frac{\partial g}{\partial W_{i j}}\right| \leq O\left(n^{-5 / 2}\right)+\frac{3 \lambda M_{S}}{2 n \sqrt{p}}\|\varphi\|_{\infty}\left\|\varphi^{\prime}\right\|_{\infty}\left(\|\varphi\|_{\infty}^{4}+R\right) .
$$

Therefore, $\|\nabla g(\mathbf{W})\|^{2} \leq \frac{9 \lambda^{2} M_{S}^{2}}{4 n}\|\varphi\|_{\infty}^{2}\left\|\varphi^{\prime}\right\|_{\infty}^{2}\left(\|\varphi\|_{\infty}^{4}+R\right)^{2}+O\left(n^{-3}\right)$ and, by the Gaussian-Poincaré inequality (the negligible term $O\left(n^{-3}\right)$ is omitted),

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{S}, \mathbf{W}\right]-\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{S}\right]\right)^{2}\right] \leq \frac{9 \lambda^{2} M_{S}^{2}}{4 n}\|\varphi\|_{\infty}^{2}\left\|\varphi^{\prime}\right\|_{\infty}^{2}\left(\|\varphi\|_{\infty}^{4}+R\right)^{2} \tag{5.95}
\end{equation*}
$$

Finally, it remains to demonstrate that $\mathbb{E}\left[\ln \mathcal{Z}_{t, R / n} \mid \mathbf{S}\right]$ concentrates on its expectation. Let us show that the function

$$
g: \mathrm{S} \in\left[-M_{S}, M_{S}\right]^{p} \mapsto \mathbb{E}\left[\ln \mathcal{z}_{t, R} / n \mid \mathbf{S}=\mathrm{S}\right]
$$

satisfies the bounded difference property. To do so, we show that the partial derivatives of $g$ are uniformly bounded. Then, we apply McDiarmid's inequality (see Proposition 2.6. For a particular realization $\mathbf{S}=\mathbf{S}$, we define $\mathrm{X}:=\varphi\left(\frac{\mathbf{W S}}{\sqrt{p}}\right)$, $\mathrm{X}^{\prime}:=\varphi^{\prime}\left(\frac{\mathrm{WS}}{\sqrt{p}}\right)$ and $\mathrm{X}^{\prime \prime}:=\varphi^{\prime \prime}\left(\frac{\mathrm{WS}}{\sqrt{p}}\right)$ in typewriter font. We denote $\mathbb{E}_{\mathrm{S}}[\cdot]:=\mathbb{E}[\cdot \mid \mathbf{S}=\mathrm{S}]$ the expectation given a particular realization $\mathbf{S}=\mathbf{S}$. For $\ell \in\{1, \ldots, p\}$, the partial derivative of $g$ with respect to its $\ell^{\text {th }}$ coordinate reads

$$
\begin{align*}
\frac{\partial g}{\partial \mathrm{~S}_{\ell}}= & \frac{\lambda R}{2 n \sqrt{p}} \sum_{i=1}^{n} \mathbb{E}_{\mathrm{S}}\left[W_{i \ell} \mathrm{X}_{i}^{\prime}\left\langle x_{i}\right\rangle\right] \\
& +\frac{\lambda(1-t)}{n^{3} \sqrt{p}} \sum_{i \in \mathcal{I}} \mathbb{E}_{\mathrm{S}}\left[\left(W_{i_{1} \ell} \mathrm{X}_{i_{1}}^{\prime} \mathrm{X}_{i_{2}} \mathrm{X}_{i_{3}}+W_{i_{2} \ell} \mathrm{X}_{i_{1}} \mathrm{X}_{i_{2}}^{\prime} \mathrm{X}_{i_{3}}+W_{i_{3} \ell} \mathrm{X}_{i_{1}} \mathrm{X}_{i_{2}} \mathrm{X}_{i_{3}}^{\prime}\right)\left\langle x_{i_{1}} x_{i_{2}} x_{i_{3}}\right\rangle\right] \\
= & \frac{\lambda R}{2 n \sqrt{p}} \sum_{i=1}^{n} \mathbb{E}_{\mathrm{S}}\left[W_{i \ell} \mathrm{X}_{i}^{\prime}\left\langle x_{i}\right\rangle\right] \\
& \quad+O\left(n^{-3 / 2}\right)+\frac{\lambda(1-t)}{2 n^{3} \sqrt{p}} \sum_{i=1}^{n} \sum_{\substack{j=1 \\
j \neq i}}^{n} \sum_{\substack{k=1 \\
k \neq i, j}}^{n} \mathbb{E}_{\mathrm{S}}\left[W_{i \ell} \mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k}\left\langle x_{i} x_{j} x_{k}\right\rangle\right] \tag{5.96}
\end{align*}
$$

Once again the triplets $\underline{\boldsymbol{i}} \in \mathcal{I}$ whose elements are non unique are accounted for with the term $O\left(n^{-3 / 2}\right)$ that is negligible compared to the others. A Gaussian integration by parts with respect to $\mathbf{W}$ gives for all $(i, j, k, \ell) \in\{1, \ldots, n\}^{3} \times\{1, \ldots, p\}$ such that $j \neq i$ and $k \neq i, j$ :

$$
\begin{align*}
\mathbb{E}_{\mathrm{S}}\left[W_{i \ell} \mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k}\left\langle x_{i} x_{j} x_{k}\right\rangle\right]=\frac{1}{\sqrt{p}} \mathbb{E}_{\mathrm{S}} & {\left[\mathrm{~S}_{\ell} \mathrm{X}_{i}^{\prime \prime} \mathrm{X}_{j} \mathrm{X}_{k}\left\langle x_{i} x_{j} x_{k}\right\rangle\right]+\frac{1}{\sqrt{p}} \mathbb{E}_{\mathrm{S}}\left[\mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k}\left\langle s_{\ell} x_{i}^{\prime} x_{j} x_{k}\right\rangle\right] } \\
& -\mathbb{E}_{\mathrm{S}}\left[\mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k}\left\langle x_{i} x_{j} x_{k} \frac{\partial \mathcal{H}_{t, R}}{\partial W_{i \ell}}\right\rangle\right] \\
& +\mathbb{E}_{\mathrm{S}}\left[\mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k}\left\langle x_{i} x_{j} x_{k}\right\rangle\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial W_{i \ell}}\right\rangle\right] \tag{5.97}
\end{align*}
$$

$$
\begin{align*}
\mathbb{E}_{\mathrm{S}}\left[W_{i \ell} \mathrm{X}_{i}^{\prime}\left\langle x_{i}\right\rangle\right]= & \frac{1}{\sqrt{p}} \mathbb{E}_{\mathrm{S}}\left[\mathrm{~S}_{\ell} \mathrm{X}_{i}^{\prime \prime}\left\langle x_{i}\right\rangle\right]+\frac{1}{\sqrt{p}} \mathbb{E}_{\mathrm{S}}\left[\mathrm{X}_{i}^{\prime}\left\langle s_{\ell} x_{i}^{\prime}\right\rangle\right] \\
& -\mathbb{E}_{\mathrm{S}}\left[\mathrm{X}_{i}^{\prime}\left\langle x_{i} \frac{\partial \mathcal{H}_{t, R}}{\partial W_{i \ell}}\right\rangle\right]+\mathbb{E}_{\mathrm{S}}\left[\mathrm{X}_{i}^{\prime}\left\langle x_{i}\right\rangle\left\langle\frac{\partial \mathcal{H}_{t, R}}{\partial W_{i \ell}}\right\rangle\right] \tag{5.98}
\end{align*}
$$

where here $\mathcal{H}_{t, R}=\mathcal{H}_{t, R}(\mathbf{s} ; \mathbf{S}=\mathbf{S}, \mathbf{Z}, \widetilde{\mathbf{Z}}, \mathbf{W})$. In order to prove the concentration result that we aim for, we need to check that the expectations(5.97) and (5.98) are both $O\left(n^{-1 / 2}\right)$. The main difficulty resides in managing the terms where partial derivatives $\partial \mathcal{H}_{t, R} / \partial W_{i \ell}$ appear. We have already dealt with these partial derivatives when proving the concentration with respect to $\mathbf{W}$ and found

$$
\begin{array}{r}
\frac{\partial \mathcal{H}_{t, R}}{\partial W_{i \ell}}=O\left(n^{-3 / 2}\right)+\frac{1}{2} \sum_{\substack{j=1 \\
j \neq i}}^{n} \sum_{\substack{k=1 \\
k \neq i, j}}^{n}\left(-\frac{\lambda(1-t)}{n^{2} \sqrt{p}} x_{i} x_{i}^{\prime} s_{\ell} x_{j}^{2} x_{k}^{2}+\frac{\lambda(1-t)}{n^{2} \sqrt{p}} \mathrm{~S}_{\ell} \mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k} x_{i} x_{j} x_{k}\right. \\
\left.+\frac{\lambda(1-t)}{n^{2} \sqrt{p}} \mathrm{X}_{i} \mathrm{X}_{j} \mathrm{X}_{k} s_{\ell} x_{i}^{\prime} x_{j} x_{k}+\frac{\sqrt{\lambda(1-t)}}{n \sqrt{p}} Z_{i j k} s_{\ell} x_{i}^{\prime} x_{j} x_{k}\right) \\
+\sqrt{\frac{\lambda R}{2 p}}\left(-\sqrt{\frac{\lambda R}{2}} s_{\ell} x_{i}^{\prime} x_{i}+\sqrt{\frac{\lambda R}{2}} \mathrm{~S}_{\ell} \mathrm{X}_{i}^{\prime} x_{i}+\sqrt{\frac{\lambda R}{2}} \mathrm{X}_{i} s_{\ell} x_{i}^{\prime}+\widetilde{Z}_{i} s_{\ell} x_{i}^{\prime}\right)
\end{array}
$$

For $(i, \ell) \in\{1, \ldots, n\} \times\{1, \ldots, p\}$ define

$$
\begin{array}{r}
\mathcal{A}_{i \ell}:=\frac{\lambda(1-t)}{2 n^{2} \sqrt{p}} \sum_{\substack{j=1 \\
j \neq i}}^{n} \sum_{\substack{k=1 \\
k \neq i, j}}^{n}\left(-x_{i} x_{i}^{\prime} s_{\ell} x_{j}^{2} x_{k}^{2}+\mathrm{S}_{\ell} \mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k} x_{i} x_{j} x_{k}+\mathrm{X}_{i} \mathrm{X}_{j} \mathrm{X}_{k} s_{\ell} x_{i}^{\prime} x_{j} x_{k}\right) \\
 \tag{5.99}\\
+ \\
\frac{\lambda R}{2 \sqrt{p}}\left(-s_{\ell} x_{i}^{\prime} x_{i}+\mathrm{S}_{\ell} \mathrm{X}_{i}^{\prime} x_{i}+\mathrm{X}_{i} s_{\ell} x_{i}^{\prime}\right)
\end{array}
$$

Note that

$$
\begin{align*}
\frac{\partial \mathcal{H}_{t, R}}{\partial W_{i \ell}} & =O\left(n^{-3 / 2}\right)+\mathcal{A}_{i \ell}+\frac{\sqrt{\lambda(1-t)}}{2 n \sqrt{p}} \sum_{\substack{j=1 \\
j \neq i}}^{n} \sum_{\substack{k=1 \\
k \neq i, j}}^{n} Z_{i j k} s_{\ell} x_{i}^{\prime} x_{j} x_{k}+\sqrt{\frac{\lambda R}{2 p}} \widetilde{Z}_{i} s_{\ell} x_{i}^{\prime}  \tag{5.100}\\
\left|\mathcal{A}_{i \ell}\right| & \leq \frac{3 \lambda}{2 \sqrt{p}} M_{S}\|\varphi\|_{\infty}\left\|\varphi^{\prime}\right\|_{\infty}\left(\|\varphi\|_{\infty}^{4}+R\right) . \tag{5.101}
\end{align*}
$$

Plugging the identity (5.100) back in (5.97) and 5.98) and making use of the upper bound (5.101) yields

$$
\begin{aligned}
& \left\lvert\, \mathbb{E}_{\mathrm{S}}\left[W_{i \ell} \mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k}\left\langle x_{i} x_{j} x_{k}\right\rangle \left\lvert\, \leq O\left(n^{-3 / 2}\right)+\frac{\|\varphi\|_{\infty}^{4} C_{1}}{\sqrt{p}}\right.\right.\right. \\
& \left.\quad+\frac{1}{2 n} \sqrt{\frac{\lambda}{p}} \sum_{\substack{\prime \\
j^{\prime}=1 \\
j^{\prime} \neq i}}^{n} \sum_{k=1}^{n \neq i^{\prime}, j^{\prime}} \substack{n} \mathbb{E}_{\mathrm{S}}\left[\mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k} Z_{i j^{\prime} k^{\prime}}\left(\left\langle x_{i} x_{j} x_{k} s_{\ell} x_{i}^{\prime} x_{j^{\prime}} x_{k^{\prime}}\right\rangle-\left\langle x_{i} x_{j} x_{k}\right\rangle\left\langle s_{\ell} x_{i}^{\prime} x_{j^{\prime}} x_{k^{\prime}}\right\rangle\right)\right] \right\rvert\, \\
& \quad+\sqrt{\frac{\lambda R}{2 p}}\left|\mathbb{E}_{\mathbf{S}}\left[\mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k} \widetilde{Z}_{i}\left(\left\langle x_{i} x_{j} x_{k} s_{\ell} x_{i}^{\prime}\right\rangle-\left\langle x_{i} x_{j} x_{k}\right\rangle\left\langle s_{\ell} x_{i}^{\prime}\right\rangle\right)\right]\right|
\end{aligned}
$$

$$
\begin{aligned}
\left|\mathbb{E}_{\mathrm{S}}\left[W_{i \ell} \mathrm{X}_{i}^{\prime}\left\langle x_{i}\right\rangle\right]\right| & \leq O\left(n^{-3 / 2}\right)+\frac{C_{1}}{\sqrt{p}}+\sqrt{\frac{\lambda R}{2 p}}\left|\mathbb{E}_{\mathrm{S}}\left[\mathrm{x}_{i}^{\prime} \widetilde{Z}_{i}\left(\left\langle x_{i} s_{\ell} x_{i}^{\prime}\right\rangle-\left\langle x_{i}\right\rangle\left\langle s_{\ell} x_{i}^{\prime}\right\rangle\right)\right]\right| \\
& +\frac{1}{2 n} \sqrt{\frac{\lambda}{p}} \sum_{\substack{j^{\prime}=1 \\
j^{\prime} \neq i}}^{n} \sum_{\substack{k^{\prime}=1 \\
k \neq i^{\prime}, j^{\prime}}}^{n}\left|\mathbb{E}_{\mathrm{S}}\left[\mathrm{x}_{i}^{\prime} Z_{i j^{\prime} k^{\prime}}\left(\left\langle x_{i} s_{\ell} x_{i}^{\prime} x_{j^{\prime}} x_{k^{\prime}}\right\rangle-\left\langle x_{i}\right\rangle\left\langle s_{\ell} x_{i}^{\prime} x_{j^{\prime}} x_{k^{\prime}}\right\rangle\right)\right]\right|
\end{aligned}
$$

where $C_{1}:=M_{S}\left(\|\varphi\|_{\infty}\left\|\varphi^{\prime \prime}\right\|_{\infty}+\left\|\varphi^{\prime}\right\|_{\infty}^{2}+3 \lambda\|\varphi\|_{\infty}^{6}\left\|\varphi^{\prime}\right\|_{\infty}^{2}+3 \lambda\|\varphi\|_{\infty}^{2}\left\|\varphi^{\prime}\right\|_{\infty}^{2} R\right)$. By integrating by parts with respect to $\mathbf{Z}$ or $\widetilde{\mathbf{Z}}$, we can show that both upper bounds are $O\left(p^{-1 / 2}\right)$. This is because $Z_{i j^{\prime} k^{\prime}}$ and $\widetilde{Z}_{i}$ appear in the Hamiltonian $\mathcal{H}_{t, R}$ via the terms $\frac{\sqrt{\lambda(1-t)}}{n} x_{i} x_{j^{\prime}} x_{k^{\prime}} Z_{i j^{\prime} k^{\prime}}$ and $\sqrt{\frac{\lambda R}{2}} x_{i} \widetilde{Z}_{i}$, respectively. In the end, for all $(i, j, k, \ell) \in\{1, \ldots, n\}^{3} \times\{1, \ldots, p\}$ such that $j \neq i$ and $k \neq i, j$ :

$$
\begin{aligned}
\left|\mathbb{E}_{\mathrm{S}}\left[W_{i \ell} \mathrm{X}_{i}^{\prime} \mathrm{X}_{j} \mathrm{X}_{k}\left\langle x_{i} x_{j} x_{k}\right\rangle\right]\right| & \leq \frac{\|\varphi\|_{\infty}^{4}\left(C_{1}+C_{1}^{\prime}\right)}{\sqrt{p}} \\
\left|\mathbb{E}_{\mathrm{S}}\left[W_{i \ell} \mathrm{X}_{i}^{\prime}\left\langle x_{i}\right\rangle\right]\right| & \leq \frac{C_{1}+C_{1}^{\prime}}{\sqrt{p}}
\end{aligned}
$$

where $C_{1}^{\prime}:=M_{S}\left(3 \lambda\|\varphi\|_{\infty}^{6}\left\|\varphi^{\prime}\right\|_{\infty}^{2}+3 \lambda\|\varphi\|_{\infty}^{2}\left\|\varphi^{\prime}\right\|_{\infty}^{2} R\right)$. Returning to the identity (5.96) for the partial derivative, we see that these upper bounds imply that

$$
\left|\frac{\partial g}{\partial \mathbf{S}_{\ell}}\right| \leq \frac{\lambda\left(C_{1}+C_{1}^{\prime}\right)}{2 p}\left(\|\varphi\|_{\infty}^{4}+R\right)
$$

uniformly in $S \in\left[-M_{S}, M_{S}\right]^{p}$ and $\ell \in\{1 \ldots, p\}$. Hence, by the mean value theorem, $g$ has bounded differences, that is, $\forall \ell \in\{1, \ldots, p\}$ :

$$
\sup _{-M_{S} \leq \mathrm{S}_{1}, \ldots, \mathbf{S}_{p}, \mathrm{~S}_{\ell}^{\prime} \leq M_{S}}\left|g(\mathrm{~S})-g\left(\mathrm{~S}_{1}, \ldots, \mathrm{~S}_{\ell-1}, \mathrm{~S}_{\ell}^{\prime}, \mathrm{S}_{\ell+1}, \ldots, \mathrm{~S}_{p}\right)\right| \leq \frac{C_{2}}{p}
$$

where $C_{2}:=\lambda M_{S}\left(C_{1}+C_{1}^{\prime}\right)\left(\|\varphi\|_{\infty}^{4}+R\right)$. By McDiarmid's inequality (see Proposition 2.6),

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \mathcal{Z}_{t, R}}{n} \right\rvert\, \mathbf{S}\right]-\mathbb{E}\left[\frac{\ln \mathcal{Z}_{t, R}}{n}\right]\right)^{2}\right] \leq \frac{C_{2}^{2}}{4 p} \tag{5.102}
\end{equation*}
$$

Combining the inequalities (5.94), (5.95) and (5.102) yields the final result.

## 5.E Proof of Proposition 5.7

The proof is based on the envelope theorem [68, Corollary 4] to obtain the derivative of $h$. We proceed as follows:

1. We show that $h$ is equal to the minimization on a compact subset of a function having sufficient regularity properties to apply [68, Corollary 4].
2. The latter gives a formula for the derivative of $h$ at any point where it is differentiable.
3. We use an optimality condition on $q_{x}^{*} \in \mathcal{Q}_{x}^{*}(\lambda)$ leading to simplified formula (5.38) for $h^{\prime}(\lambda)$.

Proof of Proposition 5.7. We proceed according to the plan outlined above.

1) We define $f\left(q_{x}, q_{s}, \lambda\right):=\sup _{r_{s} \geq 0} \psi_{\lambda, \alpha}\left(q_{x}, q_{s}, r_{s}\right)$. By the definition (5.3) of $\psi_{\lambda, \alpha}$, we have for all $\left(q_{x}, q_{s}, \lambda\right) \in\left[0, \rho_{x}\right] \times\left[0, \rho_{s}\right] \times(0,+\infty)$ :

$$
\begin{equation*}
f\left(q_{x}, q_{s}, \lambda\right)=\frac{1}{\alpha} I_{P_{S}}^{*}\left(\frac{q_{s}-\rho_{s}}{2}\right)+I_{\varphi}\left(\frac{\lambda q_{x}^{2}}{2}, q_{s} ; \rho_{s}\right)+\frac{\lambda}{12}\left(\rho_{x}-q_{x}\right)^{2}\left(\rho_{x}+2 q_{x}\right), \tag{5.103}
\end{equation*}
$$

where the functions $I_{P_{S}}^{*}$ and $I_{\varphi}\left(\cdot, \cdot ; \rho_{s}\right)$ are defined in Lemma 5.8 and Lemma 5.9 , respectively. By Lemma 5.9, $I_{\varphi}\left(\cdot, \cdot ; \rho_{s}\right)$ is continuous on $[0,+\infty) \times\left[0, \rho_{s}\right]$. By Lemma 5.8, $I_{P_{S}}^{*}$ is convex and finite on $(-\infty, 0)$, hence continuous on $(-\infty, 0)$. Besides, $I_{P_{S}}^{*}$ is nondecreasing on $(-\infty, 0)$ and we distinguish between two cases:
(i) $\lim _{x \rightarrow 0^{-}} I_{P_{S}}^{*}(x)$ exists and is finite;
(ii) $\lim _{x \rightarrow 0^{-}} I_{P_{S}}^{*}(x)$ diverges to $+\infty$.

If (i) then, by monotonicity of $I_{P_{S}}^{*}, \lim _{x \rightarrow 0^{-}} I_{P_{S}}^{*}(x) \leq I_{P_{S}}^{*}(0)$. We can redefine $I_{P_{S}}^{*}$ at $x=0$ by $I_{P_{S}}^{*}(0):=\lim _{\substack{x \rightarrow 0 \\ x<0}} I_{P_{S}}^{*}(x)$, thus making $I_{P_{S}}^{*}$ continuous on $(-\infty, 0]$ while leaving $h$ unchanged. Hence, $f$ is continuous on $\left[0, \rho_{x}\right] \times\left[0, \rho_{s}\right] \times(0,+\infty)$ and $h(\lambda)=\min _{\left(q_{x}, q_{s}\right) \in\left[0, \rho_{x}\right] \times\left[0, \rho_{s}\right]} f\left(q_{x}, q_{s}, \lambda\right)$. If (ii), first note that

$$
f(0,0, \lambda)=\frac{1}{\alpha} I_{P_{S}}^{*}\left(-\frac{\rho_{s}}{2}\right)+\frac{\lambda}{12} \rho_{x}^{3}
$$

and

$$
f\left(q_{x}, q_{s}, \lambda\right) \geq \frac{1}{\alpha} I_{P_{S}}^{*}\left(\frac{q_{s}-\rho_{s}}{2}\right) \xrightarrow[q_{s} \rightarrow \rho_{s}]{q_{s}<0}+\infty .
$$

Then, for every positive $\bar{\lambda}$, there exists $\rho_{s}(\bar{\lambda}) \in\left(0, \rho_{s}\right)$ such that

- $\forall\left(q_{x}, q_{s}, \lambda\right) \in\left[0, \rho_{x}\right] \times\left[\rho_{s}(\bar{\lambda}), \rho_{s}\right] \times(0, \bar{\lambda}]: f\left(q_{x}, q_{s}, \lambda\right)>f(0,0, \lambda)$;
- $f$ is continuous on $\left[0, \rho_{x}\right] \times\left[0, \rho_{s}(\bar{\lambda})\right] \times(0,+\infty)$.

Thus, $\forall \lambda \in[0, \bar{\lambda}]: h(\lambda)=\min _{\left(q_{x}, q_{s}\right) \in\left[0, \rho_{x}\right] \times\left[0, \rho_{s}(\bar{\lambda})\right]} f\left(q_{x}, q_{s}, \lambda\right)$.
2) Fix $\bar{\lambda}>0$. The conclusion of step 1) is that there exists $\rho_{s}(\bar{\lambda}) \in\left(0, \rho_{s}\right]$ such that $\forall \lambda \in(0, \bar{\lambda}]$ :

$$
h(\lambda)=\min _{\left(q_{x}, q_{s}\right) \in\left[0, \rho_{x}\right] \times\left[0, \rho_{s}(\bar{\lambda})\right]} f\left(q_{x}, q_{s}, \lambda\right),
$$

where $f$ is defined in (5.103) with $I_{P_{S}}^{*}(0):=\lim _{x \rightarrow 0^{-}} I_{P_{s}}^{*}(x) \in[0,+\infty]$ and is continuous on $\left[0, \rho_{x}\right] \times\left[0, \rho_{s}(\bar{\lambda})\right] \times(0,+\infty)$. By Lemma $5.9, f$ admits a partial derivative with respect to $\lambda$ and for all $\left(q_{x}, q_{s}, \lambda\right) \in\left[0, \rho_{x}\right] \times\left[0, \rho_{s}(\bar{\lambda})\right] \times(0,+\infty)$ :

$$
\begin{equation*}
\left.\frac{\partial f}{\partial \lambda}\right|_{q_{x}, q_{s}, \lambda}=\frac{q_{x}^{2}}{2} \frac{\partial I_{\varphi}}{\partial r}\left(\frac{\lambda q_{x}^{2}}{2}, q_{s} ; \rho_{s}\right)+\frac{1}{12}\left(\rho_{x}-q_{x}\right)^{2}\left(\rho_{x}+2 q_{x}\right) . \tag{5.104}
\end{equation*}
$$

This partial derivative is continuous on $\left[0, \rho_{x}\right] \times\left[0, \rho_{s}(\bar{\lambda})\right] \times(0,+\infty)\left(\partial I_{\varphi} / \partial r\right.$ is given in the proof of Lemma 5.9 and its continuity is justified by domination assumptions). For all $\lambda \in(0, \bar{\lambda})$, define the following nonempty subset of $\left[0, \rho_{x}\right] \times\left[0, \rho_{s}(\bar{\lambda})\right]$ :

$$
\mathcal{Q}_{x, s}^{*}(\lambda):=\left\{\left(q_{x}^{*}, q_{s}^{*}\right) \in\left[0, \rho_{x}\right] \times\left[0, \rho_{s}\right]: f\left(q_{x}^{*}, q_{s}^{*}, \lambda\right)=h(\lambda)\right\} .
$$

By [68, Corollary 4], $h$ is differentiable at $\lambda \in(0, \bar{\lambda})$ if, and only if, the set

$$
\mathcal{F}(\lambda):=\left\{\left.\frac{\partial f}{\partial \lambda}\right|_{q_{x}^{*}, q_{s}^{*}, \lambda}:\left(q_{x}^{*}, q_{s}^{*}\right) \in \mathcal{Q}_{x, s}^{*}(\lambda)\right\}
$$

is a singleton, in which case $h^{\prime}(\lambda)=\left.\frac{\partial f}{\partial \lambda}\right|_{q_{x}^{*}, q_{s}^{*}, \lambda}$ for any $\left(q_{x}^{*}, q_{s}^{*}\right) \in \mathcal{Q}_{x, s}^{*}(\lambda)$. Note that $\mathcal{F}(\lambda)$ could be a singleton without $\mathcal{Q}_{x, s}^{*}(\lambda)$ being one. However, in the next and final step, we derive a simple expression for $\left.\frac{\partial f}{\partial \lambda}\right|_{q_{x}^{*}, q_{s}^{*}, \lambda}$ when $\left(q_{x}^{*}, q_{s}^{*}\right) \in \mathcal{Q}_{x, s}^{*}(\lambda)$ that shows that $\mathcal{F}(\lambda)$ is a singleton if, and only if, $\mathcal{Q}_{x, s}^{*}(\lambda)$ is one too.
3) Let $\lambda \in(0, \bar{\lambda})$ and $\left(q_{x}^{*}, q_{s}^{*}\right) \in \mathcal{Q}_{x, s}^{*}(\lambda)$. The function $q_{x} \mapsto f\left(q_{x}, q_{s}^{*}, \lambda\right)$ is differentiable on $\left[0, \rho_{x}\right]$ and $f\left(q_{x}^{*}, q_{s}^{*}, \lambda\right)=\min _{q_{x}^{*} \in\left[0, \rho_{x}\right]} f\left(q_{x}, q_{s}^{*}, \lambda\right)$. If $q_{x}^{*} \in\left(0, \rho_{x}\right)$ then it satisfies the optimality condition $\left.\frac{\partial f}{\partial q_{x}}\right|_{q_{x}^{*}, q_{s}^{*}, \lambda}=0$, i.e.,

$$
\begin{equation*}
q_{x}^{*} \frac{\partial I_{\varphi}}{\partial r}\left(\frac{\lambda\left(q_{x}^{*}\right)^{2}}{2}, q_{s}^{*} ; \rho_{s}\right)=\frac{q_{x}^{*}}{2}\left(\rho_{x}-q_{x}^{*}\right) . \tag{5.105}
\end{equation*}
$$

The identity (5.105) is trivially satisfied if $q_{x}^{*}=0$. If $q_{x}^{*}=\rho_{x}$ then the necessary optimality condition reads $\frac{\partial f}{\partial q_{x}}\left(\rho_{x}, q_{s}^{*}, \lambda\right)=\lambda \rho_{x} \frac{\partial \varphi_{\varphi}}{\partial r}\left(\frac{\lambda \rho_{x}^{2}}{2}, q_{s}^{*} ; \rho_{s}\right) \leq 0$. Besides, we show in the proof of Lemma 5.9 that $\frac{\partial I_{\varphi}}{\partial r} \geq 0$. Hence, if $q_{x}^{*}=\rho_{x}$, the condition (5.105) still has to be satisfied. Making use of the identity (5.105) in (5.104), we have $\forall\left(q_{x}^{*}, q_{s}^{*}\right) \in \mathcal{Q}_{x, s}^{*}(\lambda)$ :

$$
\begin{aligned}
\left.\frac{\partial f}{\partial \lambda}\right|_{q_{x}^{*}, q_{s}^{*}, \lambda} & =\frac{\left(q_{x}^{*}\right)^{2}}{2} \frac{\partial I_{\varphi}}{\partial r}\left(\frac{\lambda\left(q_{x}^{*}\right)^{2}}{2}, q_{s}^{*} ; \rho_{s}\right)+\frac{1}{12}\left(\rho_{x}-q_{x}^{*}\right)^{2}\left(\rho_{x}+2 q_{x}^{*}\right) \\
& =\frac{\left(q_{x}^{*}\right)^{2}}{4}\left(\rho_{x}-q_{x}^{*}\right)+\frac{1}{12}\left(\rho_{x}-q_{x}^{*}\right)^{2}\left(\rho_{x}+2 q_{x}^{*}\right) \\
& =\frac{\rho_{x}^{3}-\left(q_{x}^{*}\right)^{3}}{12} .
\end{aligned}
$$

It follows that $\mathcal{F}(\lambda)$ is a singleton if, and only if, $\mathcal{Q}_{x}^{*}(\lambda)$ is a singleton. We conclude that $h$ is differentiable if, and only if, $\mathcal{Q}_{x}^{*}(\lambda)$ is a singleton in which case, letting $\mathcal{Q}_{x}^{*}(\lambda)=\left\{q_{x}^{*}(\lambda)\right\}$, we have $h^{\prime}(\lambda)=\frac{\rho_{x}^{3}-\left(q_{x}^{*}(\lambda)\right)^{3}}{12}$.

## Part II

## Generalized linear models

## Entropy and mutual information in feedforward neural networks

6.1 Introduction

In this second part, we turn to generalized linear models (GLMs) as described by Model 1.4 that we now repeat.

Model 1.4 ( $L$-layer GLM). Let X be a $n$-dimensional random vector of interest and $L$ a natural number. For $\ell \in\{1, \ldots, L\}$, let $k_{\ell}$ be a natural number, $P_{A}^{(\ell)}$ a probability distribution on $\mathbb{R}^{k_{\ell}}$, and $\varphi_{\ell}: \mathbb{R} \times \mathbb{R}^{k_{\ell}} \rightarrow \mathbb{R}$ an activation function. For $\ell \in\{1, \ldots, L\}$, let $n_{\ell}$ be a positive integer and $\mathbf{W}^{(\ell)}$ a $n_{\ell} \times n_{\ell-1}$ matrix, with $n_{0}:=n$. Starting from $\mathbf{X}^{(0)}:=\mathbf{X}$, define recursively $\forall \ell \in\{1, \ldots, L\}$ :

$$
\mathbf{X}^{(\ell)}:=\varphi_{\ell}\left(\frac{\mathbf{W}^{(\ell)} \mathbf{X}^{(\ell-1)}}{\sqrt{n_{\ell-1}}}, \mathbf{A}^{(\ell)}\right)
$$

where $\mathbf{A}^{(\ell)}:=\left\{\mathbf{A}_{i}^{(\ell)}\right\}_{i=1}^{n_{\ell}} \stackrel{\text { i.i.d. }}{\sim} P_{A}^{(\ell)}$ and $\varphi_{\ell}$ is applied componentwise. In the $L$-layer GLM, we are given $\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(L)}$ and $n_{L}$ observations of the form

$$
Y_{i}:=X_{i}^{(L)}+\sqrt{\Delta} Z_{i}
$$

where $\mathbf{Z}:=\left\{Z_{i}\right\}_{i=1}^{n_{L}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ is an AWGN and $\Delta \geq 0$ is a noise variance parameter.

Phrased another way, $\mathbf{Y}$ is the output of a $L$-layer feedforward neural network with input $\mathbf{X}$, and $\mathbf{W}^{(\ell)} / \sqrt{n_{\ell-1}}$ are the weights leading to the $\ell^{\text {th }}$ hidden unit $\mathbf{X}^{(\ell)}$. Our main goal is to establish statistical limits for the estimation of $\mathbf{X}$ from the observations $\mathbf{Y}$ and the matrices $\mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \ldots, \mathbf{W}^{(\ell)}$. In this chapter, we consider a 2-layer GLM in the high-dimensional limit where the width of every layer scales linearly in the size $n$ of the input $\mathbf{X}$, that is, $n, n_{1}, n_{2} \rightarrow+\infty$ while $n_{1} / n \rightarrow \alpha_{1}>0, n_{2} / n_{1} \rightarrow \alpha_{2}>0$. We prove a RS formula for the conditional differential entropy $h\left(\mathbf{Y} \mid \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)$ normalized by $n_{0}$, from which a formula for the normalized conditional mutual information $n_{0}^{-1} I\left(\mathbf{X}^{(1)} ; \mathbf{Y} \mid \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)$ is easily
obtained. Like in the previous chapter, the proof demonstrates the modularity of the adaptive interpolation method.

Mutual information is in general difficult to compute [131]. Recently, owing to the development of deep learning, the computation of the mutual information between different layers of a neural network has recently attracted interest. The information bottleneck theory [132] applied to deep learning suggests that a successful training should maximize the mutual information between the labels and the learned hidden representations while minimizing the mutual information between the features and these same hidden representations [133], [134], hence mitigating overfitting. From a practical standpoint, this intuition has already led to new learning algorithms and regularizers [135, [136]. It was experimentally verified by 134 that this mechanism is at play when training with stochastic gradient descent (SGD) very small neural networks for which the authors can estimate the mutual information by binning. However, using the continuous entropy estimator of [137], [138] found that the nature of the nonlinear activation functions greatly affects the overall behavior of the mutual information during learning. In order to further investigate the evolution of the mutual information in large neural networks trained with SGD, 75 later proposed to use replicasymmetric (RS) ansätze. When the matrices of weights $\mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \ldots, \mathbf{W}^{(L)}$ are independent and rotationally invariant, there are RS ansätze for the asymptotic normalized mutual information between two layers that have been conjectured 27, [139]. These RS formulas are used in [75] to approximate the mutual information of large neural networks trained with SGD in such a way ${ }^{11}$ that the weight matrices remain rotationally invariant and close to being independent during training. Their findings agree with $[138$. Note however that the validity of these RS predictions is not proved yet. In this chapter, we show that the RS formula is exact when the feedforward neural network has two layers and in the restricted case of rotationally invariant weights matrices with i.i.d. Gaussian entries.

### 6.2 Two-layer generalized linear model

### 6.2.1 Problem setting

Let $n_{0}, n_{1}, n_{2}$ be positive integers and define the triplet $\boldsymbol{n}=\left(n_{0}, n_{1}, n_{2}\right)$. Let $k_{1}, k_{2}$ be nonnegative integers, $\varphi_{1}: \mathbb{R} \times \mathbb{R}^{k_{1}} \rightarrow \mathbb{R}, \varphi_{2}: \mathbb{R} \times \mathbb{R}^{k_{2}} \rightarrow \mathbb{R}$ two measurable functions, and $P_{A}^{(1)}, P_{A}^{(2)}$ two probability distributions on $\mathbb{R}^{k_{1}}$ and $\mathbb{R}^{k_{2}}$, respectively. For $i \in\{1,2\}$, it is understood that $\varphi_{i}$ acts component-wise on $(\mathbf{x}, \mathbf{A}) \in \mathbb{R}^{m} \times \mathbb{R}^{m \times k_{i}}$, that is, $\varphi_{i}(\mathbf{x}, \mathbf{A})$ is the $m$-dimensional vector with entries $\left[\varphi_{i}(\mathbf{x}, \mathbf{A})\right]_{\mu}:=\varphi_{i}\left(x_{\mu}, \mathbf{A}_{\mu}\right)$, where $\mathbf{A}_{\mu}$ denotes the $\mu^{\text {th }}$ row of $\mathbf{A}$.

[^17]Let $P_{X}$ be a probability distribution over $\mathbb{R}$ and $\mathbf{X} \in \mathbb{R}^{n_{0}}$ a random vector with components $\left(X_{i}\right)_{i=1}^{n_{0}} \stackrel{\text { i.i.d. }}{\sim} P_{X}$. For every $j \in\left\{1, \ldots, n_{1}\right\}$,

$$
\begin{equation*}
X_{j}^{(1)}:=\varphi\left(\frac{1}{\sqrt{n_{0}}} \sum_{i=1}^{n_{0}} W_{j i}^{(1)} X_{i}, \mathbf{A}_{j}^{(1)}\right), \tag{6.1}
\end{equation*}
$$

where $\left(\mathbf{A}_{j}^{(1)}\right)_{j=1}^{n_{1}} \stackrel{\text { i.i.d. }}{\sim} P_{A}^{(1)}$ and $\left(W_{j i}^{(1)}\right)_{j=1 \ldots n_{1}, i=1 \ldots n_{0}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. Equivalently,

$$
\begin{equation*}
\mathbf{X}^{(1)}:=\varphi\left(\frac{\mathbf{W}^{(1)} \mathbf{X}}{\sqrt{n_{0}}}, \mathbf{A}^{(1)}\right) \tag{6.2}
\end{equation*}
$$

where $\mathbf{W}^{(1)} \in \mathbb{R}^{n_{1} \times n_{0}}$ is the matrix with entries $W_{j i}^{(1)}$ and $\mathbf{A}^{(1)} \in \mathbb{R}^{n_{1} \times k_{1}}$ the matrix whose $j^{\text {th }}$ row is $\mathbf{A}_{j}^{(1)}$. We are given $n_{2}$ noisy observations

$$
\begin{equation*}
Y_{\mu}:=\varphi_{2}\left(\frac{1}{\sqrt{n_{1}}} \sum_{j=1}^{n_{1}} W_{\mu j}^{(2)} X_{j}^{(1)}, \mathbf{A}_{\mu}^{(2)}\right)+\sqrt{\Delta} Z_{\mu}, \quad 1 \leq \mu \leq n_{2} \tag{6.3}
\end{equation*}
$$

where $\left(\mathbf{A}_{\mu}^{(2)}\right)_{\mu=1}^{n_{2}} \stackrel{\text { i.i.d. }}{\sim} P_{A}^{(2)},\left(W_{\mu j}^{(2)}\right)_{\mu=1 \ldots n_{2}, j=1 \ldots n_{1}},\left(Z_{\mu}\right)_{\mu=1}^{n_{2}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$, and $\Delta$ is a positive real number. Equivalently,

$$
\begin{equation*}
\mathbf{Y}:=\varphi_{2}\left(\frac{\mathbf{W}^{(2)} \mathbf{X}^{(1)}}{\sqrt{n_{1}}}, \mathbf{A}^{(2)}\right)+\sqrt{\Delta} \mathbf{Z} \tag{6.4}
\end{equation*}
$$

where $\mathbf{W}^{(2)} \in \mathbb{R}^{n_{2} \times n_{1}}$ is the matrix with entries $W_{\mu j}^{(2)}, \mathbf{A}^{(2)} \in \mathbb{R}^{n_{2} \times k_{2}}$ the matrix whose $\mu^{\text {th }}$ row is $\mathbf{A}_{\mu}^{(2)}$, and $\mathbf{Z} \in \mathbb{R}^{n_{2}}$ the vector with entries $Z_{\mu}$. Note that

$$
\begin{equation*}
Y_{\mu} \sim P_{\mathrm{out}}\left(\cdot \left\lvert\, \frac{1}{\sqrt{n_{1}}} \sum_{j=1}^{n_{1}} W_{\mu j}^{(2)} X_{j}^{(1)}\right.\right), \tag{6.5}
\end{equation*}
$$

where $P_{\text {out }}$ is the conditional probability density function defined by

$$
\begin{equation*}
P_{\mathrm{out}}(y \mid x):=\int d P_{A}^{(2)}(\mathbf{a}) \frac{1}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}\left(y-\varphi_{2}(x, \mathbf{a})\right)^{2}} . \tag{6.6}
\end{equation*}
$$

Our analysis uses both representations (6.3) and (6.5).
The observations $Y_{\mu}$ are the noisy outputs of a two-layer feedforward neural network where $\mathbf{W}^{(1)}$ are the weights between the input and hidden layers, $\mathbf{W}^{(2)}$ the weights between the hidden and output layers, $\varphi_{1}\left(\cdot, \mathbf{A}_{j}^{(1)}\right)$ the stochastic activation functions of the hidden layer, and $\varphi_{2}\left(\cdot, \mathbf{A}_{\mu}^{(2)}\right)$ the stochastic activation functions of the output layer. The inference problem is to estimate $\mathbf{X}$ from the noisy outputs $\mathbf{Y}$ and knowledge of the weight matrices $\mathbf{W}^{(1)}, \mathbf{W}^{(2)}$.

### 6.2.2 Free entropy associated to the two-layer GLM

We denote $d P_{X}(\mathbf{x}):=\prod_{i=1}^{n_{0}} d P_{X}\left(x_{i}\right)$ and $P_{A}^{(1)}(\mathbf{a}):=\prod_{i=1}^{n_{1}} d P_{A}^{(1)}\left(\mathbf{a}_{j}\right)$ for $\mathbf{x} \in \mathbb{R}^{n_{0}}$ and $\mathbf{a} \in \mathbb{R}^{n_{1} \times k_{1}}$. Using Bayes' rule, the joint posterior distribution of $\left(\mathbf{X}, \mathbf{A}^{(1)}\right)$
given $\left(\mathbf{Y}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)$ is

$$
\begin{align*}
& d P\left(\mathbf{x}, \mathbf{a} \mid \mathbf{Y}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right) \\
& \quad:=\frac{d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a})}{\mathcal{Z}\left(\mathbf{Y}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)} \prod_{\mu=1}^{n_{2}} P_{\text {out }}\left(Y_{\mu} \left\lvert\,\left[\frac{\mathbf{W}^{(2)}}{\sqrt{n_{1}}} \varphi_{1}\left(\frac{\mathbf{W}^{(1)} \mathbf{x}}{\sqrt{n_{0}}}, \mathbf{a}\right)\right]_{\mu}\right.\right), \tag{6.7}
\end{align*}
$$

where the normalization factor is

$$
\begin{align*}
& \mathcal{Z}_{\boldsymbol{n}}\left(\mathbf{Y}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right) \\
& \quad:=\int d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a}) \prod_{\mu=1}^{n_{2}} P_{\mathrm{out}}\left(Y_{\mu} \left\lvert\,\left[\frac{\mathbf{W}^{(2)}}{\sqrt{n_{1}}} \varphi_{1}\left(\frac{\mathbf{W}^{(1)} \mathbf{x}}{\sqrt{n_{0}}}, \mathbf{a}\right)\right]_{\mu}\right.\right) . \tag{6.8}
\end{align*}
$$

The main quantity of interest in this chapter is the averaged free entropy associated to the posterior (6.7), that is,

$$
\begin{equation*}
f_{n}:=\frac{1}{n_{0}} \mathbb{E} \ln \mathcal{Z}_{n}\left(\mathbf{Y}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right) . \tag{6.9}
\end{equation*}
$$

Let us stress that $\mathbf{y} \mapsto \mathcal{Z}\left(\mathbf{y}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)$ is nothing else than the conditional density function of $\mathbf{Y}$ given $\mathbf{W}^{(1)}, \mathbf{W}^{(2)}$. Hence, $-n_{0} f_{n}$ is the conditional differential entropy of $\mathbf{Y}$ given $\mathbf{W}^{(1)}, \mathbf{W}^{(2)}$, i.e., $f_{n}=-h\left(\mathbf{Y} \mid \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right) / n_{0}$.

Our main result is a replica symmetric formula for the average free entropy $f_{n}$ in the high-dimensional regime where $n_{0}, n_{1}, n_{2} \rightarrow+\infty$ such that

$$
\frac{n_{2}}{n_{1}} \rightarrow \alpha_{2} \quad \text { and } \quad \frac{n_{1}}{n_{0}} \rightarrow \alpha_{1}
$$

where $\alpha_{2}, \alpha_{1}$ are positive real numbers. In this high-dimensional regime, the sampling rate is $n_{2} / n_{0} \rightarrow \alpha:=\alpha_{1} \alpha_{2}$.

### 6.2.3 Mutual informations of scalar Gaussian channels

The building blocks of the RS formula for the limit of the average free entropy are linked to the mutual information associated with simple scalar Gaussian channels. We now present these channels.

Define $Y^{(r)}:=\sqrt{r} X+Z$ where $X \sim P_{X}, Z \sim \mathcal{N}(0,1)$ are independent random variables and $r \geq 0$ plays the role of a signal-to-noise ratio. We assume that $P_{X}$ has a finite second moment $\rho_{0}:=\mathbb{E} X^{2}$, and we denote the mutual information between $X$ and $Y^{(r)}$ by $I_{P_{X}}(r):=I\left(X ; Y^{(r)}\right)$. The average free entropy associated to this channel is

$$
\begin{equation*}
\psi_{P_{X}}(r):=\mathbb{E} \ln \int d P_{X}(x) e^{\sqrt{r} Y^{(r)} x-\frac{r x^{2}}{2}} . \tag{6.10}
\end{equation*}
$$

The function $\psi_{P_{X}}: r \in \mathbb{R}^{+} \rightarrow \mathbb{R}$ is nondecreasing, convex, Lipschitz continuous with Lipschitz constant $\rho_{0} / 2$, and $I_{P_{X}}(r):=\frac{r \rho_{0}}{2}-\psi_{P_{X}}(r)$ (see Lemma 2.3).

Let $k$ be a nonnegative integer, $P_{A}$ a probability distribution on $\mathbb{R}^{k}$ and $\varphi: \mathbb{R} \times \mathbb{R}^{k} \mapsto \mathbb{R}$ a measurable function. Let $U, V, Z \sim \mathcal{N}(0,1)$ and $\mathbf{A} \sim P_{A}$ be independent random variables. We define

$$
Y^{(r, q ; \rho)}:=\sqrt{r} \varphi(\sqrt{\rho-q} U+\sqrt{q} V, \mathbf{A})+Z,
$$

where $r \geq 0$ plays the role of a signal-to-noise ratio, $\rho \geq 0$, and $q \in[0, \rho]$. We denote the conditional mutual information between $U$ and $Y^{(r, q ; \rho)}$ given $V$ by $I_{\varphi, P_{A}}(r, q ; \rho)=I\left(U ; Y^{(r, q ; \rho)} \mid V\right)$. The average free entropy associated to this channel is

$$
\begin{equation*}
\psi_{\varphi, P_{A}}(r, q ; \rho):=\mathbb{E}\left[\ln \left(\int d u \frac{e^{-\frac{u^{2}}{2}}}{\sqrt{2 \pi}} \int \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi}} e^{-\frac{1}{2}\left(Y^{(r, q ; \rho)}-\sqrt{r} \varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})\right)^{2}}\right)\right] . \tag{6.11}
\end{equation*}
$$

We can easily check that

$$
\begin{equation*}
I_{\varphi, P_{A}}(r, q ; \rho)=\Psi_{\varphi, P_{A}}(r, \rho ; \rho)-\Psi_{\varphi, P_{A}}(r, q ; \rho) . \tag{6.12}
\end{equation*}
$$

### 6.2.4 Replica symmetric formulas

The main result of this chapter is a variational formula for the asymptotic average free entropy (6.8) associated to the model defined in Subsection 6.2.1.

Theorem 6.1 (RS formula for the average free entropy of a two-layer GLM). Consider the statistical model defined in Subsection 6.2.1. Suppose that the following hypotheses hold ${ }^{2}$
(H1) The probability distribution $P_{X}$ has a bounded support, that is, if $X \sim P_{X}$ then there exists $S$ such that $|X| \leq S$ almost surely.
(H2) For $i \in\{1,2\}$, the function $\varphi_{i}: \mathbb{R} \times \mathbb{R}^{k_{i}} \rightarrow \mathbb{R}$ is bounded, twice differentiable with respect to its first coordinate, and these partial derivatives, denoted $\varphi_{i}^{\prime}$ and $\varphi_{i}^{\prime \prime}$, are bounded continuous on $\mathbb{R} \times \mathbb{R}^{k_{i}}$.
(H3) The entries of $\mathbf{W}^{(1)}, \mathbf{W}^{(2)}$ are i.i.d. with respect to $\mathcal{N}(0,1)$.
Define the (finite) second moments

$$
\rho_{0}:=\mathbb{E} X^{2}, \quad \rho_{1}:=\mathbb{E}\left[\varphi_{1}^{2}\left(\sqrt{\rho_{0}} N, \mathbf{A}^{(1)}\right)\right]
$$

where $X \sim P_{X}, N \sim \mathcal{N}(0,1), \mathbf{A}^{(1)} \sim P_{A}^{(1)}$ are independent random variables. Let $\alpha_{1}, \alpha_{2}$ be fixed positive numbers. Define the RS potential

$$
\begin{align*}
f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right):= & \frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right)+\psi_{P_{X}}\left(r_{0}\right) \\
& +\alpha_{1} \Psi_{\varphi_{1}, P_{A}^{(1)}}\left(r_{1}, q_{0} ; \rho_{0}\right)+\alpha_{1} \alpha_{2} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{1} ; \rho_{1}\right) \\
& -\frac{r_{0} q_{0}}{2}+\alpha_{1} \frac{r_{1}\left(\rho_{1}-q_{1}\right)}{2} . \tag{6.13}
\end{align*}
$$

[^18]Denote by $\boldsymbol{n} \rightarrow+\infty$ the high-dimensional limit where $n_{0}, n_{1}, n_{2} \rightarrow+\infty$ such that $n_{2} / n_{1} \rightarrow \alpha_{2}, n_{1} / n_{0} \rightarrow \alpha_{1}$. Then, the average free entropy $f_{n}=-\frac{h\left(\mathbf{Y} \mid \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)}{n_{0}}$ defined by (6.9) satisfies

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} f_{n}=\sup _{q_{1} \in\left[0, \rho_{1}\right]} \inf _{r_{1} \geq 0} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right) . \tag{6.14}
\end{equation*}
$$

Corollary 6.2 (Asymptotic normalized mutual information of a two-layer GLM). Under the assumptions of Theorem 6.1, the normalized conditional mutual information between $\mathbf{X}^{(1)}$ and $\mathbf{Y}$ given $\left(\mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)$ satisfies

$$
\lim _{n \rightarrow+\infty} \frac{I\left(\mathbf{X}^{(1)} ; \mathbf{Y} \mid \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)}{n_{0}}=\inf _{q_{1} \in\left[0, \rho_{1}\right]} \sup _{r_{1} \geq 0} \inf _{q_{0} \in\left[0, \rho_{0}\right]} \sup _{r_{0} \geq 0} i_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right) .
$$

where

$$
\begin{aligned}
i_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right):= & -\frac{\alpha_{1} \ln (2 \pi e)}{2}-\alpha_{1} \Psi_{\varphi_{1}, P_{A}^{(1)}}\left(r_{1}, \rho_{0} ; \rho_{0}\right)+I_{P_{X}}\left(r_{0}\right) \\
& +\alpha_{1} I_{\varphi_{1}, P_{A}^{(1)}}\left(r_{1}, q_{0} ; \rho_{0}\right)+\alpha_{1} \alpha_{2} I_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{1} ; \rho_{1}\right) \\
& -\frac{r_{0}\left(\rho_{0}-q_{0}\right)}{2}-\alpha_{1} \frac{r_{1}\left(\rho_{1}-q_{1}\right)}{2} .
\end{aligned}
$$

Proof. Note that

$$
\frac{I\left(\mathbf{X}^{(1)} ; \mathbf{Y} \mid \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)}{n_{0}}=\frac{1}{n_{0}} \mathbb{E}\left[\ln \prod_{\mu=1}^{n_{2}} P_{\text {out }}\left(Y_{\mu} \left\lvert\,\left[\frac{\mathbf{W}^{(2)} \mathbf{X}^{(1)}}{n_{1}}\right]_{\mu}\right.\right)\right]-f_{n}
$$

where

$$
\begin{aligned}
\frac{1}{n_{0}} \mathbb{E}\left[\ln \prod_{\mu=1}^{n_{2}} P_{\text {out }}\left(Y_{\mu} \left\lvert\,\left[\frac{\mathbf{W}^{(2)} \mathbf{X}^{(1)}}{n_{1}}\right]_{\mu}\right.\right)\right] & =\frac{1}{n_{0}} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\ln P_{\text {out }}\left(Y_{\mu} \left\lvert\,\left[\frac{\mathbf{W}^{(2)} \mathbf{X}^{(1)}}{n_{1}}\right]_{\mu}\right.\right)\right] \\
& =\frac{n_{2}}{n_{0}} \mathbb{E}\left[\ln P_{\text {out }}\left(Y_{1} \left\lvert\,\left[\frac{\mathbf{W}^{(2)} \mathbf{X}^{(1)}}{n_{1}}\right]_{1}\right.\right)\right] .
\end{aligned}
$$

Conditionally on $\mathbf{X}^{(1)}$, $\left[\mathbf{W}^{(2)} \mathbf{X}^{(1)} / n_{1}\right]_{1} \sim \mathcal{N}\left(0,\left\|\mathbf{X}^{(1)}\right\|^{2} / n_{1}\right)$ so

$$
\begin{aligned}
\mathbb{E}\left[\ln P_{\text {out }}\left(Y_{1} \left\lvert\,\left[\frac{\mathbf{W}^{(2)} \mathbf{X}^{(1)}}{n_{1}}\right]_{1}\right.\right)\right] & =\mathbb{E}\left[\mathbb{E}\left[\left.\ln P_{\text {out }}\left(Y_{1} \left\lvert\,\left[\frac{\mathbf{W}^{(2)} \mathbf{X}^{(1)}}{n_{1}}\right]_{1}\right.\right) \right\rvert\, \mathbf{X}^{(1)}\right]\right] \\
& =-\frac{\ln (\Delta)}{2}+\mathbb{E}\left[\Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}} ; \frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}\right)\right] \\
& =o_{n_{0}}(1)-\frac{\ln (\Delta)}{2}+\Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \rho_{1} ; \rho_{1}\right),
\end{aligned}
$$

where the last equality follows from the dominated convergence theorem and the strong law of large numbers. Hence,

$$
\lim _{n \rightarrow+\infty} \frac{I\left(\mathbf{X}^{(1)} ; \mathbf{Y} \mid \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)}{n_{0}}=-\alpha_{1} \alpha_{2} \frac{\ln (\Delta)}{2}+\alpha_{1} \alpha_{2} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \rho_{1} ; \rho_{1}\right)-\lim _{n \rightarrow+\infty} f_{n}
$$

Remark. If the probability distribution $P_{A}^{(1)}$ is deterministic then

$$
\frac{I\left(\mathbf{X}^{(1)} ; \mathbf{Y} \mid \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)}{n_{0}}=\frac{I\left(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right)}{n_{0}}
$$

and the potential $i_{\text {RS }}$ simply reads

$$
\begin{gathered}
i_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right):=I_{P_{X}}\left(r_{0}\right)+\alpha_{1} I_{\varphi_{1}, P_{A}^{(1)}}\left(r_{1}, q_{0} ; \rho_{0}\right)+\alpha_{1} \alpha_{2} I_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{1} ; \rho_{1}\right) \\
\\
-\frac{r_{0}\left(\rho_{0}-q_{0}\right)}{2}-\alpha_{1} \frac{r_{1}\left(\rho_{1}-q_{1}\right)}{2}
\end{gathered}
$$

### 6.3 Proof of the replica symmetric formula

This section is dedicated to the proof of Theorem 6.1. We consider the statistical model defined in Subsection 6.2.1 and use the definitions and notations introduced in Section 6.2,

### 6.3.1 Setting of the adaptive interpolation

For all $n_{0} \in \mathbb{N}^{*}$, define

$$
\begin{equation*}
\rho_{1}\left(n_{0}\right):=\mathbb{E}\left[\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}\right]=\frac{1}{n_{1}} \sum_{j=1}^{n_{1}} \mathbb{E}\left[\left(X_{j}^{(1)}\right)^{2}\right]=\mathbb{E}\left[\varphi_{1}^{2}\left(\frac{1}{n_{0}} \sum_{i=1}^{n_{0}} W_{1 i}^{(1)} X_{i}, \mathbf{A}_{1}^{(1)}\right)\right] . \tag{6.15}
\end{equation*}
$$

Under the hypotheses (H1) and (H2), we have (see Lemma 6.11 in Appendix 6.A)

$$
\begin{equation*}
\lim _{n_{0} \rightarrow+\infty} \rho_{1}\left(n_{0}\right)=\rho_{1} . \tag{6.16}
\end{equation*}
$$

We also define the bounded sequence of nonnegative real numbers

$$
r^{*}\left(n_{0}\right):=\left.2 \alpha_{2}\left(\frac{\partial \Psi_{\varphi_{2}, P_{A}^{(2)}}}{\partial q}\right)\right|_{r=\Delta^{-1}, q=\rho_{1}\left(n_{0}\right), \rho=\rho_{1}\left(n_{0}\right)}
$$

This sequence converges to

$$
r^{*}:=\left.2 \alpha_{2}\left(\frac{\partial \Psi_{\varphi_{2}, P_{A}^{(2)}}}{\partial q}\right)\right|_{r=\Delta^{-1}, q=\rho_{1}, \rho=\rho_{1}}
$$

when $n_{0} \rightarrow+\infty$. For the analysis it is useful to define

$$
\begin{equation*}
r_{\max }:=\sup _{n_{0} \geq 1} r^{*}\left(n_{0}\right) . \tag{6.17}
\end{equation*}
$$

Of course, $r_{\max } \geq r^{*}$. Let $\left(s_{n_{0}}\right)_{n_{0} \geq 1}$ be a decreasing sequence of real numbers in $(0,1 / 2]$ with limit $\lim _{n_{0} \rightarrow+\infty} s_{n_{0}}=0$, and $\mathcal{B}_{n_{0}}=\left[s_{n_{0}}, 2 s_{n_{0}}\right]^{2}$. For a fixed $n_{0} \geq 1$, we define for all $\epsilon:=\left(\epsilon_{1}, \epsilon_{2}\right) \in \mathcal{B}_{n_{0}}$ the interpolation functions

$$
R_{1}(\cdot, \epsilon): t \in[0,1] \mapsto \epsilon_{1}+\int_{0}^{t} r_{\epsilon}(v) d v
$$

and

$$
R_{2}(\cdot, \epsilon): t \in[0,1] \mapsto \epsilon_{2}+\int_{0}^{t} q_{\epsilon}(v) d v
$$

where $q_{\epsilon}:[0,1] \rightarrow\left[0, \rho_{1}\left(n_{0}\right)\right]$ and $r_{\epsilon}:[0,1] \rightarrow\left[0, r_{\max }\right]$ are two continuous functions. We specify more explicitly $q_{\epsilon}$ and $r_{\epsilon}$ later in the proof. In particular, we will need the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ to satisfy the following notion of regularity.

Definition (Regular interpolation paths). Let $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ be a family of functions from $[0,1]$ to $\left[0, \rho_{1}\left(n_{0}\right)\right]$, and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ a family of functions from $[0,1]$ to $\left[0, r_{\max }\right]$ We say that the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ are regular if, for all $t \in[0,1]$, the function

$$
\epsilon \mapsto\left(R_{1}(t, \epsilon), R_{2}(t, \epsilon)\right)
$$

is a $C^{1}$-diffeomorphism from $\mathcal{B}_{n_{0}}$ onto its image whose Jacobian determinant is greater than, or equal, to one.

Let $\mathbf{U}, \mathbf{V}$ be two $n_{2}$-dimensional random vectors with entries $U_{\mu}, V_{\mu} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ For $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$, we denote by $\mathbf{S}^{(t, \epsilon)}$ the $n_{2}$-dimensional random vector whose entries are given for all $\mu \in\left\{1, \ldots, n_{2}\right\}$ by

$$
\begin{equation*}
S_{\mu}^{(t, \epsilon)}:=\sqrt{\frac{1-t}{n_{1}}}\left[\mathbf{W}^{(2)} \mathbf{X}^{(1)}\right]_{\mu}+\sqrt{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t-R_{2}(t, \epsilon)} U_{\mu}+\sqrt{R_{2}(t, \epsilon)} V_{\mu} \tag{6.18}
\end{equation*}
$$

Consider the following observations coming from two types of channels,

$$
\begin{cases}Y_{\mu}^{(t, \epsilon)} \sim P_{\text {out }}\left(\cdot \mid S_{\mu}^{(t, \epsilon)}\right) & 1 \leq \mu \leq n_{2}  \tag{6.19}\\ \widetilde{Y}_{i}^{(t, \epsilon)}=\sqrt{R_{1}(t, \epsilon)} X_{i}^{(1)}+\widetilde{Z}_{i}, & 1 \leq i \leq n_{1}\end{cases}
$$

where $\left(\widetilde{Z}_{i}\right)_{i=1}^{n_{1}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. The two random vectors $\mathbf{Y}^{(t, \epsilon)}:=\left(Y_{\mu}^{(t, \epsilon)}\right)_{\mu=1}^{n_{2}}$ and $\tilde{\mathbf{Y}}^{(t, \epsilon)}:=\left(\widetilde{Y}_{i}^{(t, \epsilon)}\right)_{i=1}^{n_{1}}$ sum up these observations. The joint posterior distribution of $\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)$ is

$$
\begin{align*}
& d P\left(\mathbf{x}, \mathbf{a}, \mathbf{u} \mid \mathbf{Y}^{(t, \epsilon)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right) \\
& \quad:=\frac{d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a}) \mathcal{D} \mathbf{u}}{\mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)} e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{x}, \mathbf{a}, \mathbf{u} ; \mathbf{Y}^{(t, \epsilon)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)}, \tag{6.20}
\end{align*}
$$

where $\mathcal{D} \mathbf{u}:=d \mathbf{u} \mathrm{e}^{-\frac{\|\mathbf{u}\|^{2}}{2}} / \sqrt{2 \pi^{n}}, \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)$ is the normalization factor, and $\mathcal{H}_{t, \epsilon}$ is the interpolating Hamiltonian

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}\left(\mathbf{x}, \mathbf{a}, \mathbf{u} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right):= & -\sum_{\mu=1}^{n_{2}} \ln P_{\text {out }}\left(y_{\mu} \mid s_{\mu}^{(t, \epsilon)}\right) \\
& +\frac{1}{2} \sum_{i=1}^{n_{1}}\left(\widetilde{y}_{i}-\sqrt{R_{1}(t, \epsilon)} x_{i}^{(1)}\right)^{2} \tag{6.21}
\end{align*}
$$

with

$$
\begin{gathered}
x_{i}^{(1)}:=\varphi_{1}\left(\left[\frac{\mathbf{W}^{(1)} \mathbf{x}}{\sqrt{n_{0}}}\right]_{i}, \mathbf{a}_{i}\right), \\
s_{\mu}^{(t, \epsilon)}:=\sqrt{\frac{1-t}{n_{1}}}\left[\mathbf{W}^{(2)} \mathbf{x}^{(1)}\right]_{\mu}+\sqrt{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t-R_{2}(t, \epsilon)} u_{\mu}+\sqrt{R_{2}(t, \epsilon)} V_{\mu} .
\end{gathered}
$$

The average free entropy associated with the interpolation at $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ is

$$
\begin{align*}
f_{\boldsymbol{n}}(t, \epsilon) & :=\frac{\mathbb{E} \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \\
& =\frac{1}{n_{0}} \mathbb{E} \ln \int d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a}) \mathcal{D} \mathbf{u} e^{-\mathcal{H} t, \epsilon\left(\mathbf{x}, \mathbf{a}, \mathbf{u} ; \mathbf{Y}^{(t, \epsilon)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)} . \tag{6.22}
\end{align*}
$$

### 6.3.2 Variations of the average free entropy along the interpolation path

The proof of Theorem 6.1 is based on the analysis of the variations of (6.22) when $t$ varies from 0 to 1 . Given that $\lim _{n_{0} \rightarrow+\infty} s_{n_{0}}$, the pair $\epsilon=\left(\epsilon_{1}, \epsilon_{2}\right)$ is close to zero in the high-dimensional limit. Hence, $\epsilon$ should be seen as a "perturbation" that induces only a small change in the average free entropy. In particular, at $t=0$, we have for all $\epsilon \in \mathcal{B}_{n_{0}}$ :

$$
\begin{equation*}
f_{n}(0, \epsilon)=f_{n}-\frac{n_{1}}{2 n_{0}}+O\left(s_{n_{0}}\right) \tag{6.23}
\end{equation*}
$$

where $\left|O\left(s_{n_{0}}\right) / s_{n_{0}}\right|$ is bounded uniformly in $\boldsymbol{n}$ and $\epsilon \in\left[0,2 s_{n_{0}}\right]^{3}$, and $f_{\boldsymbol{n}}$ is the average free entropy defined by (6.9) whose high-dimensional limit we want to compute. We refer to Appendix $6 . \mathrm{B}$ for a proof of $(6.23)$.

At $t=1$, the observations $\mathbf{Y}^{(1, \epsilon)}$ do not involve $\overline{\mathbf{X}}^{(1)}$ anymore so the channel with observations $\mathbf{Y}^{(1, \epsilon)}$ and the one with observations $\widetilde{\mathbf{Y}}^{(1, \epsilon)}$ are decoupled. The average free entropy at $t=1$ is therefore the sum of the average free entropies associated with these decoupled channels. The high-dimensional limit of these two average free entropies is already known in the literature and we obtain the following result.

Lemma 6.3 (Interpolating average free entropy at $t=1$ ). Under the same assumptions than Theorem 6.1, the interpolating average free entropy at $t=1$ is $\forall \epsilon \in \mathcal{B}_{n_{0}}$ :

$$
\begin{aligned}
f_{\boldsymbol{n}}(1, \epsilon)=o_{\boldsymbol{n}}(1) & +\widetilde{f}_{\mathrm{RS}}\left(R_{1}(1, \epsilon) ; \rho_{0}\right)+\frac{\alpha_{1}}{2} \ln \left(2 \pi \Delta^{-\alpha_{2}}\right) \\
& +\alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)\right)
\end{aligned}
$$

where $O_{\boldsymbol{n}}(1)$ vanishes uniformly in $\epsilon \in \mathcal{B}_{n_{0}}$, and

$$
\widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right):=\sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} \psi_{P_{X}}\left(r_{0}\right)+\alpha_{1} \Psi_{\varphi_{1}, P_{A}^{(1)}}\left(r, q_{0} ; \rho_{0}\right)-\frac{r_{0} q_{0}}{2} .
$$

Proof. By the definition (6.22) of the interpolating average free entropy, we have

$$
\begin{aligned}
f_{\boldsymbol{n}}(1, \epsilon)= & \widetilde{f}_{n_{0}, n_{1}}\left(R_{1}(1, \epsilon)\right)+\frac{n_{1}}{n_{0}} \frac{\ln 2 \pi}{2} \\
& -\frac{n_{2}}{2 n_{0}} \ln \Delta+\frac{n_{2}}{n_{0}} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, R_{2}(1, \epsilon) ; \rho_{1}\left(n_{0}\right)+2 s_{n_{0}}\right)
\end{aligned}
$$

where

$$
\widetilde{f}_{n_{0}, n_{1}}\left(R_{1}(1, \epsilon)\right):=\frac{1}{n_{0}} \mathbb{E} \ln \int d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a}) \frac{e^{-\frac{1}{2}\left\|\mathbf{Y}^{(1, \epsilon)}-\sqrt{R_{1}(1, \epsilon)} \varphi_{1}\left(\frac{\mathbf{w}^{(1)} \mathbf{x}}{{\sqrt{n_{0}}}^{\mathbf{a}}} \mathbf{a}\right)\right\|^{2}}}{\sqrt{2 \pi}^{n_{1}}}
$$

and $\mathbf{Y}^{(1, \epsilon)}:=\sqrt{R_{1}(1, \epsilon)} \varphi_{1}\left(\mathbf{W}^{(1)} \mathbf{X} / \sqrt{n_{0}}, \mathbf{A}^{(1)}\right)+\widetilde{\mathbf{Z}}$. Note that $\widetilde{f}_{n_{0}, n_{1}}\left(R_{1}(1, \epsilon)\right)$ is the average free entropy associated with a one-layer GLM, and its high-dimensional limit is given by [29, Theorem 1]. We have $\forall \epsilon \in \mathcal{B}_{n_{0}}$ :

$$
\widetilde{f}_{n_{0}, n_{1}}\left(R_{1}(1, \epsilon)\right)=o_{n_{0}}(1)+\widetilde{f}_{\mathrm{RS}}\left(R_{1}(1, \epsilon) ; \rho_{0}\right)
$$

where $O_{n_{0}}(1)$ vanishes uniformly in $\epsilon \in \mathcal{B}_{n_{0}}$, and

$$
\widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right):=\sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} \psi_{P_{X}}\left(r_{0}\right)+\alpha_{1} \Psi_{\varphi_{1}, P_{A}^{(1)}}\left(r, q_{0} ; \rho_{0}\right)-\frac{r_{0} q_{0}}{2}
$$

The last summand appearing on the right-hand side of (6.3.2) satisfies

$$
\begin{align*}
& \frac{n_{2}}{n_{0}} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, R_{2}(1, \epsilon) ; \rho_{1}\left(n_{0}\right)+2 s_{n_{0}}\right) \\
& \quad=\frac{n_{2}}{n_{0}} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)+2 s_{n_{0}}\right)+O\left(s_{n_{0}}\right) \\
& \quad=\frac{n_{2}}{n_{0}} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)\right)+O\left(s_{n_{0}}\right) \tag{6.24}
\end{align*}
$$

The first equality is because $R_{2}(1, \epsilon):=\epsilon_{2}+\int_{0}^{1} q_{\epsilon}(t) d t$ and

$$
q \mapsto \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q ; \rho_{1}\left(n_{0}\right)+2 s_{n_{0}}\right)
$$

is Lipschitz continuous on $\left[0, \rho_{1}\left(n_{0}\right)+2 s_{n_{0}}\right]$. The second equality is because $\int_{0}^{1} q_{\epsilon}(t) d t \leq \rho_{1}\left(n_{0}\right) \leq \rho_{1}\left(n_{0}\right)+2 s_{n_{0}} \leq \rho_{1}\left(n_{0}\right)+1$ and

$$
\rho \mapsto \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho\right)
$$

is Lipschitz continuous on $\left[\int_{0}^{1} q_{\epsilon}(t) d t, \rho_{1}\left(n_{0}\right)+1\right]$. The quantity $O\left(s_{n_{0}}\right)$ in (6.24) is such that $\left|O\left(s_{n_{0}}\right) / s_{n_{0}}\right|$ is uniformly bounded in $\boldsymbol{n}$ and $\epsilon \in \mathcal{B}_{n_{0}}$. By definition,

$$
\begin{aligned}
& \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q, \rho\right) \\
& \quad:=\mathbb{E}\left[\ln \int \mathcal{D} u d P_{A}^{(2)}(\mathbf{a}) \frac{e^{-\frac{1}{2 \Delta}\left(\varphi_{2}(\sqrt{\rho-q} U+\sqrt{q} V, \mathbf{A})-\varphi_{2}(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})+\sqrt{\Delta} Z\right)^{2}}}{\sqrt{2 \pi}}\right]
\end{aligned}
$$

where $\mathcal{D} u:=d u e^{-\frac{u^{2}}{2}} / \sqrt{2 \pi}$ and $U, V, Z \sim \mathcal{N}(0,1), \mathbf{A} \sim P_{A}^{(2)}$ are independent random variables. It directly follows that

$$
-\frac{\ln (2 \pi)}{2}+\mathbb{E}\left[\ln e^{-\frac{1}{2}\left(2 \frac{\left\|\varphi_{2}\right\| \infty}{\sqrt{\Delta}}+|Z|\right)^{2}}\right] \leq \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q, \rho\right) \leq-\frac{\ln (2 \pi)}{2}
$$

hence

$$
-\frac{\ln (2 \pi)}{2}-\left(4 \frac{\left\|\varphi_{2}\right\|_{\infty}^{2}}{\Delta}+1\right) \leq \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q, \rho\right) \leq-\frac{\ln (2 \pi)}{2} .
$$

The absolute value of $\Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q, \rho\right)$ is thus uniformly bounded for $\rho \geq 0$, $q \in[0, \rho]$. Combining this uniform bound with (6.24) yields

$$
\begin{aligned}
& \frac{n_{2}}{n_{0}} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, R_{2}(1, \epsilon) ; \rho_{1}\left(n_{0}\right)+2 s_{n_{0}}\right) \\
& \quad=\alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)\right)+o_{n_{0}}(1)
\end{aligned}
$$

where $O_{n_{0}}(1)$ vanishes uniformly $\epsilon \in \mathcal{B}_{n_{0}}$.
In a nutshell, we interpolate from the original problem at $t=0$ to two decoupled and analytically tractable problems at $t=1$. At $t=0$, the interpolating problem essentially reduces to the two-layer GLM defined in Subsection 6.2.1. At $t=1$, it reduces to a one-layer GLM whose asymptotic average free entropy is given in [29|), plus $n_{2}$ decoupled scalar Gaussian channels whose average free entropies are equal and given by the function $\Psi_{\varphi_{2}, P_{A}^{(2)}}$.

By the fundamental theorem of calculus, making use of (6.23) and Lemma 6.3, we have $\forall \epsilon \in \mathcal{B}_{n_{0}}$ :

$$
\begin{align*}
f_{\boldsymbol{n}}= & o_{\boldsymbol{n}}(1)+f_{\boldsymbol{n}}(0, \epsilon)+\frac{\alpha_{1}}{2} \\
= & o_{\boldsymbol{n}}(1)+f_{\boldsymbol{n}}(1, \epsilon)-\int_{0}^{1} f_{\boldsymbol{n}}^{\prime}(t, \epsilon) d t+\frac{\alpha_{1}}{2} \\
= & o_{\boldsymbol{n}}(1)+\widetilde{f}_{\mathrm{RS}}\left(R_{1}(1, \epsilon) ; \rho_{0}\right)+\frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right) \\
& \quad+\alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)\right)-\int_{0}^{1} d t f_{\boldsymbol{n}}^{\prime}(t, \epsilon), \tag{6.25}
\end{align*}
$$

where $O_{\boldsymbol{n}}(1)$ vanishes uniformly in $\epsilon \in \mathcal{B}_{n_{0}}$ and we denote by $f_{\boldsymbol{n}}^{\prime}(\cdot, \epsilon)$ the derivative of $f_{n}(\cdot, \epsilon)$.

The next step is to compute $f_{n}^{\prime}(\cdot, \epsilon)$. We denote by ( $\mathbf{x}, \mathbf{a}, \mathbf{u}$ ) a triplet sampled from the joint posterior distribution 6.20 . The angular brackets $\langle-\rangle_{\boldsymbol{n}, t, \epsilon}$ denote the expectation with respect to this posterior,

$$
\begin{equation*}
\langle g(\mathbf{x}, \mathbf{a}, \mathbf{u})\rangle_{\boldsymbol{n}, t, \epsilon}:=\int g(\mathbf{x}, \mathbf{a}, \mathbf{u}) d P\left(\mathbf{x}, \mathbf{a}, \mathbf{u} \mid \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right) \tag{6.26}
\end{equation*}
$$

We call overlap the normalized inner product between $\mathbf{x}^{(1)}:=\varphi_{1}\left(\mathbf{W}^{(1)} \mathbf{x} / \sqrt{n_{0}}, \mathbf{a}\right)$ and $\mathbf{X}^{(1)}$, that is,

$$
\begin{equation*}
Q:=\frac{\left(\mathbf{x}^{(1)}\right)^{\top} \mathbf{X}^{(1)}}{n_{1}}=\frac{1}{n_{1}} \sum_{i=1}^{n_{1}} x_{i}^{(1)} X_{i}^{(1)} . \tag{6.27}
\end{equation*}
$$

We compute the derivative $f_{n}^{\prime}(\cdot, \epsilon)$ in Appendix 6.C.
Proposition 6.4 (Derivative of the interpolating average free entropy). Assume that (H1), (H2), (H3) hold and $n_{0}, n_{1}, n_{2} \rightarrow+\infty$ such that $n_{2} / n_{1} \rightarrow \alpha_{2}, n_{1} / n_{0} \rightarrow \alpha_{1}$. Let $f_{n}(t, \epsilon)$ be the interpolating average free entropy defined by (6.22). Then, the derivative of $f_{\boldsymbol{n}}(\cdot, \epsilon)$, denoted by $f_{\boldsymbol{n}}^{\prime}(\cdot, \epsilon)$, satisfies $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\begin{aligned}
& f_{\boldsymbol{n}}^{\prime}(t, \epsilon)=-\frac{1}{2} \frac{n_{1}}{n_{0}} \mathbb{E}\left\langle\left(\frac{1}{n_{1}} \sum_{\mu=1}^{n_{2}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-r_{\epsilon}(t)\right)\left(Q-q_{\epsilon}(t)\right)\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
&+\frac{n_{1}}{n_{0}} \frac{r_{\epsilon}(t)}{2}\left(q_{\epsilon}(t)-\rho_{1}\left(n_{0}\right)\right)+o_{\boldsymbol{n}}(1),
\end{aligned}
$$

where $\ell_{y}^{\prime}(\cdot)$ is the derivative of $\ell_{y}: x \mapsto \ln P_{\text {out }}(y \mid x), o_{n}(1)$ is a quantity that vanishes uniformly in $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ when $n_{0} \rightarrow+\infty$, and $Q:=\frac{\left(\mathbf{x}^{(1)}\right)^{\top} \mathbf{x}^{(1)}}{n_{1}}$.

### 6.3.3 Sum-rule of the adaptive interpolation

The overlap $Q$ defined by (6.27) naturally appears in the derivative of the average free entropy. An important property of the overlap is that it concentrates around its expectation provided that we integrate over the perturbations $\epsilon \in \mathcal{B}_{n_{0}}$.

Proposition 6.5 (Overlap concentration). Let $s_{n_{0}}:=\frac{1}{2} n_{0}^{-1 / 16}$. Assume that (H1), (H2), (H3) hold and the families $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}},\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ are regular. There exists a constant $C$ only depending on $\varphi_{1}, \varphi_{2}, \alpha_{1}, \alpha_{2}, \Delta$ and $S$ such that

$$
\begin{equation*}
\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \leq \frac{C}{n_{0}^{1 / 8}} \tag{6.29}
\end{equation*}
$$

Proof. It follows directly from Proposition 6.22 combined with the upper bound 6.112), see Appendix 6.E

In order to prove Theorem 6.1, we need to get rid of the first term on the right-hand side of $(6.28)$. We can achieve the latter by choosing $q_{\epsilon}$ in such a way that $q_{\epsilon}(t)=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}$. Indeed, by Proposition 6.5, $Q$ concentrates around its expectation. However, $\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}$ is a function of $t$ and $R(t, \epsilon)$, so $q_{\epsilon}(t)=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}$ is an ordinary differential equation (ODE) of order 1 . We address in details the problem of picking $q_{\epsilon}$ such that $q_{\epsilon}(t)=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}$ in the next section. For now, we assume that $q_{\epsilon}(t)=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}$ and state the sum-rule of the adaptive interpolation.

Proposition 6.6. Let $s_{n_{0}}:=\frac{1}{2} n_{0}^{-1 / 16}$. Assume that (H1), (H2), (H3) hold and $n_{0}, n_{1}, n_{2} \rightarrow+\infty$ such that $n_{2} / n_{1} \rightarrow \alpha_{2}, n_{1} / n_{0} \rightarrow \alpha_{1}$. Further assume that the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}},\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ are regular and $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :
$q_{\epsilon}(t)=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}$. Then,

$$
\begin{aligned}
& f_{\boldsymbol{n}}=o_{\boldsymbol{n}}(1)+\frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right) \\
& \quad+\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}}\left\{\widetilde{f}_{\mathrm{RS}}\left(R_{1}(1, \epsilon) ; \rho_{0}\right)+\alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)\right)\right. \\
& \\
& \left.\quad+\int_{0}^{1} d t \frac{\alpha_{1}}{2} r_{\epsilon}(t)\left(\rho_{1}-q_{\epsilon}(t)\right)\right\}
\end{aligned}
$$

where $o_{\boldsymbol{n}}(1)$ is a quantity that vanishes when $\boldsymbol{n} \rightarrow+\infty$.

Proof. Integrating both sides of (6.25) over $\epsilon \in \mathcal{B}_{n_{0}}$, and making use of the formula for $f_{n}^{\prime}(t, \epsilon)$ in Proposition 6.4 yields

$$
\begin{aligned}
& f_{\boldsymbol{n}}= \int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} f_{\boldsymbol{n}} \\
&=o_{\boldsymbol{n}}(1)+\frac{n_{1}}{2 n_{0}} \mathcal{R}_{\boldsymbol{n}}+\frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right) \\
&+\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}}\left\{\widetilde{f}_{\mathrm{RS}}\left(R_{1}(1, \epsilon) ; \rho_{0}\right)+\alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)\right)\right. \\
&\left.\quad+\int_{0}^{1} d t \frac{\alpha_{1}}{2} r_{\epsilon}(t)\left(\rho_{1}-q_{\epsilon}(t)\right)\right\},
\end{aligned}
$$

where

$$
\mathcal{R}_{n}:=\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(\frac{1}{n_{1}} \sum_{\mu=1}^{n_{2}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-r_{\epsilon}(t)\right)\left(Q-q_{\epsilon}(t)\right)\right\rangle_{n, t, \epsilon}
$$

To conclude the proof we need to check that $\mathcal{R}_{\boldsymbol{n}}=o_{\boldsymbol{n}}(1)$. By assumption $q_{\epsilon}(t)=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}$ so

$$
\mathcal{R}_{n}^{2}=\left|\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(\frac{1}{n_{1}} \sum_{\mu=1}^{n_{2}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right)\left(Q-q_{\epsilon}(t)\right)\right\rangle_{n, t, \epsilon}\right|^{2},
$$

and, by Cauchy-Schwarz inequality,

$$
\begin{align*}
& \mathcal{R}_{n}^{2} \leq \int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(\frac{1}{n_{1}} \sum_{\mu=1}^{n_{2}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
& \cdot \int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \tag{6.30}
\end{align*}
$$

The integrand of the first integral on the right-hand side of (6.30) satisfies

$$
\begin{aligned}
& \mathbb{E}\left\langle\left(\frac{1}{n_{1}} \sum_{\mu=1}^{n_{2}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
& \quad \leq \frac{n_{2}}{n_{1}^{2}} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right)^{2}\left\langle\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
& \quad=\frac{n_{2}}{n_{1}^{2}} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\left\langle\ell_{Y_{\mu}^{\prime}}^{\prime}(t, \epsilon)\left(s_{\mu}^{(t, \epsilon)}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}^{2}\right] \\
& \quad \leq \frac{n_{2}}{n_{1}^{2}} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left\langle\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)^{4}\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
& \quad=\frac{n_{2}}{n_{1}^{2}} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right)^{4}\right]
\end{aligned}
$$

where the two inequalities follow from Jensen's inequality, and the two equalities from the Nishimori identity (see Lemma 2.1 in Chapter 22). By the inequality (6.53) in Appendix 6.B. $\mathbb{E}\left[\ell_{\left.Y_{\mu}^{\prime}, \epsilon\right)}^{\prime(t,)}\left(S_{\mu}^{(t, \epsilon)}\right)^{4}\right] \leq 8\left\|\varphi_{2}^{\prime} / \sqrt{\Delta}\right\|_{\infty}^{4}\left(3+16\left\|\varphi_{2} / \sqrt{\Delta}\right\|_{\infty}^{4}\right)$. Putting everything together, we obtain

$$
\left|\mathcal{R}_{\boldsymbol{n}}\right| \leq \frac{2 n_{2}}{n_{1}}\left\|\frac{\varphi_{2}^{\prime}}{\sqrt{\Delta}}\right\|_{\infty}^{2} \sqrt{6+32\left\|\frac{\varphi_{2}}{\sqrt{\Delta}}\right\|_{\infty}^{4}} \sqrt{\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}} .
$$

By Proposition 6.5, we conclude that $\mathcal{R}_{\boldsymbol{n}}=o_{\boldsymbol{n}}(1)$.

### 6.3.4 Interpolation functions solutions to ODEs

In the next subsection, we prove matching lower and upper bounds on the asymptotic average free entropy $f_{n}$. Each bound is obtained by choosing $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ so that $R(\cdot, \epsilon)$ is the solution to a first-order ODE with initial value $R(0, \epsilon)=\epsilon$. In this subsection we define these ODEs and prove important properties of their solutions; properties that will allow us to apply Proposition 6.6

Let $\left(s_{n_{0}}\right)_{n_{0} \geq 1}$ be a decreasing sequence of real numbers in $(0,1 / 2$ ]. Let $\mathbf{U}$, $\mathbf{V}$ be $n_{2}$-dimensional random vectors with entries $U_{\mu}, V_{\mu} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. For fixed $t \in[0,1]$ and $R=\left(R_{1}, R_{2}\right) \in[0,+\infty) \times\left[0,2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t\right]$, we denote by $\mathbf{S}^{\left(R_{2}\right)}$ the $n_{2}$-dimensional random vector whose entries are given for all $\mu \in\left\{1, \ldots, n_{2}\right\}$ by

$$
S_{\mu}^{\left(t, R_{2}\right)}:=\sqrt{\frac{1-t}{n_{1}}}\left[\mathbf{W}^{(2)} \mathbf{X}^{(1)}\right]_{\mu}+\sqrt{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t-R_{2}} U_{\mu}+\sqrt{R_{2}} V_{\mu}
$$

and we consider the observations

$$
\begin{cases}Y_{\mu}^{\left(t, R_{2}\right)} & \sim P_{\text {out }}\left(\cdot \mid S_{\mu}^{\left(t, R_{2}\right)}\right), \quad 1 \leq \mu \leq n_{2} \\ \widetilde{Y}_{i}^{\left(t, R_{1}\right)} & =\sqrt{R_{1}} X_{i}^{(1)}+\widetilde{Z}_{i}, \quad 1 \leq i \leq n_{1}\end{cases}
$$

where $\left(\widetilde{Z}_{i}\right)_{i=1}^{n_{1}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. The two random vectors $\mathbf{Y}^{\left(t, R_{2}\right)}:=\left(Y_{\mu}^{\left(t, R_{2}\right)}\right)_{\mu=1}^{n_{2}}$ and $\widetilde{\mathbf{Y}}^{\left(t, R_{1}\right)}:=\left(\widetilde{Y}_{i}^{\left(t, R_{1}\right)}\right)_{i=1}^{n_{1}}$ sum up these observations. The joint posterior distribution of $\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{\left(t, R_{2}\right)}, \widetilde{\mathbf{Y}}^{\left(t, R_{1}\right)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)$ is

$$
\begin{aligned}
& d P\left(\mathbf{x}, \mathbf{a}, \mathbf{u} \mid \mathbf{Y}^{\left(t, R_{2}\right)}, \widetilde{\mathbf{Y}}^{\left(t, R_{1}\right)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right) \\
\propto & \left.\prod_{i=1}^{n_{1}} d P_{X}\left(x_{i}\right) d P_{A}^{(1)}\left(\mathbf{a}_{i}\right) e^{-\frac{1}{2}\left(\tilde{Y}_{i}^{\left(t, R_{1}\right)}-\sqrt{R_{1}} x_{i}^{(1)}\right.}\right)^{2} \cdot \prod_{\mu=1}^{n_{2}} \frac{d u_{\mu}}{\sqrt{2 \pi}} e^{-\frac{u_{\mu}^{2}}{2}} P_{\text {out }}\left(Y_{\mu}^{\left(t, R_{2}\right)} \mid s_{\mu}^{\left(t, R_{2}\right)}\right),
\end{aligned}
$$

with $x_{i}^{(1)}:=\varphi_{1}\left(\left[\mathbf{W}^{(1)} \mathbf{x} / \sqrt{n_{0}}\right]_{i}, \mathbf{a}_{i}\right)$ and

$$
s_{\mu}^{\left(t, R_{2}\right)}:=\sqrt{\frac{1-t}{n_{1}}}\left[\mathbf{W}^{(2)} \mathbf{x}^{(1)}\right]_{\mu}+\sqrt{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t-R_{2}} u_{\mu}+\sqrt{R_{2}} V_{\mu}
$$

We denote by the angular brackets $\langle-\rangle_{\boldsymbol{n}, t, R}$ the expectation with respect to this posterior distribution and define

$$
\begin{aligned}
& F_{1}^{(\boldsymbol{n})}(t, R):=\left.2 \alpha_{2}\left(\frac{\partial \Psi_{\varphi_{2}, P_{A}^{(2)}}}{\partial q}\right)\right|_{r=\Delta^{-1}, q=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, R}, \rho=\rho_{1}\left(n_{0}\right)} \\
& F_{2}^{(\boldsymbol{n})}(t, R):=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, R}
\end{aligned}
$$

where $Q:=\frac{\left(\mathbf{x}^{(1)}\right)^{\top} \mathbf{x}^{(1)}}{n_{1}}$. Let $r \in\left[0, r_{\max }\right]$. We consider the first-order ODEs

$$
\begin{equation*}
\frac{d y}{d t}=\left(r, F_{2}^{(\boldsymbol{n})}(t, y)\right) \tag{6.31}
\end{equation*}
$$

and

$$
\begin{equation*}
\frac{d y}{d t}=\left(F_{1}^{(\boldsymbol{n})}(t, y), F_{2}^{(\boldsymbol{n})}(t, y)\right) \tag{6.32}
\end{equation*}
$$

Proposition 6.7. Suppose that (H1), (H2), (H3) hold. For every $\epsilon \in \mathcal{B}_{n_{0}}$, there exists a unique global solution $R(\cdot, \epsilon):[0,1] \rightarrow[0,+\infty)^{2}$ to the initial value problem

$$
\frac{d y}{d t}=\left(F_{1}^{(\boldsymbol{n})}(t, y), F_{2}^{(\boldsymbol{n})}(t, y)\right), \quad y(0)=\epsilon
$$

$R(\cdot, \epsilon)$ is continuously differentiable and the image of its derivative $R^{\prime}(\cdot, \epsilon)$ satisfies

$$
R^{\prime}([0,1], \epsilon) \subseteq\left[0, r_{\max }\right] \times\left[0, \rho_{1}\left(n_{0}\right)\right]
$$

where $r_{\max } \geq 0$ is defined by 6.17). Besides, for all $t \in[0,1], R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $\mathcal{B}_{n_{0}}$ onto its image whose Jacobian determinant is greater than, or equal to, one, i.e., $\forall \epsilon \in \mathcal{B}_{n_{0}}$ :

$$
\operatorname{det} J_{R(t, \cdot)}(\epsilon) \geq 1
$$

where $J_{R(t,)}$ denotes the Jacobian matrix of $R(t, \cdot)$. Finally, the same statement holds true if, for a fixed $r \in\left[0, r_{\max }\right]$, we instead consider the initial value problem

$$
\frac{d y}{d t}=\left(r, F_{2}^{(n)}(t, y)\right), \quad y(0)=\epsilon
$$

Proof. We only give the proof for the ODE $d y / d t=\left(F_{1}^{(n)}(t, y), F_{2}^{(n)}(t, y)\right)$ since the one for the $\mathrm{ODE} d y / d t=\left(r, F_{2}^{(n)}(t, y)\right)$ is simpler and follows the same arguments.

By Jensen's inequality and the Nishimori identity (see Lemma 2.1),
$\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, R}:=\frac{\mathbb{E}\left[\left\langle\mathbf{x}^{(1)}\right\rangle_{\boldsymbol{n}, t, R}^{\top} \mathbf{X}^{(1)}\right]}{n_{1}}=\frac{\mathbb{E}\left\|\left\langle\mathbf{x}^{(1)}\right\rangle_{\boldsymbol{n}, t, R}\right\|^{2}}{n_{1}} \leq \frac{\mathbb{E}\left\langle\left\|\mathbf{x}^{(1)}\right\|^{2}\right\rangle_{\boldsymbol{n}, t, R}}{n_{1}}=\frac{\mathbb{E}\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}$,
hence $\mathbb{E}\langle Q\rangle_{n, t, R} \in\left[0, \rho_{1}\left(n_{0}\right)\right]$. The function $q \mapsto \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q, \rho_{1}\left(n_{0}\right)\right)$ is continuously twice differentiable, convex and nondecreasing on $\left[0, \rho_{1}\left(n_{0}\right)\right]$. Therefore, $\left.q \mapsto 2 \alpha_{2}\left(\partial \Psi_{\varphi_{2}, P_{A}^{(2)}} / \partial q\right)\right|_{\Delta^{-1}, q, \rho_{1}\left(n_{0}\right)}$ is nonnegative and nondecreasing on $\left[0, \rho_{1}\left(n_{0}\right)\right]$, which implies that

$$
0 \leq\left. 2 \alpha_{2}\left(\frac{\partial \Psi_{\varphi_{2}, P_{A}^{(2)}}}{\partial q}\right)\right|_{\Delta^{-1}, q, \rho_{1}\left(n_{0}\right)} \leq\left. 2 \alpha_{2}\left(\frac{\partial \Psi_{\varphi_{2}, P_{A}^{(2)}}}{\partial q}\right)\right|_{\Delta^{-1}, q=\rho_{1}\left(n_{0}\right), \rho_{1}\left(n_{0}\right)} \leq r_{\max }
$$

We have thus shown that the function $F^{(\boldsymbol{n})}:(t, R) \mapsto\left(F_{1}^{(\boldsymbol{n})}(t, R), F_{2}^{(\boldsymbol{n})}(t, R)\right)$ is well-defined on

$$
\mathcal{D}_{n_{0}}:=\left\{\left(t, R_{1}, R_{2}\right) \in[0,1] \times[0,+\infty)^{2}: R_{2} \leq 2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t\right\}
$$

and that $F^{(\boldsymbol{n})}\left(\mathcal{D}_{n_{0}}\right) \subseteq\left[0, r_{\max }\right] \times\left[0, \rho_{1}\left(n_{0}\right)\right]$.
To invoke the Picard-Lindelöf theorem [104, Theorem 1.1], we have to check that $F^{(\boldsymbol{n})}$ is continuous in $t$ and uniformly Lipschitz continuous in $R$ (meaning that the Lipschitz constant is independent of $t$ ). We can show that $F^{(\boldsymbol{n})}$ is continuous on $\mathcal{D}_{n_{0}}$ and that, for all $t \in[0,1], F^{(\boldsymbol{n})}(t, \cdot)$ is differentiable on $(0,+\infty) \times\left(0,2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t\right)$ thanks to the standard theorems of continuity and differentiation under the integral sign. The domination hypotheses are indeed verified because (H1), (H2) hold. To check the uniform Lipschitzness, we show that the Jacobian matrix of $F^{(\boldsymbol{n})}(t, \cdot)$, that we denote by $J_{F^{(n)}(t,)}(R)$, is uniformly bounded in $(t, R)$. For all $\left(R_{1}, R_{2}\right) \in(0,+\infty) \times\left(0,2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t\right)$ :

$$
J_{F(t, \cdot)}(R)=\left[\begin{array}{cc}
c(t, R) & c(t, R)  \tag{6.33}\\
1 & 1
\end{array}\right]\left[\begin{array}{cc}
\left.\frac{\partial F_{2}^{(n)}}{\partial R_{1}}\right|_{t, R} & 0 \\
0 & \left.\frac{\partial F_{2}^{(n)}}{\partial R_{2}}\right|_{t, R}
\end{array}\right]
$$

where

$$
\begin{align*}
c(t, R) & :=\left.2 \alpha_{2}\left(\frac{\partial^{2} \Psi_{\varphi_{2}, P_{A}^{(2)}}}{\partial q^{2}}\right)\right|_{\Delta^{-1}, q=F_{2}^{(n)}(t, R), \rho_{1}\left(n_{0}\right)} \\
\left.\frac{\partial F_{2}^{(n)}}{\partial R_{1}}\right|_{t, R} & =\frac{1}{n_{1}} \sum_{i, j=1}^{n_{1}} \mathbb{E}\left[\left(\left\langle x_{i}^{(1)} x_{j}^{(1)}\right\rangle_{\boldsymbol{n}, t, R}-\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, R}\left\langle x_{j}^{(1)}\right\rangle_{\boldsymbol{n}, t, R}\right)^{2}\right]  \tag{6.34}\\
\left.\frac{\partial F_{2}^{(n)}}{\partial R_{2}}\right|_{t, R} & =\frac{1}{n_{1}} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left\|\left\langle\ell_{Y_{\mu}^{\left(t, R_{2}\right)}}^{\prime}\left(s_{\mu}^{\left(t, R_{2}\right)}\right) \mathbf{x}^{(1)}\right\rangle_{\boldsymbol{n}, t, R}-\left\langle\ell_{Y_{\mu}^{\left(t, R_{2}\right)}}^{\prime}\left(s_{\mu}^{\left(t, R_{2}\right)}\right)\right\rangle_{\boldsymbol{n}, t, R}\left\langle\mathbf{x}^{(1)}\right\rangle_{\boldsymbol{n}, t, R}\right\|^{2} \tag{6.35}
\end{align*}
$$

Remember that $\ell_{y}^{\prime}(\cdot)$ is the derivative of $\ell_{y}: x \mapsto \ln P_{\text {out }}(y \mid x)$. Both $\partial F_{2}^{(n)} / \partial R_{1}$ and $\partial F_{2}^{(n)} / \partial R_{2}$ are clearly nonnegative. Using the assumption (H2), we easily obtain from (6.34) that

$$
\begin{equation*}
0 \leq\left.\frac{\partial F_{2}^{(n)}}{\partial R_{1}}\right|_{t, R} \leq 4 n_{1}\left\|\varphi_{1}\right\|_{\infty}^{4} \tag{6.36}
\end{equation*}
$$

In the proof of Lemma 6.13, under the hypothesis (H2), we obtain the upper bound (6.53) on $\left|\ell_{y}^{\prime}(x)\right|$. It yields $\forall x \in \mathbb{R}:\left|\ell_{Y_{\mu}^{\left(t, R_{2}\right)}}^{\prime}(x)\right| \leq\left\|\varphi_{2}^{\prime} / \sqrt{\Delta}\right\|_{\infty}\left(\left|Z_{\mu}\right|+2\left\|\varphi_{2} / \sqrt{\Delta}\right\|_{\infty}\right)$. Thus, we easily see from (6.35) that

$$
\begin{equation*}
0 \leq\left.\frac{\partial F_{2}^{(n)}}{\partial R_{2}}\right|_{t, R} \leq 2 n_{2}\left\|\varphi_{1}\right\|_{\infty}^{2}\left\|\frac{\varphi_{2}^{\prime}}{\sqrt{\Delta}}\right\|_{\infty}^{2}\left(2+8\left\|\frac{\varphi_{2}}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right) \tag{6.37}
\end{equation*}
$$

Finally, $\left.q \mapsto\left(\partial^{2} \Psi_{\varphi_{2}, P_{A}^{(2)}} / \partial q^{2}\right)\right|_{\Delta^{-1}, q, \rho_{1}\left(n_{0}\right)}$ is nonnegative continuous on $\left[0, \rho_{1}\left(n_{0}\right)\right]$, so there exists $C \geq 0$ such that $\forall q \in\left[0, \rho_{1}\left(n_{0}\right)\right]$ :

$$
\left.\left(\frac{\partial^{2} \Psi_{\varphi_{2}, P_{A}^{(2)}}}{\partial q^{2}}\right)\right|_{\Delta^{-1}, q, \rho_{1}\left(n_{0}\right)} \in[0, C]
$$

and $c(t, R) \in\left[0,2 \alpha_{2} C\right]$. Combining the latter with (6.33), (6.36) and (6.37) shows that $J_{\left.F^{(n)}(t,)\right)}(R)$ is uniformly bounded in

$$
(t, R) \in\left\{\left(t, R_{1}, R_{2}\right) \in[0,1] \times(0,+\infty)^{2}: R_{2}<2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t\right\} .
$$

By the mean-value theorem, $F^{(\boldsymbol{n})}$ is uniformly Lipschitz continuous in $R$.
By the Picard-Lindelöf theorem [104, Theorem 1.1], for all $\epsilon \in \mathcal{B}_{n_{0}}$ there exists a unique solution to the initial value problem ${ }^{d y} / d t=F^{(n)}(t, y), y(0)=\epsilon$ that we denote $R(\cdot, \epsilon):[0, \delta] \rightarrow[0,+\infty)^{2}$. Here $\delta \in[0,1]$ is such that $[0, \delta]$ is the maximal interval of existence of the solution. The function $F^{(n)}$ takes its values in $\left[0, r_{\max }\right] \times\left[0, \rho_{1}\left(n_{0}\right)\right]$ and $\epsilon \in \mathcal{B}_{n_{0}}$ so $\forall t \in[0, \delta]$ :

$$
R(t, \epsilon) \in\left[s_{n_{0}}, 2 s_{n_{0}}+r_{\max } t\right] \times\left[s_{n_{0}}, 2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t\right] .
$$

It means that $\delta=1$ (the solution never leaves the domain of definition of $F^{(\boldsymbol{n})}$ ).
Each initial condition $\epsilon \in \mathcal{B}_{n_{0}}$ is tied to a unique solution $R(\cdot, \epsilon)$. This implies that the function $\epsilon \mapsto R(t, \epsilon)$ is injective. Its Jacobian determinant is given by Liouville's formula 104, Chapter V, Corollary 3.1]:

$$
\begin{aligned}
\operatorname{det} J_{R(t,)}(\epsilon) & =\left.\exp \int_{0}^{t} d s\left(\frac{\partial F_{1}^{(\boldsymbol{n})}}{\partial R_{1}}+\frac{\partial F_{2}^{(\boldsymbol{n})}}{\partial R_{2}}\right)\right|_{s, R(s, \epsilon)} \\
& =\exp \int_{0}^{t} d s\left(\left.c(s, R(s, \epsilon)) \frac{\partial F_{2}^{(\boldsymbol{n})}}{\partial R_{1}}\right|_{s, R(s, \epsilon)}+\left.\frac{\partial F_{2}^{(\boldsymbol{n})}}{\partial R_{2}}\right|_{s, R(s, \epsilon)}\right) .
\end{aligned}
$$

This Jacobian determinant is greater than, or equal to, one since we have shown earlier in the proof that $c(t, R), \partial F_{1}^{(n)} / \partial R_{1}$ and $\partial F_{2}^{(n)} / \partial R_{2}$ are nonnegative. The fact that the Jacobian determinant is bounded away from 0 uniformly in $\epsilon$ implies, by the inverse function theorem, that the injective function $\epsilon \mapsto R(t, \epsilon)$ is a $C^{1}$-diffeomorphism from $\mathcal{B}_{n_{0}}$ onto its image.

### 6.3.5 Matching lower and upper bounds

We know choose interpolation functions that are solutions to the first-order ODEs (6.31) and (6.32) in order to prove a lower and upper bounds on the high-dimensional limit of the average free entropy $f_{n}$.

Theorem 6.8 (Lower bound on the asympotic average free entropy). Under the assumptions of Theorem 6.1, the average free entropy (6.9) satisfies

$$
\liminf _{n \rightarrow \infty} f_{n} \geq \sup _{r_{1} \geq 0} \inf _{q_{1} \in\left[0, \rho_{1}\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right) .
$$

Proof. Fix $r \in\left[0, r_{\text {max }}\right]$. For all $\epsilon \in \mathcal{B}_{n_{0}}$, we choose $R(\cdot, \epsilon)=\left(R_{1}(\cdot, \epsilon), R_{2}(\cdot, \epsilon)\right)$ to be the unique solution to the first-order ODE (6.31) with initial condition $R(0, \epsilon)=\epsilon$ (see Proposition 6.7). Then, we use the derivative $R^{\prime}(\cdot, \epsilon)$ of $R(\cdot, \epsilon)$ to define $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
q_{\epsilon}(t):=R_{2}^{\prime}(t, \epsilon)=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}, \quad r_{\epsilon}(t):=R_{1}^{\prime}(t, \epsilon)=r .
$$

By Proposition 6.7, the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ are regular. We can now apply Proposition 6.6 to get

$$
\begin{align*}
& f_{\boldsymbol{n}}=o_{\boldsymbol{n}}(1)+\frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right) \\
& +\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}}\left\{\widetilde{f}_{\mathrm{RS}}\left(\epsilon_{2}+r ; \rho_{0}\right)+\alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)\right)\right. \\
&  \tag{6.38}\\
& \left.\quad+\frac{\alpha_{1} r}{2}\left(\rho_{1}-\int_{0}^{1} d t q_{\epsilon}(t)\right)\right\},
\end{align*}
$$

where

$$
\widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right):=\sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} \psi_{P_{X}}\left(r_{0}\right)+\alpha_{1} \Psi_{\varphi_{1}, P_{A}^{(1)}}\left(r, q_{0} ; \rho_{0}\right)-\frac{r_{0} q_{0}}{2}
$$

By Lemma 6.12 in Appendix 6.A, $r \mapsto \widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right)+{ }_{1} r \rho_{1} / 2$ is nondecreasing so $\forall \epsilon \in \mathcal{B}_{n_{0}}: \widetilde{f}_{\mathrm{RS}}\left(\epsilon_{2}+r ; \rho_{0}\right)+\alpha_{1}\left(\epsilon_{2}+r\right) \rho_{1} / 2 \geq \widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right)+{ }^{\alpha_{1} r \rho_{1}} / 2$. Besides,

$$
\begin{aligned}
\sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q, r ; \rho_{0}, \rho_{1}\left(n_{0}\right)\right)= & \frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right)+\widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right) \\
& +\alpha_{1} \alpha_{2} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q ; \rho_{1}\left(n_{0}\right)\right)+\frac{\alpha_{1} r}{2}\left(\rho_{1}\left(n_{0}\right)-q\right) .
\end{aligned}
$$

So it follows directly from (6.38) that

$$
\begin{align*}
f_{\boldsymbol{n}} & \geq o_{\boldsymbol{n}}(1)+\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, \int_{0}^{1} q_{\epsilon}(t) d t, r ; \rho_{0}, \rho_{1}\left(n_{0}\right)\right) \\
& \geq o_{\boldsymbol{n}}(1)+\inf _{q_{1} \in\left[0, \rho_{1}\left(n_{0}\right)\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, \int_{0}^{1} q_{\epsilon}(t) d t, r ; \rho_{0}, \rho_{1}\left(n_{0}\right)\right), \tag{6.39}
\end{align*}
$$

where the lower bound is because $\int_{0}^{1} q_{\epsilon}(t) d t \in\left[0, \rho_{1}\left(n_{0}\right)\right]$. By continuity,

$$
\begin{aligned}
\lim _{n_{0} \rightarrow+\infty} \inf _{q_{1} \in\left[0, \rho_{1}\left(n_{0}\right)\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}} & \left(q_{0}, r_{0}, q_{1}, r ; \rho_{0}, \rho_{1}\left(n_{0}\right)\right) \\
& =\inf _{q_{1} \in\left[0, \rho_{1}\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r ; \rho_{0}, \rho_{1}\right) .
\end{aligned}
$$

Taking the limit inferior on both sides of (6.39), and using the latter limit, yields

$$
\liminf _{n_{0} \rightarrow+\infty} f_{n} \geq \inf _{q_{1} \in\left[0, \rho_{1}\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r ; \rho_{0}, \rho_{1}\right) .
$$

This inequality is true for all $r \in\left[0, r_{\text {max }}\right]$, hence

$$
\begin{equation*}
\liminf _{n_{0} \rightarrow \infty} f_{n} \geq \sup _{r_{1} \in\left[0, r_{\max }\right]} \inf _{q_{1} \in\left[0, \rho_{1}\right]} \sup _{\left.q_{0} \in\left[0, \rho_{0}\right]\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right) . \tag{6.40}
\end{equation*}
$$

To conclude the proof, it remains to extend the supremum over $r_{1} \in\left[0, r_{\max }\right]$ on the r.h.s. of 6.40 to a supremum over $r_{1} \geq 0$. Define the function

$$
\psi:\left(r_{1}, q_{1}\right) \in[0,+\infty) \times\left[0, \rho_{1}\right] \mapsto f\left(r_{1}\right)+g\left(q_{1}\right)+\frac{\alpha_{1}}{2} r_{1}\left(\rho_{1}-q_{1}\right),
$$

where $f:[0,+\infty) \rightarrow \mathbb{R}$ and $g:\left[0, \rho_{1}\right] \rightarrow \mathbb{R}$ are defined by

$$
\begin{aligned}
& f\left(r_{1}\right):=\frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right)+\widetilde{f}_{\mathrm{RS}}\left(r_{1} ; \rho_{0}\right), \\
& g\left(q_{1}\right):=\alpha_{1} \alpha_{2} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{1} ; \rho_{1}\right),
\end{aligned}
$$

Note that

$$
\begin{equation*}
\psi\left(r_{1}, q_{1}\right)=\sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right) . \tag{6.41}
\end{equation*}
$$

By [29, Proposition 18, Appendix B.2], $q_{1} \mapsto \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q ; \rho_{1}\right)$ is convex, so its derivative is nondecreasing. Besides, by the definition (6.17) of $r_{\text {max }}$, we have $r_{\max } \geq r^{*}:=\left.2 \alpha_{2}\left({ }^{(\Psi}{ }_{\varphi_{2}, P_{A}^{(2)}} / \partial q\right)\right|_{\Delta^{-1}, \rho_{1}, \rho_{1}}$. Thus, if $r_{1} \geq r_{\max }$, for every $q_{1} \in\left[0, \rho_{1}\right]:$

$$
\left.\frac{\partial \psi}{\partial q_{1}}\right|_{r_{1}, q_{1}}=\left.\alpha_{1} \alpha_{2}\left(\frac{\partial \Psi_{\varphi_{2}, P_{A}^{(2)}}}{\partial q}\right)\right|_{\Delta^{-1}, q_{1}, \rho_{1}}-\frac{\alpha_{1} r_{1}}{2} \leq \frac{\alpha_{1}}{2}\left(r^{*}-r_{\max }\right) \leq 0 .
$$

The latter implies that, for all $r_{1} \in\left[r_{\max },+\infty\right), \inf _{q_{1} \in\left[0, \rho_{1}\right]} \psi\left(r_{1}, q_{1}\right)=\psi\left(r_{1}, \rho_{1}\right)$ so

$$
\begin{aligned}
\inf _{q_{1} \in\left[0, \rho_{1}\right]} \psi\left(r_{1}, q_{1}\right)-\inf _{q_{1} \in\left[0, \rho_{1}\right]} \psi\left(r_{\max }, q_{1}\right) & =\psi\left(r_{1}, \rho_{1}\right)-\psi\left(r_{\max }, \rho_{1}\right) \\
& =\widetilde{f}_{\mathrm{RS}}\left(r_{1} ; \rho_{0}\right)-\widetilde{f}_{\mathrm{RS}}\left(r_{\max } ; \rho_{0}\right) \leq 0 .
\end{aligned}
$$

The nonpositivity follows from Lemma 6.12 in Appendix 6.A where we show that $\widetilde{f}_{\mathrm{RS}}\left(\cdot ; \rho_{0}\right)$ is nondecreasing. Combining the upper bound (6.40), the identity (6.41), and the fact that $\inf _{q_{1} \geq\left[0, \rho_{1}\right]} \psi\left(r_{1}, q_{1}\right) \leq \inf _{q_{1} \geq\left[0, \rho_{1}\right]} \psi\left(r_{\text {max }}, q_{1}\right)$ for every $r_{1} \in\left[r_{\text {max }},+\infty\right)$, ends the proof.

Theorem 6.9 (Upper bound on the asympotic averaged free entropy). Under the assumptions of Theorem 6.1, the average free entropy (6.9) satisfies

$$
\limsup _{n \rightarrow+\infty} f_{n} \leq \sup _{r_{1} \geq 0} \inf _{q_{1} \in\left[0, \rho_{1}\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right) .
$$

Proof. For all $\epsilon \in \mathcal{B}_{n_{0}}$, we choose $R(\cdot, \epsilon)=\left(R_{1}(\cdot, \epsilon), R_{2}(\cdot, \epsilon)\right)$ to be the unique solution to the first-order ODE 6.32 with initial condition $R(0, \epsilon)=\epsilon$ (see Proposition 6.7). Then, we use the derivative $R^{\prime}(\cdot, \epsilon)$ of $R(\cdot, \epsilon)$ to define $\forall(t, \epsilon) \in$ $[0,1] \times \mathcal{B}_{n_{0}}:$

$$
\begin{aligned}
& q_{\epsilon}(t):=R_{2}^{\prime}(t, \epsilon)=\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon} \\
& r_{\epsilon}(t):=R_{1}^{\prime}(t, \epsilon)=\left.2 \alpha_{2}\left(\frac{\partial \Psi_{\varphi 2, P_{A}^{(2)}}}{\partial q}\right)\right|_{r=\Delta^{-1}, q=q_{\epsilon}(t), \rho=\rho_{1}\left(n_{0}\right)}
\end{aligned}
$$

By Proposition 6.7, the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ are regular. We can now apply Proposition 6.6 to get

$$
\begin{align*}
& f_{\boldsymbol{n}}=o_{\boldsymbol{n}}(1)+\frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right) \\
&+\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}}\left\{\widetilde{f}_{\mathrm{RS}}\left(R_{1}(1, \epsilon) ; \rho_{0}\right)+\alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)\right)\right. \\
&\left.\quad+\int_{0}^{1} d t \frac{\alpha_{1}}{2} r_{\epsilon}(t)\left(\rho_{1}\left(n_{0}\right)-q_{\epsilon}(t)\right)\right\} \tag{6.42}
\end{align*}
$$

The function $q \in\left[0, \rho_{1}\left(n_{0}\right)\right] \mapsto \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q ; \rho_{1}\left(n_{0}\right)\right)$ is convex so, by Jensen's inequality, $\forall \epsilon \in \mathcal{B}_{n_{0}}$ :

$$
\begin{equation*}
\Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, \int_{0}^{1} q_{\epsilon}(t) d t ; \rho_{1}\left(n_{0}\right)\right) \leq \int_{0}^{1} d t \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{\epsilon}(t) ; \rho_{1}\left(n_{0}\right)\right) \tag{6.43}
\end{equation*}
$$

The function $r \in[0,+\infty) \mapsto \widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right)$ is convex and Lipschitz continuous on $[0,+\infty)$ so $\forall \epsilon \in \mathcal{B}_{n_{0}}$ :

$$
\begin{align*}
\widetilde{f}_{\mathrm{RS}}\left(R_{1}(1, \epsilon) ; \rho_{0}\right) & =\widetilde{f}_{\mathrm{RS}}\left(\epsilon_{1}+\int_{0}^{1} d t r_{\epsilon}(t) ; \rho_{0}\right) \\
& =O\left(s_{n_{0}}\right)+\widetilde{f}_{\mathrm{RS}}\left(\int_{0}^{1} d t r_{\epsilon}(t) ; \rho_{0}\right) \\
& \leq O\left(s_{n_{0}}\right)+\int_{0}^{1} d t \widetilde{f}_{\mathrm{RS}}\left(r_{\epsilon}(t) ; \rho_{0}\right) \tag{6.44}
\end{align*}
$$

where the second equality is due to the Lipschitzness of $\widetilde{f}_{\mathrm{RS}}\left(\cdot ; \rho_{0}\right)$ and the subsequent inequality to Jensen's inequality. We use (6.43) and (6.44) to upper bound (6.42). We obtain

$$
\begin{align*}
f_{\boldsymbol{n}} \leq o_{\boldsymbol{n}}(1) & +\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \int_{0}^{1} d t\left\{\frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right)+\widetilde{f}_{\mathrm{RS}}\left(r_{\epsilon}(t) ; \rho_{0}\right)\right. \\
& \left.+\alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{\epsilon}(t) ; \rho_{1}\left(n_{0}\right)\right)+\frac{\alpha_{1}}{2} r_{\epsilon}(t)\left(\rho_{1}\left(n_{0}\right)-q_{\epsilon}(t)\right)\right\} \tag{6.45}
\end{align*}
$$

Fix $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$. Our choice of interpolation functions implies that

$$
r_{\epsilon}(t)=\left.2 \alpha_{2}\left(\frac{\partial \Psi_{\varphi_{2}, P_{A}^{(2)}}}{\partial q}\right)\right|_{r=\Delta^{-1}, q=q_{\epsilon}(t), \rho=\rho_{1}\left(n_{0}\right)},
$$

so $q_{\epsilon}(t)$ is a stationary point of the convex function

$$
q_{1} \in\left[0, \rho_{1}\left(n_{0}\right)\right] \mapsto \alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{1} ; \rho_{1}\left(n_{0}\right)\right)+\frac{\alpha_{1}}{2} r_{\epsilon}(t)\left(\rho_{1}\left(n_{0}\right)-q_{1}\right) .
$$

Hence, $q_{\epsilon}(t)$ is a global minimum of the latter convex function and

$$
\begin{aligned}
\alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}} & \left(\Delta^{-1}, q_{\epsilon}(t) ; \rho_{1}\left(n_{0}\right)\right)+\frac{\alpha_{1}}{2} r_{\epsilon}(t)\left(\rho_{1}\left(n_{0}\right)-q_{\epsilon}(t)\right) \\
& =\inf _{q_{1} \in\left[0, \rho_{1}\left(n_{0}\right)\right]} \alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{1} ; \rho_{1}\left(n_{0}\right)\right)+\frac{\alpha_{1}}{2} r_{\epsilon}(t)\left(\rho_{1}\left(n_{0}\right)-q_{1}\right) .
\end{aligned}
$$

Plugging this identity back in 6.45 yields

$$
\begin{aligned}
& f_{n} \leq o_{\boldsymbol{n}}(1)+\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \int_{0}^{1} d t\left\{\frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right)+\widetilde{f}_{\mathrm{RS}}\left(r_{\epsilon}(t) ; \rho_{0}\right)\right. \\
& \left.\underset{q_{1} \in\left[0, \rho_{1}\left(n_{0}\right)\right]}{+\inf _{2}} \alpha_{2} \alpha_{1} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{1} ; \rho_{1}\left(n_{0}\right)\right)+\frac{\alpha_{1}}{2} r_{\epsilon}(t)\left(\rho_{1}\left(n_{0}\right)-q_{1}\right)\right\} \\
& =o_{\boldsymbol{n}}(1)+\int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \int_{0}^{1} d t \inf _{q_{1} \in\left[0, \rho_{1}\left(n_{0}\right)\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{\epsilon}(t) ; \rho_{0}, \rho_{1}\left(n_{0}\right)\right) \\
& \leq o_{\boldsymbol{n}}(1)+\sup _{r_{1} \geq 0} \inf _{q_{1} \in\left[0, \rho_{1}\left(n_{0}\right)\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{\epsilon}(t) ; \rho_{0}, \rho_{1}\left(n_{0}\right)\right) \text {. }
\end{aligned}
$$

Taking the limit superior on both sides of this last inequality yields the desired result.

The lower bound of Theorem 6.8 matches the upper bound of Theorem 6.9, hence

$$
\lim _{n \rightarrow+\infty} f_{n}=\sup _{r_{1} \geq 0} \inf _{q_{1} \in\left[0, \rho_{1}\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right) .
$$

We conclude this section by showing that we can invert the order of the optimizations on $r_{1}$ and $q_{1}$, thus ending the proof of Theorem 6.1.

Lemma 6.10 (Switching the optimization order). Under the assumptions of Theorem 6.1, we have

$$
\begin{aligned}
& \sup _{r_{1} \geq 0} \inf _{q_{1} \in\left[0, \rho_{1}\right]} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right) \\
&=\sup _{q_{1} \in\left[0, \rho_{1}\right]} \inf _{r_{1} \geq 0} \sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right) .
\end{aligned}
$$

Proof. Define the function

$$
h:\left(r_{1}, q_{1}\right) \in[0,+\infty) \times\left[0, \rho_{1}\right] \mapsto f\left(r_{1}\right)+g\left(q_{1}\right)-\frac{\alpha_{1}}{2} r_{1} q_{1}
$$

where $f:[0,+\infty) \rightarrow \mathbb{R}$ and $g:\left[0, \rho_{1}\right] \rightarrow \mathbb{R}$ are defined by

$$
\begin{aligned}
f\left(r_{1}\right) & :=\frac{\alpha_{1}}{2} \ln \left(2 \pi e \Delta^{-\alpha_{2}}\right)+\widetilde{f}_{\mathrm{RS}}\left(r_{1} ; \rho_{0}\right)+\frac{\alpha_{1} \rho_{1}}{2} r_{1}, \\
g\left(q_{1}\right) & :=\alpha_{1} \alpha_{2} \Psi_{\varphi_{2}, P_{A}^{(2)}}\left(\Delta^{-1}, q_{1} ; \rho_{1}\right)
\end{aligned}
$$

Note that $h\left(r_{1}, q_{1}\right)=\sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} f_{\mathrm{RS}}\left(q_{0}, r_{0}, q_{1}, r_{1} ; \rho_{0}, \rho_{1}\right)$. The function $f$ is convex nondecreasing ( $\alpha_{1} \rho_{1} / 2$ )-Lipschitz continuous on $[0,+\infty$ ) (see Lemma 6.12 in Appendix (6.A) while the function $g$ is convex nondecreasing Lipschitz continuous on $\left[0, \rho_{1}\right]$ 29, Proposition 18, Appendix B.2]. Hence, we can apply [29, Corollary 7, Appendix D] and get

$$
\sup _{r_{1} \geq 0} \inf _{q_{1} \in\left[0, \rho_{1}\right]} h\left(r_{1}, q_{1}\right)=\sup _{q_{1} \in\left[0, \rho_{1}\right]} \inf _{r_{1} \geq 0} h\left(r_{1}, q_{1}\right) .
$$

## Appendices

## 6.A Miscellaneous useful results

Lemma 6.11 (Convergence of the sequence $\left.\left(\rho_{1}\left(n_{0}\right)\right)_{n_{0} \geq 1}\right)$. Let $k$ be a nonnegative integer, $\varphi: \mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R}^{k}$ a measurable bounded function, $P_{A}$ a probability distribution on $\mathbb{R}^{k}$, and $P_{X}$ a probability distribution on $\mathbb{R}$ with finite second moment. For $n_{0} \in \mathbb{N}^{*}$, let $\mathbf{X} \in \mathbb{R}^{n_{0}}$ be a random vector with entries $X_{1}, X_{2}, \ldots, X_{n_{0}} \stackrel{\text { i.i.d. }}{\sim} P_{X}$, $\mathbf{W} \in \mathbb{R} n_{0}$ a random vector with entries $W_{1}, X_{2}, \ldots, W_{n_{0}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$, and $\mathbf{A} \sim P_{A}$ such that $(\mathbf{X}, \mathbf{W}, \mathbf{A})$ are independent. Define $\rho_{1}\left(n_{0}\right):=\mathbb{E}\left[\varphi\left(\mathbf{W}^{\top} \mathbf{x} / \sqrt{n_{0}}, \mathbf{A}\right)^{2}\right]$ and the second moments $\rho_{0}:=\mathbb{E} X_{1}^{2}, \rho_{1}:=\mathbb{E}\left[\varphi\left(\sqrt{\rho_{0}} W_{1}, \mathbf{A}\right)^{2}\right]$. Then, the sequence $\left\{\rho_{1}\left(n_{0}\right)\right\}_{n_{0} \geq 1}$ converges and

$$
\lim _{n_{0} \rightarrow+\infty} \rho_{1}\left(n_{0}\right)=\rho_{1} .
$$

Proof. If $\rho_{0}=0$ then $\mathbf{X}=0$ almost surely (a.s.) and $\rho_{1}\left(n_{0}\right)=\mathbb{E} \varphi^{2}(0, \mathbf{A})=\rho_{1}$ for all $n_{0} \in \mathbb{N}^{*}$. From now on, we assume that $\rho_{0}>0$. Define

$$
h: v \in(0,+\infty) \mapsto \int d t d P_{A}(\mathbf{a}) \varphi^{2}(t, \mathbf{a}) \frac{1}{\sqrt{2 \pi v}} \exp \left(-t^{2} / 2 v\right) .
$$

Conditionally on $\mathbf{X}$, we have $\mathbf{W}^{\top} \mathbf{x} / \sqrt{n_{0}} \sim \mathcal{N}\left(0,\|\mathbf{X}\|^{2} / n_{0}\right)$ so

$$
\rho_{1}\left(n_{0}\right)=\mathbb{E}\left[\mathbb{E}\left[\varphi^{2}\left(\mathbf{W}^{\top} \mathbf{x} / \sqrt{n_{0}}, \mathbf{A}\right) \mid \mathbf{X}\right]\right]=\mathbb{E}\left[h\left(\frac{\|\mathbf{X}\|^{2}}{n_{0}}\right)\right] .
$$

By the dominated convergence theorem, $h$ is continuous on $(0,+\infty)$. By the strong law of large numbers, $\|\mathbf{X}\|^{2} / n_{0}$ converges a.s. to $\rho_{0}$. Combined with the continuity of $h$, it comes

$$
\lim _{n_{0} \rightarrow+\infty} h\left(\frac{\|\mathbf{X}\|^{2}}{n_{0}}\right)=h\left(\rho_{0}\right)=\rho_{1} \quad \text { almost surely. }
$$

Note that $\left|h\left(\|\mathbf{X}\|^{2} / n_{0}\right)\right|$ is upper bounded by $\|\varphi\|_{\infty}^{2}$. By the dominated convergence theorem, we conclude that

$$
\rho_{1}\left(n_{0}\right)=\mathbb{E}\left[h\left(\frac{\|\mathbf{X}\|^{2}}{n_{0}}\right)\right] \underset{n_{0} \rightarrow+\infty}{ } \mathbb{E}\left[\lim _{n_{0} \rightarrow+\infty} h\left(\frac{\|\mathbf{X}\|^{2}}{n_{0}}\right)\right]=\rho_{1} .
$$

Lemma 6.12. Let $k$ be a nonnegative integer, $\varphi: \mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R}^{k}$ a measurable bounded function, and $P_{A}$ a probability distribution on $\mathbb{R}^{k}$. Let $U, V, Z \sim \mathcal{N}(0,1)$, $A \sim P_{A}$ be independent random variables. For $(r, \rho) \in[0,+\infty)$ and $q \in[0, \rho]$, define $Y^{(r, q ; \rho)}=\sqrt{r} \varphi(\sqrt{\rho-q} U+\sqrt{q} V, \mathbf{A})+Z$ and the average free entropy

$$
\psi_{\varphi, P_{A}}(r, q ; \rho):=\mathbb{E}\left[\ln \left(\int d u \frac{e^{-\frac{u^{2}}{2}}}{\sqrt{2 \pi}} \int \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi}} e^{-\frac{1}{2}\left(Y^{(r, q ; \rho)}-\sqrt{r} \varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})\right)^{2}}\right)\right]
$$

Then, for all $\rho \in[0,+\infty)$ and $q \in[0, \rho]$, the function $\Psi_{\varphi, P_{A}}(q, \cdot ; \rho)$ is twicedifferentiable, convex, nonincreasing and $\left(\mathbb{E}\left[\varphi(\sqrt{\rho} U, \mathbf{A})^{2}\right] / 2\right)$-Lipschitz continuous on $[0,+\infty)$. Let $P_{X}$ be a probability distribution on $\mathbb{R}$ with finite second moment $\rho_{0}:=\mathbb{E}_{X \sim P_{X}}\left[X^{2}\right], \alpha_{1}$ a positive real number, and

$$
\widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right):=\sup _{q_{0} \in\left[0, \rho_{0}\right]} \inf _{r_{0} \geq 0} \psi_{P_{X}}\left(r_{0}\right)+\alpha_{1} \Psi_{\varphi, P_{A}}\left(q_{0}, r ; \rho_{0}\right)-\frac{r_{0} q_{0}}{2}
$$

where $\psi_{P_{X}}$ is defined by 6.10). Then, the function $\widetilde{f}_{\mathrm{RS}}\left(\cdot ; \rho_{0}\right)$ is convex nonincreasing $\left(\alpha_{1} \mathbb{E}\left[\varphi\left(\sqrt{\rho_{0}} U, \mathbf{A}\right)^{2}\right] / 2\right)$-Lipschitz continuous on $[0,+\infty)$ while

$$
r \mapsto \widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right)+\frac{\alpha_{1} r}{2} \mathbb{E}\left[\varphi\left(\sqrt{\rho_{0}} U, \mathbf{A}\right)^{2}\right]
$$

is convex nondecreasing and $\left(\alpha_{1} \mathbb{E}\left[\varphi\left(\sqrt{\rho_{0}} U, \mathbf{A}\right)^{2}\right] / 2\right)$-Lipschitz continuous on $[0,+\infty)$.
Proof. Fix $\rho \in[0,+\infty)$ and $q \in[0, \rho]$. Let $\psi(r):=\Psi_{\varphi, P_{A}}(q, r ; \rho)$. Note that

$$
\begin{equation*}
\psi(r)=\widetilde{\psi}(r)-\frac{1+r \mathbb{E}\left[\varphi\left(\sqrt{\rho_{0}} U, \mathbf{A}\right)^{2}\right]}{2} \tag{6.46}
\end{equation*}
$$

where

$$
\begin{align*}
\widetilde{\psi}(r):=\mathbb{E} \ln & \int \mathcal{D} u \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi}} \exp (r \varphi(\sqrt{\rho-q} U+\sqrt{q} V, \mathbf{A}) \varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a}) \\
& \left.-\frac{r}{2} \varphi^{2}(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})+\sqrt{r} \varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a}) Z\right), \tag{6.47}
\end{align*}
$$

where $\mathcal{D} u:=d u e^{-u^{2} / 2} / \sqrt{2 \pi}$. The boundedness of $\varphi$ ensures that all the domination hypotheses to prove the twice-differentiability of $\psi$ are reunited. We denote by the angular brackets $\langle-\rangle_{r}$ the expectation w.r.t. the joint posterior distribution

$$
d P\left(u, \mathbf{a} \mid Y^{(r, q ; \rho)}, V\right)=\frac{\mathcal{D} u d P_{A}(\mathbf{a})}{\mathcal{Z}(r, q ; \rho)} e^{\sqrt{r} Y^{(r, q ;)} \varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})-\frac{r}{2} \varphi^{2}(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})}
$$

where $\mathcal{Z}(r, q ; \rho)$ is a normalization factor. Differentiating (6.46) under the expectation sign, we have $\forall r \geq 0$ :

$$
\begin{align*}
\widetilde{\psi}^{\prime}(r)= & \mathbb{E}\langle\varphi(\sqrt{\rho-q} U+\sqrt{q} V, \mathbf{A}) \varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})\rangle_{r} \\
& -\frac{1}{2} \mathbb{E}\left\langle\varphi^{2}(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})\right\rangle_{r}+\frac{1}{2 \sqrt{r}} \mathbb{E}\left[Z\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})\rangle_{r}\right] \\
= & \frac{1}{2} \mathbb{E}\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{A})\rangle_{r}^{2}, \tag{6.48}
\end{align*}
$$

where the second equality follows from a Gaussian integration by parts w.r.t. $Z$ and the Nishimori identity (see Lemma 2.1)
$\mathbb{E}\langle\varphi(\sqrt{\rho-q} U+\sqrt{q} V, \mathbf{A}) \varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})\rangle_{r}=\mathbb{E}\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})\rangle_{r}^{2}$.
Thus, for all $r \in[0,+\infty), \widetilde{\psi^{\prime}}(r) \geq 0$ and, by Jensen's inequality,

$$
\begin{aligned}
\psi^{\prime}(r) & \leq \frac{1}{2} \mathbb{E}\left\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})^{2}\right\rangle_{r} \\
& =\frac{1}{2} \mathbb{E}\left[\varphi(\sqrt{\rho-q} U+\sqrt{q} V, \mathbf{A})^{2}\right]=\frac{\mathbb{E}\left[\varphi(\sqrt{\rho} U, \mathbf{A})^{2}\right]}{2},
\end{aligned}
$$

where the first equality is due to the Nishimori identity. We see that $\forall r \in[0,+\infty)$ :

$$
\psi^{\prime}(r) \in\left[-\frac{\mathbb{E}\left[\varphi(\sqrt{\rho} U, \mathbf{A})^{2}\right]}{2}, 0\right]
$$

so $\psi$ is nonincreasing and $\left(\mathbb{E}\left[\varphi(\sqrt{\rho} U, \mathbf{A})^{2}\right] / 2\right)$-Lipschitz continuous. Further differentiating, integrating by parts w.r.t. $Z$ and applying the Nishimori identity, it comes $\forall r \geq 0$ :

$$
\psi^{\prime \prime}(r)=\widetilde{\psi}^{\prime \prime}(r)=\frac{1}{2} \mathbb{E}\left[\left(\left\langle\varphi^{2}(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})\right\rangle_{r}-\langle\varphi(\sqrt{\rho-q} u+\sqrt{q} V, \mathbf{a})\rangle_{r}^{2}\right)^{2}\right] .
$$

Thus, $\psi^{\prime \prime}(r)$ is nonnegative and $\psi$ is convex on $[0,+\infty)$.
By definition, $\widetilde{f}_{\mathrm{RS}}\left(\cdot ; \rho_{0}\right)$ is the supremum of the functions

$$
r \in[0,+\infty) \mapsto \alpha_{1} \Psi_{\varphi, P_{A}}\left(q_{0}, r ; \rho_{0}\right)+\inf _{r_{0} \geq 0} \psi_{P_{X}}\left(r_{0}\right)-\frac{r_{0} q_{0}}{2}
$$

that are convex nonincreasing $\left(\alpha_{1} \mathbb{E}\left[\varphi(\sqrt{\rho} U, \mathbf{A})^{2}\right] / 2\right)$-Lipschitz continuous, hence its properties. Similarly, the properties of $r \mapsto \widetilde{f}_{\mathrm{RS}}\left(r ; \rho_{0}\right)+\left(\alpha_{1} \mathbb{E}\left[\varphi\left(\sqrt{\rho_{0}} U, \mathbf{A}\right)^{2}\right] / 2\right) r$ follow directly from the function being the supremum of the convex nondecreasing $\left(\alpha_{1} \mathbb{E}\left[\varphi(\sqrt{\rho} U, \mathbf{A})^{2}\right] / 2\right)$-Lipschitz continuous functions

$$
r \in[0,+\infty) \mapsto \frac{\alpha_{1} r \mathbb{E}\left[\varphi\left(\sqrt{\rho_{0}} U, \mathbf{A}\right)^{2}\right]}{2}+\alpha_{1} \Psi_{\varphi, P_{A}}\left(q_{0}, r ; \rho_{0}\right)+\inf _{r_{0} \geq 0} \psi_{P_{X}}\left(r_{0}\right)-\frac{r_{0} q_{0}}{2}
$$

## 6.B Interpolating average free entropy at $t=0$

Let $\left(s_{n_{0}}\right)_{n_{0} \geq 1}$ be a decreasing sequence of real numbers in $(0,1 / 2]$ with limit $\lim _{n_{0} \rightarrow+\infty} s_{n_{0}}=0$. Consider the setting of Subsection 6.2.1. Let $\mathbf{U}, \mathbf{V}$ be $n_{2}$-dimensional random vectors with entries $U_{\mu}, V_{\mu} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. For $\epsilon=\left(\epsilon_{1}, \epsilon_{2}, \epsilon_{3}\right) \in$ $\left[0,2 s_{n_{0}}\right]^{3}$, we denote by $\mathbf{S}^{\left(\epsilon_{2}, \epsilon_{3}\right)}$ the $n_{2}$-dimensional random vector whose entries are given for all $\mu \in\left\{1, \ldots, n_{2}\right\}$ by

$$
S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}:=\frac{1}{\sqrt{n_{1}}}\left[\mathbf{W}^{(2)} \mathbf{X}^{(1)}\right]_{\mu}+\sqrt{\epsilon_{3}} U_{\mu}+\sqrt{\epsilon_{2}} V_{\mu}
$$

and we consider the observations

$$
\begin{cases}Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} & \sim P_{\mathrm{out}}\left(\cdot \mid S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right), \quad 1 \leq \mu \leq n_{2} \\ \widetilde{Y}_{i}^{\left.\epsilon_{1}\right)} & =\sqrt{\epsilon_{1}} X_{i}^{(1)}+\widetilde{Z}_{i}, \quad 1 \leq i \leq n_{1}\end{cases}
$$

where $\left(\widetilde{Z}_{i}\right)_{i=1}^{n_{1}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. The two random vectors $\mathbf{Y}^{\left(\epsilon_{2}, \epsilon_{3}\right)}:=\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)_{\mu=1}^{n_{2}}$ and $\tilde{\mathbf{Y}}^{\left(\epsilon_{1}\right)}:=\left(\widetilde{Y}_{i}^{\left(\epsilon_{1}\right)}\right)_{i=1}^{n_{1}}$ sum up these observations. The joint posterior distribution of $\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{\left(\epsilon_{2}, \epsilon_{3}\right)}, \widetilde{\mathbf{Y}}^{\left(\epsilon_{1}\right)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)$ is

$$
\begin{equation*}
d P\left(\mathbf{x}, \mathbf{a}, \mathbf{u} \mid \mathbf{Y}^{\left(\epsilon_{2}, \epsilon_{3}\right)}, \widetilde{\mathbf{Y}}^{\left(\epsilon_{1}\right)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right):=\frac{d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a}) \mathcal{D} \mathbf{u}}{\widetilde{\mathcal{Z}}_{\boldsymbol{n}}(\epsilon)} e^{-\widetilde{\mathcal{H}}_{n}(\epsilon, \mathbf{x}, \mathbf{a}, \mathbf{u})} \tag{6.49}
\end{equation*}
$$

where $\mathcal{D} \mathbf{u}:=d \mathbf{u} e^{-\frac{\|\mathbf{u}\|^{2}}{2}} / \sqrt{2 \pi^{n}}, \widetilde{\mathcal{Z}}_{n}(\epsilon)$ is the normalization factor, and $\widetilde{\mathcal{H}}_{n}$ is the Hamiltonian

$$
\widetilde{\mathcal{H}}_{n}(\epsilon, \mathbf{x}, \mathbf{a}, \mathbf{u}):=-\sum_{\mu=1}^{n_{2}} \ln P_{\text {out }}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} \mid s_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)+\frac{1}{2} \sum_{i=1}^{n_{1}}\left(\widetilde{Y}_{i}^{\left(\epsilon_{1}\right)}-\sqrt{\epsilon_{1}} x_{i}^{(1)}\right)^{2}
$$

with $x_{i}^{(1)}:=\varphi_{1}\left(\left[\mathbf{W}^{(1)} \mathbf{x} / \sqrt{n_{0}}\right]_{i}, \mathbf{a}_{i}\right)$ and $\left.s_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}:=\left[\mathbf{W}^{(2)} \mathbf{x}^{(1)} / \sqrt{n_{1}}\right]\right]_{\mu}+\sqrt{\epsilon_{3}} u_{\mu}+\sqrt{\epsilon_{2}} V_{\mu}$. The average free entropy associated with the latter posterior is

$$
\widetilde{f}_{\boldsymbol{n}}(\epsilon):=\frac{\mathbb{E} \ln \widetilde{\mathcal{Z}}_{\boldsymbol{n}}(\epsilon)}{n_{0}}=\frac{1}{n_{0}} \mathbb{E} \ln \int d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a}) \mathcal{D} \mathbf{u} e^{-\widetilde{\mathcal{H}}_{n}(\epsilon, \mathbf{x}, \mathbf{a}, \mathbf{u})}
$$

Note that, for all $\epsilon=\left(\epsilon_{1}, \epsilon_{2}\right) \in\left[s_{n_{0}}, 2 s_{n_{0}}\right]^{2}, \widetilde{f}_{\boldsymbol{n}}\left(\epsilon_{1}, \epsilon_{2}, 2 s_{n}-\epsilon_{2}\right)=f_{\boldsymbol{n}}(0, \epsilon)$, where $f_{n}(t, \epsilon)$ is the average free entropy defined by (6.22) evaluated at $t=0$. Besides, $\widetilde{f}_{n}(0)=f_{n}-\frac{n_{1}}{2 n_{0}}$.

Lemma 6.13. Assume that (H1), (H2), (H3) hold and $\boldsymbol{n}$ is such that

$$
\lim _{n_{0} \rightarrow+\infty} \frac{n_{1}}{n_{0}}=\alpha_{1}, \lim _{n_{0} \rightarrow+\infty} \frac{n_{2}}{n_{1}}=\alpha_{2}
$$

Then, $\forall \epsilon \in\left[0,2 s_{n_{0}}\right]^{3}$ :

$$
\widetilde{f}_{n}(\epsilon)=f_{n}-\frac{n 1}{2 n_{0}}+O\left(s_{n_{0}}\right)
$$

where $\left|O\left(s_{n_{0}}\right) / s_{n_{0}}\right|$ is bounded uniformly in $\boldsymbol{n}$ and $\epsilon \in\left[0,2 s_{n_{0}}\right]^{3}$, and $f_{\boldsymbol{n}}$ is the average free entropy defined by 6.9.

Proof. Let ( $\mathbf{x}, \mathbf{a}, \mathbf{u}$ ) be a triplet sampled from the posterior distribution (6.49). We denote by angular brackets $\langle-\rangle_{\boldsymbol{n}, \epsilon}$ the expectation with respect to this posterior. Derivation under the expectation sign yields

$$
\begin{aligned}
\left.\frac{\partial \widetilde{f}_{n}}{\partial \epsilon_{1}}\right|_{\epsilon}=-\frac{1}{n_{0}} \mathbb{E}\left\langle\frac{\partial \widetilde{\mathcal{H}}_{n}}{\partial \epsilon_{1}}\right\rangle_{n, \epsilon} & =-\frac{1}{n_{0}} \sum_{i=1}^{n_{1}} \mathbb{E}\left\langle\frac{\left(X_{i}^{(1)}-x_{i}^{(1)}\right)^{2}}{2}+\frac{\widetilde{Z}_{i}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)}{2 \sqrt{\epsilon_{1}}}\right\rangle_{n, \epsilon} \\
& =-\frac{1}{2 n_{0}} \mathbb{E}\left[\left\|\mathbf{X}^{(1)}\right\|^{2}-\left\|\left\langle\mathbf{x}^{(1)}\right\rangle_{\boldsymbol{n}, \epsilon}\right\|^{2}\right]
\end{aligned}
$$

where the last equality follows from integrating by parts with respect to the standard Gaussian random variables $\widetilde{Z}_{i}$ and the Nishimori identity. Therefore,

$$
\left|\frac{\partial \widetilde{f}_{n}}{\partial \epsilon_{1}}\right|_{\epsilon} \left\lvert\,=\frac{1}{2 n_{0}} \mathbb{E}\left[\left\|\mathbf{X}^{(1)}\right\|^{2}-\left\|\left\langle\mathbf{x}^{(1)}\right\rangle_{\boldsymbol{n}, \epsilon}\right\|^{2}\right] \leq \frac{1}{2 n_{0}} \mathbb{E}\left\|\mathbf{X}^{(1)}\right\|^{2} \leq \frac{n_{1}}{2 n_{0}} \rho_{1}\left(n_{0}\right) .\right.
$$

We see that, for $n_{0}, n_{1}$ large enough, for all $\epsilon \in\left[0,2 s_{n_{0}}\right]^{3}$ :

$$
\begin{equation*}
\left.\left|\frac{\partial \widetilde{f}_{n}}{\partial \epsilon_{1}}\right|_{\epsilon} \right\rvert\, \leq \alpha_{1}\left\|\varphi_{1}\right\|_{\infty} \tag{6.50}
\end{equation*}
$$

To compute $\partial \tilde{f}_{n} / \partial \epsilon_{2}$ and $\partial \tilde{f}_{n} / \partial \epsilon_{3}$, we proceed like in Appendix 6.C where we compute the derivative of the interpolating average free entropy. We find

$$
\begin{align*}
&\left.\frac{\partial \widetilde{f}_{n}}{\partial \epsilon_{2}}\right|_{\epsilon}= \frac{1}{2} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} \mid S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)}{P_{\text {out }}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} \mid S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)} \frac{\ln \widetilde{\mathcal{Z}}_{n}(\epsilon)}{n_{0}}\right] \\
&+\frac{1}{2 n_{0}} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\left\langle\ell_{Y_{\mu}^{\prime\left(\epsilon_{2}, \epsilon_{3}\right)}}^{\prime}\left(s_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)\right\rangle_{n, \epsilon}^{2}\right]  \tag{6.51}\\
&\left.\frac{\partial \widetilde{f}_{n}}{\partial \epsilon_{3}}\right|_{\epsilon}= \frac{1}{2} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} \mid S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)}{P_{\text {out }}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} \mid S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)} \frac{\ln }{\widetilde{\mathcal{Z}}_{n}(\epsilon)}\right.  \tag{6.52}\\
& n_{0}
\end{align*},
$$

where $\ell_{y}^{\prime}$ is the derivative of $\ell_{y}: x \mapsto \ln P_{\text {out }}(y \mid x)$ and $P_{\text {out }}^{\prime \prime}(y \mid \cdot)$ the second derivative of $x \mapsto P_{\text {out }}(y \mid x)$. By Jensen's inequality and the Nishimori identity, we have for all $\mu \in\left\{1, \ldots, n_{2}\right\}$ :

$$
\mathbb{E}\left[\left\langle\ell_{Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}}^{\prime}\left(s_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)\right\rangle_{\boldsymbol{n}, \epsilon}^{2}\right] \leq \mathbb{E}\left[\left\langle\ell_{Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}}^{\prime}\left(s_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)^{2}\right\rangle_{\boldsymbol{n}, \epsilon}\right]=\mathbb{E}\left[\ell_{Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}}^{\prime}\left(S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)^{2}\right] .
$$

As $\ell_{y}(x):=\ln \int \frac{d P_{A}^{(2)}(\mathbf{a})}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}\left(y-\varphi_{2}(x, \mathbf{a})\right)^{2}}$ and $Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}=\varphi_{2}\left(S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}, \mathbf{A}_{\mu}^{(2)}\right)+\sqrt{\Delta} Z_{\mu}$, it comes

$$
\begin{align*}
\left|\ell_{Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}}^{\prime}(x)\right| & =\frac{\left|\int d P_{A}^{(2)}(\mathbf{a}) \varphi_{2}^{\prime}(x, \mathbf{a})\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}-\varphi_{2}(x, \mathbf{a})\right) e^{-\frac{1}{2 \Delta}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}-\varphi_{2}(x, \mathbf{a})\right)^{2}}\right|}{\Delta \int d P_{A}^{(2)}(\mathbf{a}) e^{-\frac{1}{2}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}-\varphi(x, \mathbf{a})\right)^{2}}} \\
& \leq\left\|\frac{\varphi_{2}^{\prime}}{\sqrt{\Delta}}\right\|_{\infty}\left(\left|Z_{\mu}\right|+2\left\|\frac{\varphi_{2}}{\sqrt{\Delta}}\right\|_{\infty}\right) . \tag{6.53}
\end{align*}
$$

Hence, $\forall \mu \in\left\{1, \ldots, n_{2}\right\}$ :

$$
\begin{align*}
& \mathbb{E}\left[\left\langle\ell_{Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}}^{\prime}\left(s_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)\right\rangle_{\boldsymbol{n}, \epsilon}^{2}\right] \\
& \quad \leq\left\|\frac{\varphi_{2}^{\prime}}{\sqrt{\Delta}}\right\|_{\infty}^{2} \mathbb{E}\left[\left(\left|Z_{\mu}\right|+2\left\|\frac{\varphi_{2}}{\sqrt{\Delta}}\right\|_{\infty}\right)^{2}\right] \leq\left\|\frac{\varphi_{2}^{\prime}}{\sqrt{\Delta}}\right\|_{\infty}^{2}\left(2+8\left\|\frac{\varphi_{2}}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right) . \tag{6.54}
\end{align*}
$$

The first summand on the right-hand side of (6.51) is similar to the quantity $A_{n}(t, \epsilon)$ studied in Appendix 6.C. Proceeding exactly like in the latter appendix,
we obtain

$$
\begin{aligned}
&\left|\frac{1}{2} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} \mid S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)}{P_{\text {out }}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} \mid S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)} \frac{\ln \widetilde{\mathcal{Z}}_{n}(\epsilon)}{n_{0}}\right]\right| \\
&\left.\leq \frac{1}{2} \sqrt{n_{2}\left(4\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{4}+2\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right) \operatorname{Var}\left(\frac{\ln }{\widetilde{\mathcal{Z}}_{n}(\epsilon)}\right.} n_{0}\right)
\end{aligned} .
$$

Similarly to the proof of Proposition 6.14 in Appendix 6.D, we can show that there exists a constant $C$ such that $\forall \epsilon \in\left[0,2 s_{n_{0}}\right]^{3}$ :

$$
\operatorname{Var}\left(\frac{\ln \widetilde{\mathcal{Z}}_{n}(\epsilon)}{n_{0}}\right) \leq \frac{C}{n_{0}}
$$

Therefore, for $n_{0}, n_{2}$ large enough, $\forall \epsilon \in\left[0,2 s_{n_{0}}\right]^{3}$ :

$$
\begin{align*}
&\left|\frac{1}{2} \sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} \mid S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)}{P_{\text {out }}\left(Y_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)} \mid S_{\mu}^{\left(\epsilon_{2}, \epsilon_{3}\right)}\right)} \frac{\ln }{\ln } \frac{\widetilde{\mathcal{Z}}_{n}(\epsilon)}{n_{0}}\right]\right| \\
& \leq \sqrt{\alpha_{1} \alpha_{2} C\left(2\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{4}+\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right)} . \tag{6.55}
\end{align*}
$$

It follows from the bounds (6.54), (6.55) and the formulas (6.51), (6.52) that $\forall \epsilon \in\left[0,2 s_{n_{0}}\right]^{3}$ :

$$
\begin{aligned}
& \left|\frac{\partial \widetilde{f}_{n}}{\partial \epsilon_{2}}\right|_{\epsilon} \left\lvert\, \leq \frac{n_{2}}{2 n_{0}}\left\|\frac{\varphi_{2}^{\prime}}{\sqrt{\Delta}}\right\|_{\infty}^{2}\left(2+8\left\|\frac{\varphi_{2}}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right)+\sqrt{\alpha_{1} \alpha_{2} C\left(2\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{4}+\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right)}\right. \\
& \left|\frac{\partial \widetilde{f}_{n}}{\partial \epsilon_{3}}\right|_{\epsilon} \left\lvert\, \leq \sqrt{\alpha_{1} \alpha_{2} C\left(2\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{4}+\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right)} .\right.
\end{aligned}
$$

By the mean value theorem, and the upper bounds on the absolute values of the partial derivatives of $\widetilde{f}_{\boldsymbol{n}}$, for $n_{0}, n_{1}, n_{2}$ large enough, we have $\forall \epsilon \in\left[0,2 s_{n_{0}}\right]^{3}$ :

$$
\begin{aligned}
&\left|\widetilde{f}_{n}(\epsilon)-\widetilde{f}_{\boldsymbol{n}}(0)\right| \leq \alpha_{1}\left\|\varphi_{1}\right\|_{\infty}\left|\epsilon_{1}\right|+\alpha_{1} \alpha_{2}\left\|\frac{\varphi_{2}^{\prime}}{\sqrt{\Delta}}\right\|_{\infty}^{2}\left(2+8\left\|\frac{\varphi_{2}}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right)\left|\epsilon_{2}\right| \\
&+\sqrt{\alpha_{1} \alpha_{2} C\left(2\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{4}+\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right)}\left(\left|\epsilon_{2}\right|+\left|\epsilon_{3}\right|\right) .
\end{aligned}
$$

Hence, $\widetilde{f}_{\boldsymbol{n}}(\epsilon)=\widetilde{f}_{\boldsymbol{n}}(0)+O\left(s_{n_{0}}\right)$ where $\left|O\left(s_{n_{0}}\right) / s_{n_{0}}\right|$ is bounded uniformly in $\boldsymbol{n}$ and $\epsilon \in\left[0,2 s_{n_{0}}\right]^{3}$. Finally, note that $\widetilde{f}_{n}(0)=f_{n}-\frac{n_{1}}{2 n_{0}}$.

## 6.C Derivative of the averaged interpolating free entropy

In this appendix we compute the partial derivative of the average free entropy $f_{n}(t, \epsilon)$ with respect to $t$.

Proposition 6.4 (Derivative of the interpolating average free entropy). Assume that (H1), (H2), (H3) hold and $n_{0}, n_{1}, n_{2} \rightarrow+\infty$ such that $n_{2} / n_{1} \rightarrow \alpha_{2}, n_{1} / n_{0} \rightarrow \alpha_{1}$. Let $f_{n}(t, \epsilon)$ be the interpolating average free entropy defined by 6.22). Then, the derivative of $f_{\boldsymbol{n}}(\cdot, \epsilon)$, denoted by $f_{n}^{\prime}(\cdot, \epsilon)$, satisfies $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\begin{array}{r}
f_{\boldsymbol{n}}^{\prime}(t, \epsilon)=-\frac{1}{2} \frac{n_{1}}{n_{0}} \mathbb{E}\left\langle\left(\frac{1}{n_{1}} \sum_{\mu=1}^{n_{2}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-r_{\epsilon}(t)\right)\left(Q-q_{\epsilon}(t)\right)\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
 \tag{6.28}\\
+\frac{n_{1}}{n_{0}} \frac{r_{\epsilon}(t)}{2}\left(q_{\epsilon}(t)-\rho_{1}\left(n_{0}\right)\right)+o_{\boldsymbol{n}}(1),
\end{array}
$$

where $\ell_{y}^{\prime}(\cdot)$ is the derivative of $\ell_{y}: x \mapsto \ln P_{\text {out }}(y \mid x), o_{\boldsymbol{n}}(1)$ is a quantity that vanishes uniformly in $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ when $n_{0} \rightarrow+\infty$, and $Q:=\frac{\left(\mathbf{x}^{(1)}\right)^{\top} \mathbf{X}^{(1)}}{n_{1}}$. Proof. Remember that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}: f_{\boldsymbol{n}}(t, \epsilon):=\frac{\mathbb{E} \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}$ with

$$
\mathcal{Z}_{\boldsymbol{n}}(t, \epsilon):=\int d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a}) \mathcal{D} \mathbf{u} e^{-\mathcal{H} t, \epsilon\left(\mathbf{x}, \mathbf{a}, \mathbf{u} ; \mathbf{Y}^{(t, \epsilon)}, \tilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)}
$$

where $\mathcal{D} \mathbf{u}:=d \mathbf{u} e^{-\frac{\|\mathbf{u}\|^{2}}{2}} / \sqrt{2 \pi^{n}}$,

$$
\begin{aligned}
\mathcal{H}_{t, \epsilon}\left(\mathbf{x}, \mathbf{a}, \mathbf{u} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right):= & -\sum_{\mu=1}^{n_{2}} \ln P_{\mathrm{out}}\left(y_{\mu} \mid s_{\mu}^{(t, \epsilon)}\right) \\
& +\frac{1}{2} \sum_{i=1}^{n_{1}}\left(\widetilde{y}_{i}-\sqrt{R_{1}(t, \epsilon)} x_{i}^{(1)}\right)^{2},
\end{aligned}
$$

and $x_{i}^{(1)}:=\varphi_{1}\left(\left[\mathbf{W}^{(1)} \mathbf{x} / \sqrt{n_{0}}\right]_{i}, \mathbf{a}_{i}\right)$,

$$
s_{\mu}^{(t, \epsilon)}:=\sqrt{\frac{1-t}{n_{1}}}\left[\mathbf{W}^{(2)} \mathbf{x}^{(1)}\right]_{\mu}+\sqrt{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t-R_{2}(t, \epsilon)} u_{\mu}+\sqrt{R_{2}(t, \epsilon)} V_{\mu} .
$$

The posterior distribution of $\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)$ is

$$
\begin{aligned}
d P\left(\mathbf{x}, \mathbf{a}, \mathbf{u} \mid \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right. & \left., \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right) \\
& \propto d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a}) \mathcal{D} \mathbf{u} e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{x}, \mathbf{a}, \mathbf{u} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)}
\end{aligned}
$$

We denote by ( $\mathbf{x}, \mathbf{a}_{1}, \mathbf{u}$ ) a triplet sampled from this joint posterior distribution and by the angular brackets $\langle-\rangle_{n, t, \epsilon}$ the expectation with respect to this same distribution.

Computation of the derivative The conditional probability density function of $\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)$ given $\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)$ is
$f\left(\mathbf{Y}^{(t, \epsilon)}=\mathbf{y}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}=\widetilde{\mathbf{y}} \mid \mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right):=\frac{e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U} ; \mathbf{y}, \tilde{\mathbf{y}}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)}}{\sqrt{2 \pi^{n_{1}}}}$.

Therefore, the average free entropy reads

$$
\begin{aligned}
& f_{\boldsymbol{n}}(t, \epsilon)=\frac{1}{n_{0}} \mathbb{E}\left[\int \frac{d \mathbf{y} d \widetilde{\mathbf{y}}}{{\sqrt{2 \pi^{n}}}^{n_{1}}} e^{-\mathcal{H} t, \epsilon\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U} ; \mathbf{y}, \tilde{\mathbf{y}}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)}\right. \\
&\left.\cdot \ln \int d P_{X}(\mathbf{x}) d P_{A[1]}(\mathbf{a}) \mathcal{D} \mathbf{u} e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{x}, \mathbf{a}, \mathbf{u}, \mathbf{y}, \tilde{\mathbf{y}}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)}\right] .
\end{aligned}
$$

where the expectation $\mathbb{E}$ is with respect to $\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)$. Differentiating this last identity with respect to $t$ under the expectation sign yields

$$
\begin{align*}
f_{\boldsymbol{n}}^{\prime}(t, \epsilon)=-\frac{1}{n_{0}} \mathbb{E}[ & \left.\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
& -\frac{1}{n_{0}} \mathbb{E}\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{x}, \mathbf{a}, \mathbf{u} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon} \tag{6.56}
\end{align*}
$$

where $\mathcal{H}_{t, \epsilon}^{\prime}$ is defined as $\left(\ell_{y}^{\prime}(\cdot)\right.$ is the derivative of $\left.\ell_{y}: x \mapsto \ln P_{\text {out }}(y \mid x)\right)$

$$
\begin{aligned}
\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{x}, \mathbf{a}, \mathbf{u} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right):= & -\sum_{\mu=1}^{n_{2}} \frac{\partial s_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{y_{\mu}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right) \\
& -\frac{1}{2} \frac{r_{\epsilon}(t)}{\sqrt{R_{1}(t, \epsilon)}} \sum_{i=1}^{n_{1}} x_{i}^{(1)}\left(\widetilde{y}_{i}-\sqrt{R_{1}(t, \epsilon)} x_{i}^{(1)}\right) .
\end{aligned}
$$

Evaluating the latter at $(\mathbf{x}, \mathbf{a}, \mathbf{u}, \mathbf{Y}, \tilde{\mathbf{Y}})=\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U}, \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)$, we obtain

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)= & -\sum_{\mu=1}^{n_{2}} \frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \\
& -\frac{r_{\epsilon}(t)}{2 \sqrt{R_{1}(t, \epsilon)}} \sum_{i=1}^{n_{1}} X_{i}^{(1)} \widetilde{Z}_{i} \tag{6.57}
\end{align*}
$$

First expectation on the r.h.s. of 6.56 Using (6.57), the first term on the r.h.s. of (6.56) reads

$$
\begin{align*}
E_{1} & :=-\frac{1}{n_{0}} \mathbb{E}\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
& =\sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right]+\frac{r_{\epsilon}(t)}{2 \sqrt{R_{1}(t, \epsilon)}} \sum_{i=1}^{n_{1}} \mathbb{E}\left[X_{i}^{(1)} \widetilde{Z}_{i} \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \tag{6.58}
\end{align*}
$$

From the definition (6.18) of $S_{\mu}^{(t, \epsilon)}, \forall \mu \in\left\{1, \ldots, n_{2}\right\}$ :

$$
\begin{align*}
& \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right]=-\frac{1}{2} \mathbb{E}\left[\frac{\left[\mathbf{W}^{(2)} \mathbf{X}^{(1)}\right]}{\sqrt{n_{1}(1-t)}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
& \quad+\frac{1}{2} \mathbb{E}\left[\left(\frac{q_{\epsilon}(t) V_{\mu}}{\sqrt{R_{2}(t, \epsilon)}}+\frac{\left(\rho_{1}\left(n_{0}\right)-q_{\epsilon}(t)\right) U_{\mu}}{\sqrt{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t-R_{2}(t, \epsilon)}}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \tag{6.59}
\end{align*}
$$

A Gaussian integration by parts w.r.t $W_{\mu i}^{(2)}, 1 \leq i \leq n_{1}$, yields

$$
\begin{align*}
& \mathbb{E}\left[\frac{\left[\mathbf{W}^{(2)} \mathbf{X}^{(1)}\right]}{\sqrt{n_{1}(1-t)}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
& =\sum_{i=1}^{n_{1}} \mathbb{E}\left[\int \frac{d \mathbf{y} d \widetilde{\mathbf{y}}}{\sqrt{2 \pi}^{n_{1}}} e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)} \frac{W_{\mu i}^{(2)} X_{i}^{(1)}}{\sqrt{n_{1}(1-t)}} \ell_{y_{\mu}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
& =\frac{1}{\sqrt{n_{1}(1-t)}} \sum_{i=1}^{n_{1}} \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial W_{\mu i}^{(2)}} X_{i}^{(1)}\left(\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime \prime}\left(S_{\mu}^{(t, \epsilon)}\right)+\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right)^{2}\right) \ln \mathcal{Z}_{n}(t, \epsilon)\right] \\
& +\frac{1}{\sqrt{n_{1}(1-t)}} \sum_{i=1}^{n_{1}} \mathbb{E}\left\langle\frac{\partial s_{\mu}^{(t, \epsilon)}}{\partial W_{\mu i}^{(2)}} X_{i}^{(1)} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right\rangle_{n, t, \epsilon} \\
& =\mathbb{E}\left[\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)} \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
& +\mathbb{E}\left\langle Q \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon}, \tag{6.60}
\end{align*}
$$

where the last equality follows simply from $\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial W_{\mu i}^{(2)}}=\sqrt{\frac{1-t}{n_{1}}} X_{i}^{(1)}, \frac{\partial s_{\mu}^{(t, \epsilon)}}{\partial W_{\mu i}^{(2)}}=\sqrt{\frac{1-t}{n_{1}}} x_{i}^{(1)}$, and the identity

$$
\ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}=\frac{P_{\text {out }}^{\prime \prime}(y \mid x)}{P_{\text {out }}(y \mid x)}
$$

with $P_{\text {out }}^{\prime \prime}(y \mid \cdot)$ the second derivative of $x \mapsto P_{\text {out }}(y \mid x)$. To simplify the second expectation on the r.h.s. of (6.59), we do another integration by parts, this time w.r.t. $V_{\mu}, U_{\mu} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. It comes

$$
\begin{gather*}
\mathbb{E}\left[\left(\frac{q_{\epsilon}(t) V_{\mu}}{\sqrt{R_{2}(t, \epsilon)}}+\frac{\left(\rho_{1}\left(n_{0}\right)-q_{\epsilon}(t)\right) U_{\mu}}{\sqrt{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t-R_{2}(t, \epsilon)}}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
=\mathbb{E}\left[\int \frac{d \mathbf{y} d \widetilde{\mathbf{y}}}{\sqrt{2 \pi}^{n_{1}}} e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U}, \mathbf{y}, \mathbf{y}, \tilde{\mathbf{y}}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)}\right. \\
\left.\cdot\left(\frac{q_{\epsilon}(t) V_{\mu}}{\sqrt{R_{2}(t, \epsilon)}}+\frac{\left(\rho_{1}\left(n_{0}\right)-q_{\epsilon}(t)\right) U_{\mu}}{\sqrt{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right) t-R_{2}(t, \epsilon)}}\right) \ell_{y_{\mu}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
=\mathbb{E}\left[\rho_{1}\left(n_{0}\right) \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)} \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
+\mathbb{E}\left\langle q_{\epsilon}(t) \ell_{Y_{\mu}^{\prime}}^{(t, \epsilon)}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{\prime}}^{\prime(t, \epsilon)}\left(s_{\mu}^{(t, \epsilon)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon} . \tag{6.61}
\end{gather*}
$$

Plugging (6.60) and (6.61) back in (6.59) gives

$$
\begin{aligned}
& \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right] \\
& \quad=-\frac{1}{2} \mathbb{E}\left[\frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right) \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right]
\end{aligned}
$$

$$
\begin{equation*}
-\frac{1}{2} \mathbb{E}\left\langle\left(Q-q_{\epsilon}(t)\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon} \tag{6.62}
\end{equation*}
$$

We now simplify $\sum_{i=1}^{n_{1}} \mathbb{E}\left[X_{i}^{(1)} \widetilde{Z}_{i} \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right]$. A Gaussian integration by parts w.r.t. the standard Gaussian random variable $\widetilde{Z}_{i}$ yields

$$
\begin{gather*}
\sum_{i=1}^{n_{1}} \mathbb{E}\left[X_{i}^{(1)} \widetilde{Z}_{i} \ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)\right]=-\sum_{i=1}^{n_{1}} \mathbb{E}\left[X_{i}^{(1)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)+\widetilde{Z}_{i}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
=-\sqrt{R_{1}(t, \epsilon)}\left(n_{1} \rho_{1}\left(n_{0}\right)-\mathbb{E}\left\langle\left(\mathbf{x}^{(1)}\right)^{\top} \mathbf{X}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right)-\sum_{i=1}^{n_{1}} \mathbb{E}\left[X_{i}^{(1)} \widetilde{Z}_{i}\right] \\
=-n_{1} \sqrt{R_{1}(t, \epsilon)}\left(\rho_{1}\left(n_{0}\right)-\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}\right) \tag{6.63}
\end{gather*}
$$

Finally, we plug (6.62) and (6.63) in (6.58) to obtain

$$
\begin{align*}
E_{1}=- & \frac{1}{2} \mathbb{E}\left[\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right) \frac{\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}\right] \\
& -\frac{1}{2} \frac{n_{1}}{n_{0}} \mathbb{E}\left\langle\left(\frac{1}{n_{1}} \sum_{\mu=1}^{n_{2}} \ell_{\left.Y_{\mu}^{\prime}, \epsilon\right)}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{\prime(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-r_{\epsilon}(t)\right)\left(Q-q_{\epsilon}(t)\right)\right\rangle_{n, t, \epsilon} \\
& -\frac{n_{1}}{n_{0}} \frac{r_{\epsilon}(t)}{2}\left(\rho_{1}\left(n_{0}\right)-q_{\epsilon}(t)\right) . \tag{6.64}
\end{align*}
$$

Second expectation on the r.h.s. of (6.56 Let us show that the second expectation on the r.h.s. of (6.56) is zero. By the Nishimori identity (see Lemma 2.1),

$$
\begin{align*}
E_{2} & =\mathbb{E}\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{x}, \mathbf{a}, \mathbf{u} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
& =\mathbb{E}\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}, \mathbf{A}^{(1)}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)\right] \\
& =-\sum_{\mu=1}^{n_{2}} \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{\prime}}^{(t, \epsilon)}\left(S_{\mu}^{(t, \epsilon)}\right)\right]-\frac{r_{\epsilon}(t)}{2 \sqrt{R_{1}(t, \epsilon)}} \sum_{i=1}^{n_{1}} \mathbb{E}\left[X_{i}^{(1)} \widetilde{Z}_{i}\right] \\
& =-\frac{1}{2} \mathbb{E}\left[\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right)\right], \tag{6.65}
\end{align*}
$$

where the third equality is due to (6.57), and the fourth equality to $\mathbb{E}\left[X_{i}^{(1)} \widetilde{Z}_{i}\right]=0$ and a simplification of $E\left[\left(\partial S_{\mu}^{(t, \epsilon)} / \partial t\right) \ell_{\left.Y_{\mu}^{\prime}, \epsilon\right)}^{\prime(t,)}\left(S_{\mu}^{(t, \epsilon)}\right)\right]$ based on the same Gaussian integration by parts than the ones leading to (6.62). A direct computation shows that $\int P_{\text {out }}^{\prime \prime}(y \mid s) d y=0$ for all $s \in \mathbb{R}$. Hence, $\forall \mu \in\left\{1, \ldots, n_{2}\right\}$ :

$$
\begin{equation*}
\mathbb{E}\left[\left.\frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)} \right\rvert\, \mathbf{X}^{(1)}, \mathbf{S}^{(t, \epsilon)}\right]=\int d y P_{\text {out }}^{\prime \prime}\left(y \mid S_{\mu}^{(t, \epsilon)}\right)=0 . \tag{6.66}
\end{equation*}
$$

Therefore, by the tower property of the conditional expectation, we have

$$
\begin{align*}
& \mathbb{E}\left[\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right)\right] \\
& =\mathbb{E}\left[\mathbb{E}\left[\left.\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)} \right\rvert\, \mathbf{X}^{(1)}, \mathbf{S}^{(t, \epsilon)}\right]\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right)\right]=0 . \tag{6.67}
\end{align*}
$$

Combining (6.65) and (6.67), we get $E_{2}=0$.

Exact expression for $f_{\boldsymbol{n}}(t, \epsilon)$ We have shown that the second expectation on the right-hand side of 6.56 is zero. Therefore, $\forall(t, \epsilon) \in(0,1) \times \mathcal{B}_{n_{0}}$ :

$$
\begin{align*}
& f_{\boldsymbol{n}}(t, \epsilon)= E_{1} \\
&=-\frac{1}{2} \frac{n_{1}}{n_{0}} \mathbb{E}\left\langle\left(\frac{1}{n_{1}} \sum_{\mu=1}^{n_{2}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-r_{\epsilon}(t)\right)\left(Q-q_{\epsilon}(t)\right)\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
&+\frac{n_{1}}{n_{0}} \frac{r_{\epsilon}(t)}{2}\left(q_{\epsilon}(t)-\rho_{1}\left(n_{0}\right)\right)-\frac{A_{\boldsymbol{n}}(t, \epsilon)}{2}, \tag{6.68}
\end{align*}
$$

where

$$
\begin{equation*}
A_{\boldsymbol{n}}(t, \epsilon):=\mathbb{E}\left[\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right) \frac{\ln \mathcal{Z}_{n}(t, \epsilon)}{n_{0}}\right] . \tag{6.69}
\end{equation*}
$$

Next we show that $A_{\boldsymbol{n}}(t, \epsilon)$ goes to 0 uniformly in $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ when $\boldsymbol{n} \rightarrow+\infty$, thus ending the proof of the proposition.

Proof that $A_{\boldsymbol{n}}(t, \epsilon)$ vanishes uniformly It follows directly from (6.67) that

$$
\mathbb{E}\left[\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right) f_{n}(t, \epsilon)\right]=0 .
$$

Thus, we have

$$
\begin{align*}
\left|A_{\boldsymbol{n}}(t, \epsilon)\right|=\mid \mathbb{E} & { \left.\left[\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right)\left(\frac{\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}-f_{\boldsymbol{n}}(t, \epsilon)\right)\right] \right\rvert\, } \\
\leq & \mathbb{E}\left[\left(\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right)^{2}\right]^{\frac{1}{2}} \\
& \cdot \mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}-f_{\boldsymbol{n}}(t, \epsilon)\right)^{2}\right]^{\frac{1}{2}} \tag{6.70}
\end{align*}
$$

where the inequality is due to Cauchy-Schwarz inequality. By the tower property of conditional expectation, the first expectation on the r.h.s. of (6.70) reads

$$
\begin{align*}
& \mathbb{E}\left[\left(\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right)^{2}\right] \\
& =\mathbb{E}\left[\mathbb{E}\left[\left.\left(\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2} \right\rvert\, \mathbf{X}^{(1)}, \mathbf{S}^{(t, \epsilon)}\right]\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right)^{2}\right] . \tag{6.71}
\end{align*}
$$

Conditionally on $\mathbf{S}^{(t, \epsilon)}$, the random variables $\frac{P_{\text {out }}^{\prime \prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}, \mu \in\left\{1, \ldots, n_{2}\right\}$, are i.i.d. and centered. Therefore,

$$
\begin{align*}
\mathbb{E}\left[\left.\left(\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2} \right\rvert\, \mathbf{X}^{(1)}, \mathbf{S}^{(t, \epsilon)}\right] & =\mathbb{E}\left[\left.\left(\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2} \right\rvert\, \mathbf{S}^{(t, \epsilon)}\right] \\
& =n_{2} \mathbb{E}\left[\left.\left(\frac{P_{\text {out }}^{\prime \prime}\left(Y_{1}^{(t, \epsilon)} \mid S_{1}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{1}^{(t, \epsilon)} \mid S_{1}^{(t, \epsilon)}\right)}\right)^{2} \right\rvert\, S_{1}^{(t, \epsilon)}\right] \\
& =n_{2} \int \frac{P_{\text {out }}^{\prime \prime}\left(y \mid S_{1}^{(t, \epsilon)}\right)^{2}}{P_{\text {out }}\left(y \mid S_{1}^{(t, \epsilon)}\right)} d y . \tag{6.72}
\end{align*}
$$

Under the assumption (H2), it is not difficult to show that (we refer to the proof in Appendix 7.E.1 of Chapter 7)

$$
\begin{equation*}
0 \leq \int \frac{P_{\text {out }}^{\prime \prime}\left(y \mid S_{1}^{(t, \epsilon)}\right)^{2}}{P_{\text {out }}\left(y \mid S_{1}^{(t, \epsilon)}\right)} d y \leq 4\left\|\frac{\varphi_{2}^{\prime}}{\sqrt{\Delta}}\right\|_{\infty}^{4}+2\left\|\frac{\varphi_{2}^{\prime \prime}}{\sqrt{\Delta}}\right\|_{\infty}^{2}=: C . \tag{6.73}
\end{equation*}
$$

Combining (6.71), (6.72) and (6.73), we obtain

$$
\begin{equation*}
\mathbb{E}\left[\left(\sum_{\mu=1}^{n_{2}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right)^{2}\right] \leq C \frac{n_{2}}{n_{1}^{2}} \operatorname{Var}\left\|\mathbf{X}^{(1)}\right\|^{2} \tag{6.74}
\end{equation*}
$$

Let us prove that $\operatorname{Var}\left\|\mathbf{X}^{(1)}\right\|^{2} / n_{1}$ is bounded in the high-dimensional limit. We have

$$
\begin{equation*}
\operatorname{Var}\left\|\mathbf{X}^{(1)}\right\|^{2}=\mathbb{E}\left[\operatorname{Var}\left(\left\|\mathbf{X}^{(1)}\right\|^{2} \mid \mathbf{X}\right)\right]+\operatorname{Var}\left(\mathbb{E}\left[\left\|\mathbf{X}^{(1)}\right\|^{2} \mid \mathbf{X}\right]\right) \tag{6.75}
\end{equation*}
$$

Remember that $\mathbf{X}^{(1)}=\varphi_{1}\left(\mathbf{W}^{(1)} \mathbf{X} / \sqrt{n_{0}}, \mathbf{A}^{(1)}\right)$ so, conditionally on $\mathbf{X}$, the random variables $X_{i}^{(1)}, i \in\left\{1, \ldots, n_{1}\right\}$, are i.i.d. and it comes

$$
\operatorname{Var}\left(\left\|\mathbf{X}^{(1)}\right\|^{2} \mid \mathbf{X}\right)=\sum_{i=1}^{n_{1}} \operatorname{Var}\left(\left(X_{i}^{(1)}\right)^{2} \mid \mathbf{X}\right)=n_{1} \operatorname{Var}\left(\left(X_{1}^{(1)}\right)^{2} \mid \mathbf{X}\right)
$$

It follows that the first term on the r.h.s. of (6.75) satisfies

$$
\mathbb{E}\left[\operatorname{Var}\left(\left\|\mathbf{X}^{(1)}\right\|^{2} \mid \mathbf{X}\right)\right]=\mathbb{E}\left[\operatorname{Var}\left(\left(X_{1}^{(1)}\right)^{2} \mid \mathbf{X}\right)\right] \leq n_{1} \operatorname{Var}\left(\left(X_{1}^{(1)}\right)^{2}\right) \leq n_{1} \mathbb{E}\left[\left(X_{1}^{(1)}\right)^{4}\right]
$$

Under the assumption (H2), we thus have $\mathbb{E}\left[\operatorname{Var}\left(\left\|\mathbf{X}^{(1)}\right\|^{2} \mid \mathbf{X}\right)\right] \leq n_{1}\left\|\varphi_{1}\right\|_{\infty}^{4}$. For the second term on the r.h.s. of 6.75), first note that

$$
\begin{equation*}
\mathbb{E}\left[\left\|\mathbf{X}^{(1)}\right\|^{2} \mid \mathbf{X}\right]=n_{1} \mathbb{E}\left[\left.\varphi_{1}^{2}\left(\left[\frac{\mathbf{W}^{(1)} \mathbf{X}}{\sqrt{n_{0}}}\right]_{1}, \mathbf{A}_{1}^{(1)}\right) \right\rvert\, \mathbf{X}\right]=n_{1} g\left(X_{1}, \ldots, X_{n_{0}}\right), \tag{6.76}
\end{equation*}
$$

where $g(\mathbf{c}):=\mathbb{E}\left[\varphi_{1}^{2}\left(\left[\mathbf{W}^{(1)} \mathbf{c} / \sqrt{n_{0}}\right]_{1}, \mathbf{A}_{1}^{(1)}\right)\right]$ for any $\mathbf{c}=\left(c_{1}, \ldots, c_{n_{0}}\right) \in \mathbb{R}^{n_{0}}$. The partial derivatives of $g$ are $\forall \in\left\{1, \ldots, n_{0}\right\}$ :

$$
\begin{aligned}
\frac{\partial g}{\partial c_{j}} & =\mathbb{E}\left[2 \varphi_{1}\left(\left[\frac{\mathbf{W}^{(1)} \mathbf{c}}{\sqrt{n_{0}}}\right]_{1}, \mathbf{A}_{1}^{(1)}\right) \varphi_{1}^{\prime}\left(\left[\frac{\mathbf{W}^{(1)} \mathbf{c}}{\sqrt{n_{0}}}\right]_{1}, \mathbf{A}_{1}^{(1)}\right) \frac{W_{1 j}^{(1)}}{\sqrt{n_{0}}}\right] \\
& =\frac{2 c_{j}}{n_{0}} \mathbb{E}\left[\varphi_{1}^{\prime}\left(\left[\frac{\mathbf{W}^{(1)} \mathbf{c}}{\sqrt{n_{0}}}\right]_{1}, \mathbf{A}_{1}^{(1)}\right)^{2}+\varphi_{1}\left(\left[\frac{\mathbf{W}^{(1)} \mathbf{c}}{\sqrt{n_{0}}}\right]_{1}, \mathbf{A}_{1}^{(1)}\right) \varphi_{1}^{\prime \prime}\left(\left[\frac{\mathbf{W}^{(1)} \mathbf{c}}{\sqrt{n_{0}}}\right]_{1}, \mathbf{A}_{1}^{(1)}\right)\right],
\end{aligned}
$$

where the second equality is obtained by integrating by parts w.r.t. $W_{1 j}^{(1)}$. Under the assumption (H1), the support of the probability distribution $P_{X}$ is bounded and included in $[-S, S]$. For every $\mathbf{c} \in[-S, S]^{n_{0}}$ :

$$
\left|\frac{\partial g}{\partial c_{j}}\right|_{\mathbf{c}} \left\lvert\, \leq \frac{2 S}{n_{0}}\left(\left\|\varphi_{1}^{\prime}\right\|_{\infty}^{2}+\left\|\varphi_{1}\right\|_{\infty}\left\|\varphi_{1}^{\prime \prime}\right\|_{\infty}\right)=: \frac{C^{\prime}}{n_{0}},\right.
$$

where, under the assumption (H2), all of $\left\|\varphi_{1}\right\|_{\infty},\left\|\varphi_{1}^{\prime}\right\|_{\infty},\left\|\varphi_{1}^{\prime \prime}\right\|_{\infty}$ are finite. By the mean value theorem, $g$ satisfies the bounded difference property on $[-S, S]^{n_{0}}$, that is, $\forall j \in\left\{1, \ldots, n_{0}\right\}$ :

$$
\sup _{\substack{\mathbf{c} \in\left[-S, S S^{n} n_{0} \\ c_{j}^{\prime} \in[-S, S]\right.}}\left|g(\mathbf{c})-g\left(c_{1}, \ldots, c_{j}^{\prime}, \ldots, c_{n_{0}}\right)\right| \leq 2 S \frac{C^{\prime}}{n_{0}}
$$

By McDiarmid's inequality (see Proposition 2.6), it comes

$$
\operatorname{Var}(g(\mathbf{X})) \leq \frac{1}{4} \sum_{j=1}^{n_{0}}\left(\frac{2 S C^{\prime}}{n_{0}}\right)^{2}=\frac{\left(S C^{\prime}\right)^{2}}{n_{0}}
$$

Then, using (6.76),

$$
\operatorname{Var}\left(\mathbb{E}\left[\left\|\mathbf{X}^{(1)}\right\|^{2} \mid \mathbf{X}\right]\right)=n_{1}^{2} \operatorname{Var}(g(\mathbf{X})) \leq \frac{n_{1}^{2}\left(S C^{\prime}\right)^{2}}{n_{0}}
$$

Combining the latter with (6.75) and $\mathbb{E}\left[\operatorname{Var}\left(\left\|\mathbf{X}^{(1)}\right\|^{2} \mid \mathbf{X}\right)\right] \leq n_{1}\left\|\varphi_{1}\right\|_{\infty}^{4}$ yields

$$
\begin{equation*}
\frac{\operatorname{Var}\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}} \leq\left\|\varphi_{1}\right\|_{\infty}^{4}+\frac{n_{1}\left(S C^{\prime}\right)^{2}}{n_{0}} \leq\left\|\varphi_{1}\right\|_{\infty}^{4}+2 \alpha_{1}\left(S C^{\prime}\right)^{2}=: K \tag{6.77}
\end{equation*}
$$

where the last inequality is valid as long as $n_{0}, n_{1}$ are large enough so that $n_{1} / n_{0}$ is close to $\alpha_{1}$. We now combine (6.70), (6.74) and (6.77) to obtain

$$
\begin{equation*}
\left|A_{\boldsymbol{n}}(t, \epsilon)\right| \leq \sqrt{\frac{C K n_{2}}{n_{1}}} \mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}-f_{\boldsymbol{n}}(t, \epsilon)\right)^{2}\right]^{\frac{1}{2}} \tag{6.78}
\end{equation*}
$$

By Proposition 6.14 in Appendix 6.D, the variance $\mathbb{E}\left[\left(\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon) / n_{0}-f_{\boldsymbol{n}}(t, \epsilon)\right)^{2}\right]$ vanishes uniformly in $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ when $\boldsymbol{n} \rightarrow+\infty$. This theorem and (6.78) imply that $A_{\boldsymbol{n}}(t, \epsilon)$ vanishes uniformly in $(t, \epsilon)$.

## 6.D Concentration of the interpolating free entropy

In this appendix, we prove that the interpolating free entropy $\ln \mathcal{Z}_{n}(t, \epsilon) / n_{0}$ concentrates around its expectation (6.22) uniformly in $(t, \epsilon)$.

Proposition 6.14. Assume that (H1), (H2), (H3) and $n_{0}, n_{1}, n_{2} \rightarrow+\infty$ such that $n_{2} / n_{1} \rightarrow \alpha_{2}, n_{1} / n_{0} \rightarrow \alpha_{1}$. There exists a positive constant $C$, depending only on $\left(\varphi_{1}, \varphi_{2}, S, \Delta, \alpha_{1}, \alpha_{2}\right)$, such that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}-\mathbb{E}\left[\frac{\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}\right]\right)^{2}\right] \leq \frac{C}{n_{0}} \tag{6.79}
\end{equation*}
$$

Proof. Let us first rewrite $\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon) / n_{0}$. Remember that $P_{\text {out }}$ is defined by (6.6) in Subsection 6.2.1. We have

$$
\begin{align*}
P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid s_{\mu}^{(t, \epsilon)}\right) & =\int d P_{A}^{(2)}\left(\mathbf{a}_{\mu}^{(2)}\right) \frac{1}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}\left(Y_{\mu}^{(t, \epsilon)}-\varphi_{2}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right)\right)^{2}} \\
& =\int d P_{A}^{(2)}\left(\mathbf{a}_{\mu}^{(2)}\right) \frac{1}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}\left(\Gamma_{\mu}^{(t, \epsilon)}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right)+\sqrt{\Delta} Z_{\mu}\right)^{2}} \tag{6.80}
\end{align*}
$$

where

$$
\begin{equation*}
\Gamma_{\mu}^{(t, \epsilon)}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right):=\varphi_{2}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)-\varphi_{2}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right) \tag{6.81}
\end{equation*}
$$

Define where

$$
\begin{equation*}
\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}=\frac{1}{n_{0}} \ln \int d P_{X}(\mathbf{x}) d P_{A}^{(1)}(\mathbf{a}) d P_{A}^{(2)}\left(\mathbf{a}^{(2)}\right) \mathcal{D} \mathbf{u} e^{-\widehat{\mathcal{H}}_{t, \epsilon}\left(\mathbf{x}, \mathbf{a}, \mathbf{a}\left(\mathbf{a}^{(2)}, \mathbf{u}\right)\right.} \tag{6.82}
\end{equation*}
$$

where

$$
\begin{align*}
& \widehat{\mathcal{H}}_{t, \epsilon}\left(\mathbf{x}, \mathbf{a}, \mathbf{a}^{(2)}, \mathbf{u}\right):=\frac{1}{2 \Delta} \sum_{\mu=1}^{n_{2}}\left(\Gamma_{\mu}^{(t, \epsilon)}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right)^{2}+2 \sqrt{\Delta} Z_{\mu} \Gamma_{\mu}^{(t, \epsilon)}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right)\right) \\
& \quad+\frac{1}{2} \sum_{i=1}^{n_{1}} R_{1}(t, \epsilon)\left(X_{i}^{(1)}-x_{i}^{(1)}\right)^{2}+2 \widetilde{Z}_{i} \sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right) \tag{6.83}
\end{align*}
$$

A comparison with the interpolating Hamiltonian (6.21) defined in Subsection 6.3.1 shows that

$$
\frac{\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}=\frac{\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon)}{n_{0}}-\frac{1}{2 n_{0}} \sum_{\mu=1}^{n_{2}} Z_{\mu}^{2}-\frac{1}{2 n_{0}} \sum_{i=1}^{n_{1}} \widetilde{Z}_{i}^{2}-\frac{n_{2}}{2 n_{0}} \ln (2 \pi \Delta)
$$

Therefore,

$$
\begin{align*}
\operatorname{Var}\left(\frac{\ln \mathcal{Z}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}\right) & \leq 2 \operatorname{Var}\left(\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}\right)+2 \operatorname{Var}\left(\frac{1}{2 n_{0}} \sum_{\mu=1}^{n_{2}} Z_{\mu}^{2}+\frac{1}{2 n_{0}} \sum_{i=1}^{n_{1}} \widetilde{Z}_{i}^{2}\right) \\
& \leq 2 \operatorname{Var}\left(\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}\right)+\frac{n_{2}+n_{1}}{n_{0}^{2}} \tag{6.84}
\end{align*}
$$

In the remaing part of this appendix we show that the free entropy (6.82) concentrates around its expectation. This concentration together with the upper bound (6.84) yields the proposition.

Note that $\ln \hat{\mathcal{Z}}_{n}(t, \epsilon) / n_{0}$ has been written as a function of $\mathbf{Z}, \widetilde{\mathbf{Z}}, \mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}, \mathbf{A}^{(2)}$, $\mathbf{W}^{(1)}$ and $\mathbf{X}^{(1)}:=\varphi_{1}\left(\mathbf{W}^{(1)} \mathbf{X}^{(1)} / \sqrt{n_{0}}, \mathbf{A}^{(1)}\right)$. We prove that $\ln \hat{\mathcal{Z}}_{n}(t, \epsilon) / n_{0}$ concentrates with respect to each of these random variables. The order in which the concentrations are proved matters. In the end, the combination of Lemmas 6.15, 6.16, 6.17, 6.18, 6.19 and 6.20 stated below proves the existence of a positive constant $C$, depending only on $\left(\varphi_{1}, \varphi_{2}, S, \Delta, \alpha_{1}, \alpha_{2}\right)$, such that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\operatorname{Var}\left(\frac{\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon)}{n_{0}}\right) \leq \frac{C}{n_{0}}
$$

Lemma 6.15. Under the assumption of Proposition 6.14, there exists a positive constant $C$, depending only on $\left(\varphi_{1}, \varphi_{2}, S, \Delta, \alpha_{1}, \alpha_{2}\right)$, such that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$.

$$
\mathbb{E}\left[\left(\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}-\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}, \mathbf{A}^{(2)}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]\right)^{2}\right] \leq \frac{C}{n_{0}}
$$

Proof. We see that $g(\mathbf{Z}, \widetilde{\mathbf{Z}})=\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon) / n_{0}$ as a function of $\mathbf{Z}, \widetilde{\mathbf{Z}}$ only and we work conditionally to all other random variables, i.e., $\mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}, \mathbf{A}_{2}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}$. The squared norm of the gradient of $g$ reads

$$
\|\nabla g\|^{2}=\sum_{\mu=1}^{n_{2}}\left|\frac{\partial g}{\partial Z_{\mu}}\right|^{2}+\sum_{i=1}^{n_{1}}\left|\frac{\partial g}{\partial \widetilde{Z}_{i}}\right|^{2}
$$

Each of these partial derivatives is of the form $\partial g / \partial x=-n_{0}^{-1}\left\langle\partial \hat{\mathcal{H}}_{t, \epsilon} / \partial x\right\rangle_{\boldsymbol{n}, t, \epsilon}$ where the angular brackets $\langle-\rangle_{\boldsymbol{n}, t, \epsilon}$ denote an expectation with respect to the posterior distribution (6.20). We have

$$
\begin{aligned}
\left|\frac{\partial g}{\partial Z_{\mu}}\right| & =\frac{1}{n_{0} \sqrt{\Delta}}\left|\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right| \leq \frac{2\left\|\varphi_{2}\right\|_{\infty}}{n_{0} \sqrt{\Delta}} \\
\left|\frac{\partial g}{\partial \widetilde{Z}_{i}}\right| & =\frac{1}{n_{0}}\left|\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right)\right| \leq \frac{2 \sqrt{r_{\max }}\left\|\varphi_{1}\right\|_{\infty}}{n_{0}} .
\end{aligned}
$$

By the Gaussian-Poincaré inequality (see Proposition 2.7), we thus have

$$
\begin{aligned}
& \mathbb{E}\left[\left(\frac{\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon)}{n_{0}}-\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}, \mathbf{A}_{2}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]\right)^{2}\right] \\
& \leq \mathbb{E}\|\nabla g\|^{2} \leq \frac{4}{n_{0}}\left(\frac{n_{2}}{n_{0}} \frac{\left\|\varphi_{2}\right\|_{\infty}^{2}}{\Delta}+\frac{n_{1}}{n_{0}} r_{\max }\left\|\varphi_{1}\right\|_{\infty}^{2}\right) .
\end{aligned}
$$

Lemma 6.16. Under the assumption of Proposition 6.14, there exists a positive constant $C$, depending only on $\left(\varphi_{1}, \varphi_{2}, S, \Delta, \alpha_{1}, \alpha_{2}\right)$, such that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\begin{array}{r}
\mathbb{E}\left|\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}, \mathbf{A}^{(2)}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]-\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{A}^{(2)}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]\right|^{2} \\
\leq \frac{C}{n_{0}}
\end{array}
$$

Proof. We see $g\left(\mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}\right)=\mathbb{E}\left[\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon) \mid \mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}, \mathbf{A}^{(2)}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right] / n_{0}$ as a function of $\mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}$ only and we work conditionally to all other random variables, i.e., $\mathbf{A}^{(2)}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}$. From now on and until the end of the proof, we denote by $\tilde{\mathbb{E}}[\cdot]$ the conditional expectation $\mathbb{E}\left[\cdot \mid \mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}, \mathbf{A}^{(2)}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]$. We have

$$
\begin{align*}
\left|\frac{\partial g}{\partial V_{\mu}}\right| & =\frac{1}{n_{0} \Delta}\left|\tilde{\mathbb{E}}\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right) \frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial V_{\mu}}\right\rangle_{n, t, \epsilon}\right| \\
& \leq \frac{\sqrt{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right)}}{n_{0} \Delta} \tilde{\mathbb{E}}\left[\left(2\left\|\varphi_{2}\right\|_{\infty}+\sqrt{\Delta}\left|Z_{\mu}\right|\right) 2\left\|\varphi_{2}^{\prime}\right\|_{\infty}\right] \\
& \leq \frac{\sqrt{1+\rho_{1}\left(n_{0}\right)}}{n_{0} \Delta}\left(2\left\|\varphi_{2}\right\|_{\infty}+\sqrt{\frac{2 \Delta}{\pi}}\right) 2\left\|\varphi_{2}^{\prime}\right\|_{\infty} \tag{6.85}
\end{align*}
$$

The same inequality holds true for $\left|\partial g / \partial U_{\mu}\right|$. To compute the partial derivative w.r.t. $W_{\mu i}^{(2)}$, first note that

$$
\begin{equation*}
\frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial W_{\mu i}^{(2)}}=\sqrt{\frac{1-t}{n_{1}}}\left(X_{i}^{(1)} \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)-x_{i}^{(1)} \varphi_{2}^{\prime}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right)\right) \tag{6.86}
\end{equation*}
$$

Thus,

$$
\begin{align*}
\left|\frac{\partial g}{\partial W_{\mu i}^{(2)}}\right| & =\frac{1}{n_{0} \Delta}\left|\tilde{\mathbb{E}}\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right) \frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial W_{\mu i}^{(2)}}\right\rangle_{\hat{\mathcal{H}}_{t, \epsilon}}\right| \\
& \leq \frac{1}{n_{0} \sqrt{n_{1}} \Delta} \tilde{\mathbb{E}}\left[\left(2\left\|\varphi_{2}\right\|_{\infty}+\sqrt{\Delta}\left|Z_{\mu}\right|\right) 2\left\|\varphi_{1}\right\|_{\infty}\left\|\varphi_{2}^{\prime}\right\|_{\infty}\right] \\
& =\frac{1}{n_{0} \sqrt{n_{1}} \Delta}\left(2\left\|\varphi_{2}\right\|_{\infty}+\sqrt{\frac{2 \Delta}{\pi}}\right) 2\left\|\varphi_{1}\right\|_{\infty}\left\|\varphi_{2}^{\prime}\right\|_{\infty} \tag{6.87}
\end{align*}
$$

Combining (6.85) and (6.87), we get

$$
\begin{aligned}
\|\nabla g\|^{2} & =\sum_{\mu=1}^{n_{2}}\left|\frac{\partial g}{\partial V_{\mu}}\right|^{2}+\sum_{\mu=1}^{n_{2}}\left|\frac{\partial g}{\partial U_{\mu}}\right|^{2}+\sum_{\mu=1}^{n_{2}} \sum_{i=1}^{n_{1}}\left|\frac{\partial g}{\partial W_{\mu i}^{(2)}}\right|^{2} \\
& \leq \frac{n_{2}}{n_{0}^{2}}\left(2+2 \rho_{1}\left(n_{0}\right)+\left\|\varphi_{1}\right\|_{\infty}^{2}\right)\left(\frac{2\left\|\varphi_{2}^{\prime}\right\|_{\infty}}{\Delta}\right)^{2}\left(2\left\|\varphi_{2}\right\|_{\infty}+\sqrt{\frac{2 \Delta}{\pi}}\right)^{2}
\end{aligned}
$$

To end the proof, we apply the Gaussian-Poincaré inequality and use the latter upper bound on the squared norm of the gradient of $g$.

Lemma 6.17. Under the assumption of Proposition 6.14, there exists a positive constant $C$, depending only on $\left(\varphi_{1}, \varphi_{2}, S, \Delta, \alpha_{1}, \alpha_{2}\right)$, such that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{A}_{2}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]-\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]\right)^{2}\right] \leq \frac{C}{n_{0}} \tag{6.88}
\end{equation*}
$$

Proof. We see $g\left(\mathbf{A}^{(2)}\right)=\mathbb{E}\left[\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon) \mid \mathbf{A}^{(2)}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right] / n_{0}$ as a function of $\mathbf{A}^{(2)}$ only and we work conditionally to $\mathbf{X}^{(1)}, \mathbf{W}^{(1)}$. We denote by $\mathbb{E}_{G}$ the expectation w.r.t. the Gaussian random variables $\mathbf{Z}, \widetilde{\mathbf{Z}}, \mathbf{V}, \mathbf{U}, \mathbf{W}^{(2)}$, hence $g=E_{G}\left[\ln \hat{\mathcal{Z}}_{n}(t, \epsilon)\right] / n_{0}$.

We show that $g$ satisfies the bounded difference property. Fix $\nu \in\left\{1, \ldots, n_{2}\right\}$. We want to upper bound the variation $g\left(\mathbf{A}^{(2)}\right)-g\left(\mathbf{A}^{(2, \nu)}\right)$ for two configurations $\mathbf{A}^{(2)}$ and $\mathbf{A}^{(2, \nu)}$ such that $\forall \mu \neq \nu: \mathbf{A}_{\mu}^{(2, \nu)}=\mathbf{A}_{\mu}^{(2)}$. Denote $\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}$ and $\Gamma_{\mu}^{(t, \epsilon, \nu)}$ the quantities $\widehat{\mathcal{H}}_{t, \epsilon}$ and $\Gamma_{\mu}^{(t, \epsilon)}$ where $\mathbf{A}^{(2)}$ has been replaced by $\mathbf{A}^{(2, \nu)}$. We also distinguish with indices the angular brackets $\langle-\rangle_{\hat{\mathcal{H}}_{t, \epsilon}}$ and $\langle-\rangle_{\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}}$ associated with the two configurations. By Jensen's inequality,

$$
\begin{equation*}
\frac{1}{n_{0}} \mathbb{E}_{G}\left\langle\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}-\widehat{\mathcal{H}}_{t, \epsilon}\right\rangle_{\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}} \leq g\left(\mathbf{A}^{(2)}\right)-g\left(\mathbf{A}^{(2, \nu)}\right) \leq \frac{1}{n_{0}} \mathbb{E}_{G}\left\langle\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}-\widehat{\mathcal{H}}_{t, \epsilon}\right\rangle_{\widehat{\mathcal{H}}_{t, \epsilon}} . \tag{6.89}
\end{equation*}
$$

Making use of the definition (6.83) of $\widehat{\mathcal{H}}_{t, \epsilon}$, we have

$$
\begin{aligned}
\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}-\widehat{\mathcal{H}}_{t, \epsilon} & =\frac{1}{2 \Delta} \sum_{\mu=1}^{n_{2}}\left(\Gamma_{\mu}^{(t, \epsilon, \nu)}-\Gamma_{\mu}^{(t, \epsilon)}\right)\left(\Gamma_{\mu}^{(t, \epsilon, \nu)}+\Gamma_{\mu}^{(t, \epsilon)}+2 \sqrt{\Delta} Z_{\mu}\right) \\
& =\frac{1}{2 \Delta}\left(\Gamma_{\nu}^{(t, \epsilon, \nu)}-\Gamma_{\nu}^{(t, \epsilon)}\right)\left(\Gamma_{\nu}^{(t, \epsilon, \nu)}+\Gamma_{\nu}^{(t, \epsilon)}+2 \sqrt{\Delta} Z_{\nu}\right)
\end{aligned}
$$

Hence, $\left|\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}-\widehat{\mathcal{H}}_{t, \epsilon}\right| \leq \Delta^{-1}\left\|\varphi_{2}\right\|_{\infty}^{2}+2 \Delta^{-1 / 2}\left|Z_{\nu}\right|\left\|\varphi_{2}\right\|_{\infty}$. This inequality combined with (6.89) shows that $g$ satisfies the bounded difference property,

$$
\left|g\left(\mathbf{A}^{(2)}\right)-g\left(\mathbf{A}^{(2, \nu)}\right)\right| \leq \frac{\left\|\varphi_{2}\right\|_{\infty}}{\Delta n_{0}}\left(\left\|\varphi_{2}\right\|_{\infty}+2 \sqrt{\frac{2}{\pi} \Delta}\right)
$$

By McDiarmid's inequality (see Proposition 2.6), we thus have

$$
\begin{array}{r}
\mathbb{E}\left[\left.\left(\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{A}_{2}, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]-\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]\right)^{2} \right\rvert\, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right] \\
\leq \frac{n_{2}\left\|\varphi_{2}\right\|_{\infty}^{2}}{n_{0}^{2} \Delta^{2}}\left(\frac{\left\|\varphi_{2}\right\|_{\infty}}{2}+\sqrt{\frac{2}{\pi} \Delta}\right)^{2}
\end{array}
$$

almost surely. Taking the expectation on both sides ends the proof.
Lemma 6.18. Under the assumption of Proposition 6.14, there exists a positive constant $C$, depending only on $\left(\varphi_{1}, \varphi_{2}, S, \Delta, \alpha_{1}, \alpha_{2}\right)$, such that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]-\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{W}^{(1)}, \mathbf{X}\right]\right)^{2}\right] \leq \frac{C}{n_{0}} \tag{6.90}
\end{equation*}
$$

Proof. Note that $\mathbb{E}\left[\ln \hat{\mathcal{Z}}_{n}(t, \epsilon) / n_{0} \mid \mathbf{X}^{(1)}, \mathbf{W}^{(1)}\right]=\mathbb{E}\left[\ln \hat{\mathcal{Z}}_{n}(t, \epsilon) / n_{0} \mid \mathbf{X}^{(1)}, \mathbf{W}^{(1)}, \mathbf{X}\right]$ because $\ln \hat{\mathcal{E}}_{n}(t, \epsilon) / n_{0}$ depends on $\mathbf{X}$ only through $\mathbf{X}^{(1)}$. As $\mathbf{X}^{(1)}:=\varphi_{1}\left(\mathbf{W}^{(1)} \mathbf{X} / \sqrt{n_{0}}, \mathbf{A}^{(1)}\right)$ and $\mathbf{A}_{1}^{(1)}, \ldots, \mathbf{A}_{n_{1}}^{(1)}$ are i.i.d., the random variables $X_{1}^{(1)}, \ldots, X_{n_{1}}^{(1)}$ are i.i.d. conditionally on $\left(\mathbf{W}^{(1)}, \mathbf{X}\right)$. Define $g(\mathbf{c}):=\mathbb{E}\left[\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon) / n_{0} \mid \mathbf{X}^{(1)}=\mathbf{c}, \mathbf{W}^{(1)}, \mathbf{X}\right]$.

Let us show that $g$ satisfies the bounded difference property. Fix $i \in\left\{1, \ldots, n_{1}\right\}$. Consider two vectors $\mathbf{c}, \mathbf{c}^{(i)} \in\left[-\left\|\varphi_{1}\right\|_{\infty},\left\|\varphi_{1}\right\|_{\infty}\right]^{n_{1}}$ such that $\forall j \neq i: c_{j}^{(i)}=c_{j}$. Define $\psi: \theta \in[0,1] \mapsto g\left(\theta \mathbf{c}+(1-\theta) \mathbf{c}^{(i)}\right)$, so that $\psi(1)=g(\mathbf{c})$ and $\psi(0)=g\left(\mathbf{c}^{(i)}\right)$. Let us prove that there exists some constant $C$ that does not depend on $(t, \epsilon)$ and such that $\forall \theta \in[0,1]$ :

$$
\begin{equation*}
\left|\psi^{\prime}(\theta)\right| \leq \frac{C}{n_{0}} \tag{6.91}
\end{equation*}
$$

Then, the bounded difference property will follow from the mean value theorem,

$$
\sup _{\substack{\mathbf{c} \in\left[-\left\|\varphi_{1}\right\|_{\infty},\left\|\varphi_{1}\right\|_{\infty}\right]^{n_{1}} \\ c_{i}^{(i)} \in\left[-\left\|\varphi_{1}\right\| \infty,\left\|\varphi_{1}\right\|_{\infty}\right], \forall j \neq i: c_{j}^{(i)}=c_{j}}}\left|g(\mathbf{c})-g\left(\mathbf{c}^{(i)}\right)\right| \leq \frac{C}{n_{0}} .
$$

We denote by $\tilde{\mathbb{E}}[\cdot]$ the conditional expectation $\mathbb{E}\left[\cdot \mid \mathbf{X}^{(1)}=\theta \mathbf{c}+(1-\theta) \mathbf{c}^{(i)}, \mathbf{W}^{(1)}, \mathbf{X}\right]$. The derivative of $\psi$ reads

$$
\begin{align*}
\left|\psi^{\prime}(\theta)\right|= & \frac{\left|c_{i}-c_{i}^{(i)}\right|}{n_{0}} \tilde{\mathbb{E}}\left[\left\langle\frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial X_{i}^{(1)}}\right\rangle_{n, t, \epsilon}\right] \\
\leq & \frac{2\left\|\varphi_{1}\right\|_{\infty}}{n_{0} \Delta} \sum_{\mu=1}^{n_{2}}\left|\tilde{\mathbb{E}}\left[\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right) \frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial X_{i}^{(1)}}\right\rangle_{n, t, \epsilon}\right]\right| \\
& \quad+\frac{2\left\|\varphi_{1}\right\|_{\infty}}{n_{0}}\left|\tilde{\mathbb{E}}\left[\sqrt{R_{1}(t, \epsilon)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)+\widetilde{Z}_{i}\right\rangle_{n, t, \epsilon}\right]\right| \tag{6.92}
\end{align*}
$$

The last expectation on the r.h.s. of (6.92) satisfies

$$
\begin{aligned}
& \left|\tilde{\mathbb{E}}\left[\sqrt{R_{1}(t, \epsilon)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)+\widetilde{Z}_{i}\right\rangle_{n, t, \epsilon}\right]\right| \\
& \quad \leq \tilde{\mathbb{E}}\left[\sqrt{2 s_{n_{0}}+r_{\max }}\left(\sqrt{2 s_{n_{0}}+r_{\max }} 2\left\|\varphi_{1}\right\|_{\infty}+\left|\widetilde{Z}_{i}\right|\right)\right] \\
& \quad=2\left(1+r_{\max }\right)\left\|\varphi_{1}\right\|_{\infty}+\sqrt{\frac{2\left(1+r_{\max }\right)}{\pi}}
\end{aligned}
$$

Note that $\partial \Gamma_{\mu}^{(t, \epsilon)} / \partial X_{i}^{(1)}=\sqrt{{ }^{(1-t)} / n_{1}} W_{\mu i}^{(2)} \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)$ is independent of $Z_{\mu}$. Therefore, the expectations in the sum on the r.h.s. of (6.92) read $\forall \mu \in\left\{1, \ldots, n_{2}\right\}$ :

$$
\begin{aligned}
& \tilde{\mathbb{E}}\left[\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\right.\right.\right.\left.\left.\left.\sqrt{\Delta} Z_{\mu}\right) \frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial X_{i}^{(1)}}\right\rangle_{n, t, \epsilon}\right]=\sqrt{\frac{1-t}{n_{1}}} \tilde{\mathbb{E}}\left[\left\langle\Gamma_{\mu}^{(t, \epsilon)} W_{\mu i}^{(2)} \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)\right\rangle_{n, t, \epsilon}\right] \\
&=\frac{1-t}{n_{1}} \tilde{\mathbb{E}}\left\langle\Gamma_{\mu}^{(t, \epsilon)} X_{i}^{(1)} \varphi_{2}^{\prime \prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)\right\rangle_{\widehat{\mathcal{H}} t, \epsilon}+\frac{1}{\sqrt{n_{1}}} \tilde{\mathbb{E}}\left\langle\frac{\partial \Gamma_{t, \epsilon, \mu}}{\partial W_{\mu i}^{(2)}} \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)\right\rangle_{n, t, \epsilon} \\
&+\frac{1}{\sqrt{n_{1}}} \tilde{\mathbb{E}}\left[\left\langle\Gamma_{\mu}^{(t, \epsilon)} \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)\right\rangle_{n, t, \epsilon}\left\langle\frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial W_{\mu i}^{(2)}}\right\rangle_{n, t, \epsilon}\right]
\end{aligned}
$$

$$
\begin{align*}
&-\frac{1}{\sqrt{n_{1}}} \tilde{\mathbb{E}}\left\langle\Gamma_{\mu}^{(t, \epsilon)} \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right) \frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial W_{\mu i}^{(2)}}\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
&=\frac{1-t}{n_{1}} \tilde{\mathbb{E}}\left\langle\Gamma_{\mu}^{(t, \epsilon)} X_{i}^{(1)} \varphi_{2}^{\prime \prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)\right\rangle_{\mathcal{H}_{t, \epsilon}}+\frac{1}{\sqrt{n_{1}}} \tilde{\mathbb{E}}\left\langle\frac{\partial \Gamma_{t, \epsilon, \mu}}{\partial W_{\mu i}^{(2)}} \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
&+\frac{1}{\sqrt{n_{1}}} \tilde{\mathbb{E}}\left\langle\varphi_{2}^{\prime}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right) \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right) \frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial W_{\mu i}^{(2)}}\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
&-\frac{1}{\sqrt{n_{1}}} \tilde{\mathbb{E}}\left[\left\langle\varphi_{2}^{\prime}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right) \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon}\left\langle\frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial W_{\mu i}^{(2)}}\right\rangle_{n, t, \epsilon}\right], \tag{6.93}
\end{align*}
$$

where the second equality follows from an integration by parts w.r.t. $W_{\mu i}^{(2)}$, and the third from the definition of $\Gamma_{\mu}^{(t, \epsilon)}$. It is easily shown that there exists a constant $C^{\prime}$ that does not depend on $(t, \epsilon)$ and such that the four summands on the r.h.s. of (6.93) are bounded by $C^{\prime} / n_{1}$. All in all, we see that there exists $C$ such that (6.91) is satisfied uniformly in $(t, \epsilon)$. By McDiarmid's inequality, we thus have

$$
\mathbb{E}\left[\left.\left(\mathbb{E}\left[\left.\frac{\ln \hat{\mathcal{Z}}_{n, t, \epsilon}}{n_{0}} \right\rvert\, \mathbf{X}^{(1)}, \mathbf{W}^{(1)}, \mathbf{X}\right]-\mathbb{E}\left[\left.\frac{\ln \hat{\mathcal{Z}}_{n, t, \epsilon}}{n_{0}} \right\rvert\, \mathbf{W}^{(1)}, \mathbf{X}\right]\right)^{2} \right\rvert\, \mathbf{W}^{(1)}, \mathbf{X}\right] \leq \frac{n_{1} C^{2}}{4 n_{0}^{2}}
$$

almost surely. Taking the expectation on both sides of this inequality ends the proof.

Lemma 6.19. Under the assumption of Proposition 6.14, there exists a positive constant $C$, depending only on $\left(\varphi_{1}, \varphi_{2}, S, \Delta, \alpha_{1}, \alpha_{2}\right)$, such that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{W}^{(1)}, \mathbf{X}\right]-\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{X}\right]\right)^{2}\right] \leq \frac{C}{n_{0}} \tag{6.94}
\end{equation*}
$$

Proof. We see $g\left(\mathbf{W}^{(1)}\right)=\mathbb{E}\left[\ln \hat{\mathcal{Z}}_{n}(t, \epsilon) \mid \mathbf{W}^{(1)}, \mathbf{X}\right] / n_{0}$ as a function of $\mathbf{W}^{(1)}$ only and we work conditionally to $\mathbf{X}$. Denote $\tilde{\mathbb{E}}[\cdot]$ the conditional expectation $\mathbb{E}\left[\cdot \mid \mathbf{W}^{(1)}, \mathbf{X}\right]$. We introduce the notations

$$
\begin{array}{rlrl}
\mathbf{X}^{\left(1^{\prime}\right)} & =\varphi_{1}^{\prime}\left(\mathbf{W}^{(1)} \mathbf{x} / \sqrt{n_{0}}, \mathbf{A}^{(1)}\right) ; & \mathbf{x}^{\left(1^{\prime}\right)}=\varphi_{1}^{\prime}\left(\mathbf{W}^{(1)} \mathbf{x} / \sqrt{n_{0}}, \mathbf{a}\right) ; \\
\mathbf{X}^{\left(1^{\prime \prime}\right)}=\varphi_{1}^{\prime \prime}\left(\mathbf{W}^{(1)} \mathbf{X} / \sqrt{n_{0}}, \mathbf{A}^{(1)}\right) ; & \mathbf{x}^{\left(1^{\prime \prime}\right)}=\varphi_{1}^{\prime \prime}\left(\mathbf{W}^{(1)} \mathbf{x} / \sqrt{n_{0}}, \mathbf{a}\right) ; \\
\mathbf{X}^{\left(2^{\prime}\right)}=\varphi_{2}^{\prime}\left(\mathbf{S}^{(t, \epsilon)}, \mathbf{A}^{(2)}\right) \quad ; \quad \mathbf{x}^{\left(2^{\prime}\right)}=\varphi_{2}^{\prime}\left(\mathbf{s}^{(t, \epsilon)}, \mathbf{A}^{(2)}\right) ; \\
\mathbf{X}^{\left(2^{\prime \prime}\right)}=\varphi_{2}^{\prime \prime}\left(\mathbf{S}^{(t, \epsilon)}, \mathbf{A}^{(2)}\right) \quad ; \quad \mathbf{x}^{\left(2^{\prime \prime}\right)}=\varphi_{2}^{\prime \prime}\left(\mathbf{s}^{(t, \epsilon)}, \mathbf{A}^{(2)}\right) . \tag{6.98}
\end{array}
$$

Fix $(i, j) \in\left\{1, \ldots, n_{1}\right\} \times\left\{1, \ldots, n_{0}\right\}$. The partial derivative of $g$ w.r.t. $W_{i j}^{(1)}$ reads

$$
\begin{align*}
& \frac{\partial g}{\partial W_{i j}^{(1)}}=-\frac{1}{n_{0} \Delta} \sum_{\mu=1}^{n_{2}} \tilde{\mathbb{E}}\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right) \frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial W_{i j}^{(1)}}\right\rangle_{n, t, \epsilon} \\
& -\frac{\sqrt{R_{1}(t, \epsilon)}}{n_{0}^{3 / 2}} \mathbb{E}\left\langle\left(\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)+\widetilde{Z}_{i}\right)\left(X_{j} X_{i}^{\left(1^{\prime}\right)}-x_{j} x_{i}^{\left(1^{\prime}\right)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon} \tag{6.99}
\end{align*}
$$

The absolute value of the second term on the r.h.s. of (6.99) is easily upper bounded by

$$
\frac{\sqrt{1+r_{\max }}}{n_{0}^{3 / 2}}\left(2 \sqrt{1+r_{\max }}\left\|\varphi_{1}\right\|_{\infty}+\sqrt{\frac{2}{\pi}}\right) 2 S\left\|\varphi_{1}^{\prime}\right\|_{\infty}
$$

We turn to the terms in the sum on the r.h.s. of (6.99). First, note that

$$
\frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial W_{i j}^{(1)}}=\sqrt{\frac{1-t}{n_{0} n_{1}}} W_{\mu i}^{(2)}\left(X_{j} X_{i}^{\left(1^{\prime}\right)} \varphi_{2}^{\prime}\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}^{(2)}\right)-x_{j} x_{i}^{\left(1^{\prime}\right)} \varphi_{2}^{\prime}\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}^{(2)}\right)\right)
$$

Plugging the latter in $\tilde{\mathbb{E}}\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right)^{\partial \Gamma_{\mu}^{(t, \epsilon)} / \partial W_{i j}^{(1)}}\right\rangle_{\boldsymbol{n}, t, \epsilon}$, and integrating by parts w.r.t. $W_{\mu i}^{(2)}$, yields

$$
\begin{align*}
& \tilde{\mathbb{E}}\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right) \frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial W_{i j}^{(1)}}\right\rangle_{n, t, \epsilon} \\
& =\frac{1-t}{n_{1} \sqrt{n_{0}} \Delta} \tilde{\mathbb{E}}\left[\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right)\left(X_{j} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}-x_{j} x_{i}^{\left(1^{\prime}\right)} x_{\mu}^{\left(2^{\prime}\right)}\right)\right\rangle_{n, t, \epsilon}\right. \\
& \left.\cdot\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right)\left(X_{i}^{1} X_{\mu}^{\left(2^{\prime}\right)}-x_{i}^{1} x_{\mu}^{\left(2^{\prime}\right)}\right)\right\rangle_{n, t, \epsilon}\right] \\
& \quad-\frac{1-t}{n_{1} \sqrt{n_{0}} \Delta} \tilde{\mathbb{E}}\left[\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right)^{2}\left(X_{j} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}-x_{j} x_{i}^{\left(1^{\prime}\right)} x_{\mu}^{\left(2^{\prime}\right)}\right)\right.\right. \\
& \left.\left.\quad \cdot\left(X_{i}^{1} X_{\mu}^{\left(2^{\prime}\right)}-x_{i}^{1} x_{\mu}^{\left(2^{\prime}\right)}\right)\right\rangle_{n, t, \epsilon}\right] \\
& \quad+\frac{1-t}{n_{1} \sqrt{n_{0}}} \tilde{\mathbb{E}}\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right)\left(X_{j} X_{i}^{\left(1^{\prime}\right)} X_{i}^{(1)} X_{\mu}^{\left(2^{\prime \prime}\right)}-x_{j} x_{i}^{\left(1^{\prime}\right)} x_{i}^{(1)} x_{\mu}^{\left(2^{\prime \prime \prime}\right)}\right)\right\rangle_{n, t, \epsilon} \\
& \quad+\frac{1-t}{n_{1} \sqrt{n_{0}}} \tilde{\mathbb{E}}\left\langle\left(X_{i}^{1} X_{\mu}^{\left(2^{\prime}\right)}-x_{i}^{1} x_{\mu}^{\left(2^{\prime}\right)}\right)\left(X_{j} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}-x_{j} x_{i}^{\left(1^{\prime}\right)} x_{\mu}^{\left(2^{\prime}\right)}\right)\right\rangle_{n, t, \epsilon} . \tag{6.100}
\end{align*}
$$

The absolute value of each of the four conditional expectations on the r.h.s. of 6.100) is easily upper bounded by a constant that does not depend on $(t, \epsilon)$. Putting everything together, there exists a positive constant $C$ such that almost surely $\forall(i, j) \in\left\{1, \ldots, n_{1}\right\} \times\left\{1, \ldots, n_{0}\right\}$ :

$$
\left|\frac{\partial g}{\partial W_{i j}^{(1)}}\right| \leq \frac{C}{n_{0}^{3 / 2}}
$$

Hence,

$$
\|\nabla g\|^{2}=\sum_{i=1}^{n_{1}} \sum_{j=1}^{n_{0}}\left|\frac{\partial g}{\partial W_{i j}^{(1)}}\right|^{2} \leq \frac{1}{n_{0}} \frac{n_{1}}{n_{0}} C^{2}
$$

almost surely. We end the proof with an application of the Gaussian-Poincaré inequality (see Proposition 2.7).
Lemma 6.20. Under the assumption of Proposition 6.14, there exists a positive constant $C$, depending only on $\left(\varphi_{1}, \varphi_{2}, S, \Delta, \alpha_{1}, \alpha_{2}\right)$, such that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\begin{equation*}
\mathbb{E}\left[\left(\mathbb{E}\left[\left.\frac{\ln \widehat{\mathcal{Z}}_{n}(t, \epsilon)}{n_{0}} \right\rvert\, \mathbf{X}\right]-\mathbb{E}\left[\frac{\ln \widehat{\mathcal{Z}}_{\boldsymbol{n}}(t, \epsilon)}{n_{0}}\right]\right)^{2}\right] \leq \frac{C}{n_{0}} \tag{6.101}
\end{equation*}
$$

Proof. We see $g(\mathbf{X})=\mathbb{E}\left[\ln \hat{\mathcal{E}}_{n}(t, \epsilon) \mid \mathbf{X}\right] / n_{0}$ as a function of $\mathbf{X}$. We denote by $\tilde{\mathbb{E}}$ the conditional expectation $\mathbb{E}[\cdot \mid \mathbf{X}]$. To lighten the equations, we use the notations (6.95) to (6.98) introduced in the proof of Lemma 6.19.

We show that the partial derivatives of $g$ are almost surely bounded by $C / n_{0}$, where $C$ is a positive constant that does not depend on $(t, \epsilon)$. Exactly like in the proof of Lemma 6.18, it implies that $g$ satisfies a bounded difference property and we end the proof with an application of McDiarmid's inequality (see Proposition (2.6). For all $j \in\left\{1, \ldots, n_{0}\right\}$ :

$$
\begin{align*}
\frac{\partial g}{\partial X_{j}}= & -\frac{\sqrt{1-t}}{n_{0}^{3 / 2} \sqrt{n_{1}} \Delta} \sum_{\mu=1}^{n_{2}} \sum_{i=1}^{n_{1}} \tilde{\mathbb{E}}\left[W_{i j}^{(1)} W_{\mu i}^{(2)} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}\left\langle\Gamma_{\mu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\mu}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
& -\frac{\sqrt{R_{1}(t, \epsilon)}}{n_{0}^{3 / 2}} \sum_{i=1}^{n_{1}} \tilde{\mathbb{E}}\left[W_{i j}^{(1)} X_{i}^{\left(1^{\prime}\right)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)+\widehat{Z}_{i}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
= & -\frac{\sqrt{1-t}}{n_{0}^{3 / 2} \sqrt{n_{1}} \Delta} \sum_{\mu=1}^{n_{2}} \sum_{i=1}^{n_{1}} \tilde{\mathbb{E}}\left[W_{i j}^{(1)} W_{\mu i}^{(2)} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
& -\frac{\sqrt{R_{1}(t, \epsilon)}}{n_{0}^{3 / 2}} \sum_{i=1}^{n_{1}} \tilde{\mathbb{E}}\left[W_{i j}^{(1)} X_{i}^{\left(1^{\prime}\right)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \tag{6.102}
\end{align*}
$$

We first look at the summands in the second sum on the r.h.s. of (6.102). A Gaussian integration by parts w.r.t. $W_{i j}^{(1)}$ yields

$$
\begin{align*}
& \tilde{\mathbb{E}}\left[W_{i j}^{(1)} X_{i}^{\left(1^{\prime}\right)}\langle \right.\left.\left.\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
&=\frac{1}{\sqrt{n_{0}}} \tilde{\mathbb{E}}\left[X_{j} X_{i}^{\left(1^{\prime \prime}\right)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
&+\frac{1}{\sqrt{n_{0}}} \tilde{\mathbb{E}}\left[X_{i}^{\left(1^{\prime}\right)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{j} X_{i}^{\left(1^{\prime}\right)}-x_{j} x_{i}^{\left(1^{\prime}\right)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
& \quad \tilde{\mathbb{E}}\left[X_{i}^{\left(1^{\prime}\right)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right) \frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial W_{i j}^{(1)}}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
&+\tilde{\mathbb{E}}\left[X_{i}^{\left(1^{\prime}\right)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon}\left\langle\frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial W_{i j}^{(1)}}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] . \tag{6.103}
\end{align*}
$$

The absolute values of the first two conditional expectations on the r.h.s. of 6.103) are $O\left(1 / \sqrt{n_{0}}\right)$ uniformly in $(t, \epsilon)$. This is also the case of the last two conditional expectations; to show this we can proceed as in the proof of Lemma 6.19 where $\partial \widehat{\mathcal{H}}_{t, \epsilon} / \partial W_{i j}^{(1)}$ was already computed. All in all, there exists a positive constant $C$ that does not depend on $(t, \epsilon)$ and such that almost surely $\forall i \in\left\{1, \ldots, n_{1}\right\}$ :

$$
\begin{equation*}
\left|\tilde{\mathbb{E}}\left[W_{i j}^{(1)} X_{i}^{\left(1^{\prime}\right)}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)\right\rangle_{\boldsymbol{n}, t, \epsilon}\right]\right| \leq \frac{C}{\sqrt{n_{0}}} \tag{6.104}
\end{equation*}
$$

We now look at the summands in the first sum on the r.h.s. of 6.102). For every
pair $(\mu, i) \in\left\{1, \ldots, n_{2}\right\} \times\left\{1, \ldots, n_{1}\right\}$ :

$$
\begin{align*}
& \tilde{\mathbb{E}}[ {\left[W_{i j}^{(1)}\right.} \\
&\left.=W_{\mu i}^{(2)} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
&=\tilde{\mathbb{E}} {\left[W_{\mu i}^{(2)} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\left\langle\frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial W_{i j}^{(1)}}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] } \\
&-\tilde{\mathbb{E}}\left[W_{\mu i}^{(2)} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}\left\langle\Gamma_{\mu}^{(t, \epsilon)} \frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial W_{i j}^{(1)}}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
&+\frac{1}{\sqrt{n_{0}}} \tilde{\mathbb{E}}\left[W_{\mu i}^{(2)} X_{j} X_{i}^{\left(1^{\prime \prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] \\
&+\sqrt{\frac{1-t}{n_{1} n_{0}}} \tilde{\mathbb{E}}\left[\left(W_{\mu i}^{(2)}\right)^{2} X_{j}\left(X_{i}^{\left(1^{\prime}\right)}\right)^{2} X_{\mu}^{\left(2^{\prime \prime}\right)}\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right]  \tag{6.105}\\
&+\sqrt{\frac{1-t}{n_{1} n_{0}}} \tilde{\mathbb{E}}\left[\left(W_{\mu i}^{(2)}\right)^{2} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}\left\langle X_{j} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}-x_{j} x_{i}^{\left(1^{\prime}\right)} x_{\mu}^{\left(2^{\prime}\right)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right] .
\end{align*}
$$

We can show that the absolute value of the third conditional expectation on the r.h.s. of 6.105) is a $O\left(n_{0}^{-1}\right)$ after integrating by parts w.r.t. $W_{\mu i}^{(2)}$. The fourth and fifth expectations have their absolute values upper bounded by $\frac{2 S}{\sqrt{n_{0} n_{1}}}\left\|\varphi_{1}^{\prime}\right\|_{\infty}^{2}\left\|\varphi_{2}\right\|_{\infty}\left\|\varphi_{2}^{\prime \prime}\right\|_{\infty}$ and $\frac{2 S}{\sqrt{n_{0} n_{1}}}\left\|\varphi_{1}^{\prime}\right\|_{\infty}^{2}\left\|\varphi_{2}^{\prime}\right\|_{\infty}^{2}$, respectively (remember that $\left.\mathbb{E}\left(W_{\mu i}^{(2)}\right)^{2}=1\right)$. Regarding the first two conditional expectations on the r.h.s. of 6.105), first note that

$$
\begin{align*}
\frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial W_{i j}^{(1)}} & =\frac{1}{\Delta} \sqrt{\frac{1-t}{n_{0} n_{1}}} \sum_{\nu=1}^{n_{2}} W_{\nu i}^{(2)}\left(\Gamma_{\nu}^{(t, \epsilon)}+\sqrt{\Delta} Z_{\nu}\right)\left(X_{j} X_{i}^{\left(1^{\prime}\right)} X_{\nu}^{\left(2^{\prime}\right)}-x_{j} x_{i}^{\left(1^{\prime}\right)} x_{\nu}^{\left(2^{\prime}\right)}\right) \\
& +\sqrt{\frac{R_{1}(t, \epsilon)}{n_{0}}}\left(\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{(1)}-x_{i}^{(1)}\right)+\widetilde{Z}_{i}\right)\left(X_{j} X_{i}^{\left(1^{\prime}\right)}-x_{j} x_{i}^{\left(1^{\prime}\right)}\right) \tag{6.106}
\end{align*}
$$

We plug (6.106) in the expressions of the first two conditional expectations. The latter can then be split into $n_{2}+1$ terms (one term for each summand on the r.h.s. of (6.106) such that:

- the term due to $\nu=\mu$ is directly upperbounded by $C \mathbb{E}\left(W_{\mu i}^{(2)}\right)^{2} / \sqrt{n_{1} n_{0}}$ for some constant $C$;
- each term due to a $\nu \neq \mu$ is upperbounded by $C^{\prime} / n_{0}^{2}$ after integrating by parts w.r.t. $W_{\mu i}^{(2)}$ and $W_{\nu i}^{(2)}$;
- the term due to the last summand on the r.h.s. of (6.106) is upper bounded by $C^{\prime \prime} / n_{0}$ after integrating by parts w.r.t. $W_{\mu i}^{(2)}$.

All in all, there exists a positive constant $C$ that does not depend on $(t, \epsilon)$ and such that almost surely $\forall(\mu, i) \in\left\{1, \ldots, n_{2}\right\} \times\left\{1, \ldots, n_{1}\right\}$ :

$$
\begin{equation*}
\left|\tilde{\mathbb{E}}\left[W_{i j}^{(1)} W_{\mu i}^{(2)} X_{i}^{\left(1^{\prime}\right)} X_{\mu}^{\left(2^{\prime}\right)}\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right]\right| \leq \frac{C}{n_{0}} \tag{6.107}
\end{equation*}
$$

Combining (6.102), (6.104) and (6.107) gives the existence of a positive constant $C$ that does not depend on $(t, \epsilon)$ and such that almost surely $\forall j \in\left\{1, \ldots, n_{0}\right\}$ : $\left|\partial g / \partial X_{j}\right| \leq C / n_{0}$. We can thus end the proof by McDimarmid's inequality.

## 6.E Concentration of the overlap

In this appendix we prove Proposition 6.5. The outline of the proof is similar to the one provided for the one-layer GLM in [29]. For a fixed $t \in[0,1]$, we treat the average free entropy as a function of $\left(R_{1}, R_{2}\right) \mapsto f_{n}(t, \epsilon)$ of $R_{1}=R_{1}(t, \epsilon)$ and $R_{2}=R_{2}(t, \epsilon)$. Note that this is possible because we work under the assumption that the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}},\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ are regular, hence $\epsilon \mapsto R(t, \epsilon)$ is a $C^{1}$-diffeomorphism from $\mathcal{B}_{n_{0}}$ onto its image. For a fixed $t \in[0,1]$, we also treat the free entropy

$$
\begin{equation*}
F_{\boldsymbol{n}}(t, \epsilon):=\frac{1}{n_{0}} \ln \mathcal{Z}_{\boldsymbol{n}, t, \epsilon}\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right) \tag{6.108}
\end{equation*}
$$

whose expectation is equal to $f_{n}(t, \epsilon)$, as a function of $\left(R_{1}, R_{2}\right)$.
Remember that $\mathbf{X}^{(1)}=\varphi_{1}\left(\mathbf{W}^{(1)} \mathbf{x} / \sqrt{n_{0}}, \mathbf{a}\right)$ where the triplet $(\mathbf{x}, \mathbf{u}, \mathbf{a})$ is sampled from the joint posterior distribution (6.20). Hence, $\mathbf{x}^{(1)}$ is nothing but a sample obtained from the conditionnal probability distribution of $\mathbf{X}^{(1)}$ given $\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{V}\right)$. Define

$$
\begin{equation*}
\mathcal{L}:=\frac{1}{n_{1}} \sum_{i=1}^{n_{1}} \frac{\left(x_{i}^{(1)}\right)^{2}}{2}-x_{i}^{(1)} X_{i}^{(1)}-\frac{x_{i}^{(1)} \widetilde{Z}_{i}}{2 \sqrt{R_{1}}} \tag{6.109}
\end{equation*}
$$

This quantity is closely linked to the overlap Q . We first prove an important identity.
Lemma 6.21 (Formula for $\left.\mathbb{E}\langle\mathcal{L}\rangle_{n, t, \epsilon}\right)$. Assume that (H1), (H2) and (H3) hold. Then, $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n_{0}}$ :

$$
\begin{equation*}
\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}=-\frac{1}{2} \mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon} \tag{6.110}
\end{equation*}
$$

Proof. Note that $\forall i \in\left\{1, \ldots, n_{1}\right\}$ :

$$
\begin{gathered}
\mathbb{E}\left[X_{i}^{(1)}\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right]=\mathbb{E}\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}^{2}, \\
\mathbb{E}\left[\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon} \widetilde{Z}_{i}\right]=\mathbb{E}\left[\frac{\partial\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}}{\partial \widehat{Z}_{i}}\right]=\mathbb{E}\left[\sqrt{R_{1}(t, \epsilon)}\left(\left\langle\left(x_{i}^{(1)}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}-\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}^{2}\right)\right],
\end{gathered}
$$

where the first identity follows from the Nishimory identity, and the second one from a Gaussian integration by parts w.r.t. $\widetilde{Z}_{i}$. Making use of these two identities, we directly obtain

$$
\begin{aligned}
& \mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}=\frac{1}{n_{1}} \sum_{i=1}^{n_{1}} \frac{1}{2} \mathbb{E}\left\langle\left(x_{i}^{(1)}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}-\mathbb{E}\left[X_{i}^{(1)}\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}\right]-\frac{\mathbb{E}\left[\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon} \widetilde{Z}_{i}\right]}{2 \sqrt{R_{1}(t, \epsilon)}} \\
& =-\frac{1}{2 n_{1}} \sum_{i=1}^{n_{1}} \mathbb{E}\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}^{2}=-\frac{1}{2 n_{1}} \sum_{i=1}^{n_{1}} \mathbb{E}\left[\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon} X_{i}^{(1)}\right]=-\frac{1}{2} \mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon} .
\end{aligned}
$$

The fluctuations of the overlap $Q$ are related to the fluctuations of $\mathcal{L}$ through the identity

$$
\begin{align*}
\mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}=\frac{1}{4} \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{n, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} & +\frac{1}{2} \mathbb{E}\left[\left\langle Q^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}-\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}^{2}\right] \\
& +\frac{\rho_{1}\left(n_{0}\right)}{4 n_{1} R_{1}} \tag{6.111}
\end{align*}
$$

The full derivation of the identity (6.111) is found in [87, Section 6]. It involves lengthy algebra using the Nishimori identity and integration by parts w.r.t. the Gaussian random variables $\widetilde{Z}_{i}$. The derivation in 87, Section 6] applies directly to our problem by doing the identifications $X_{i}^{(1)} \leftrightarrow S_{i}, x_{i}^{(1)} \leftrightarrow X_{i}, n_{1} \leftrightarrow n, R_{1} \leftrightarrow \widetilde{\epsilon}$. The identity 6.111) yields the important inequality

$$
\begin{equation*}
\mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \geq \frac{1}{4} \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{n, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} . \tag{6.112}
\end{equation*}
$$

Thanks to 6.112), Proposition 6.5 directly follows from the following result on the concentration of $\mathcal{L}$.

Proposition 6.22 (Concentration of $\mathcal{L}$ around its expectation). Assume that (H1), (H2), (H3) hold and $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}},\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ are regular. Then, there exists a constant $C$ such that $\forall t \in[0,1]$ :

$$
\begin{equation*}
\int d t \int_{\mathcal{B}_{n_{0}}} \frac{d \epsilon}{s_{n_{0}}^{2}} \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \leq \frac{C\left(\varphi_{1}, \varphi_{2}, \alpha_{1}, \alpha_{2}, S\right)}{s_{n_{0}}^{2} n_{0}^{1 / 4}} \tag{6.113}
\end{equation*}
$$

Proof. Proposition 6.22 follows from the identity

$$
\mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}=\mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}+\mathbb{E}\left(\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}-\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}
$$

and Lemmas 6.23 and 6.24 stated below.
Lemma 6.23 (Concentration of $\mathcal{L}$ on $\left.\langle\mathcal{L}\rangle_{n, t, \epsilon}\right)$. Assume that (H1), (H2), (H3) hold and $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}},\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ are regular. Then, for $n_{0}$ large enough, $\forall t \in[0,1]$ :

$$
\int_{\mathcal{B}_{n_{0}}} d \epsilon \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \leq \frac{\rho_{1}\left(1+\rho_{1}\right)}{\alpha_{1} n_{0}}
$$

Proof. The first and second derivatives of the free entropy 6.108 with respect to $R_{1}$ read

$$
\begin{align*}
\frac{\partial F_{\boldsymbol{n}}}{\partial R_{1}} & =-\frac{n_{1}}{n_{0}}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}-\frac{n_{1}}{2 n_{0}} \frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\frac{1}{2 n_{0} \sqrt{R_{1}}} \sum_{i=1}^{n_{1}} \widetilde{Z}_{i} X_{i}^{(1)} ;  \tag{6.114}\\
\frac{1}{n_{0}} \frac{\partial^{2} F_{\boldsymbol{n}}}{\partial R_{1}^{2}} & =\left(\frac{n_{1}}{n_{0}}\right)^{2}\left(\left\langle\mathcal{L}^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}-\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}^{2}\right)-\frac{1}{4 n_{0}^{2} R_{1}^{3 / 2}} \sum_{i=1}^{n_{1}}\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon} \widetilde{Z}_{i} . \tag{6.115}
\end{align*}
$$

Taking an expectation on both sides of these two identities gives

$$
\begin{align*}
& \frac{\partial f_{n}}{\partial R_{1}}=-\frac{n_{1}}{n_{0}} \mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}-\frac{n_{1}}{2 n_{0}} \rho_{1}\left(n_{0}\right)=-\frac{n_{1}}{2 n_{0}}\left(\rho_{1}\left(n_{0}\right)-\mathbb{E}\langle Q\rangle_{\boldsymbol{n}, t, \epsilon}\right) ;  \tag{6.116}\\
& \frac{1}{n_{0}} \frac{\partial^{2} f_{n}}{\partial R_{1}^{2}}=\left(\frac{n_{1}}{n_{0}}\right)^{2} \mathbb{E}\left[\left(\left\langle\mathcal{L}^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}-\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}^{2}\right)\right] \\
&-\frac{1}{4 n_{0}^{2} R_{1}} \sum_{i=1}^{n_{1}} \mathbb{E}\left[\left\langle\left(x_{i}^{(1)}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}-\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}^{2}\right] \tag{6.117}
\end{align*}
$$

Rearranging the terms in 6.117, we obtain

$$
\begin{align*}
\mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} & =\frac{n_{0}}{n_{1}^{2}} \frac{\partial^{2} f_{\boldsymbol{n}}}{\partial R_{1}^{2}}+\frac{1}{4 n_{1}^{2} R_{1}} \sum_{i=1}^{n_{1}} \mathbb{E}\left[\left\langle\left(x_{i}^{(1)}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}-\left\langle x_{i}^{(1)}\right\rangle_{\boldsymbol{n}, t, \epsilon}^{2}\right] \\
& \leq \frac{n_{0}}{n_{1}^{2}} \frac{\partial^{2} f_{\boldsymbol{n}}}{\partial R_{1}^{2}}+\frac{\rho_{1}\left(n_{0}\right)}{4 n_{1} \epsilon_{1}} \tag{6.118}
\end{align*}
$$

where the inequality is because $\sum_{i=1}^{n_{1}} \mathbb{E}\left\langle\left(x_{i}^{(1)}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon}=\mathbb{E}\left\|\mathbf{X}^{(1)}\right\|^{2}=n_{1} \rho_{1}\left(n_{0}\right)$ and $R_{1} \geq \epsilon_{1}$. By assumption, $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}},\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ are regular. Thus, for a fixed $t \in[0,1], R:\left(\epsilon_{1}, \epsilon_{2}\right) \mapsto\left(R_{1}(t, \epsilon), R_{2}(t, \epsilon)\right)$ is a $C^{1}$-diffeomorphism whose Jacobian determinant satisfies $\forall \epsilon \in \mathcal{B}_{n_{0}}:\left|J_{R}(\epsilon)\right| \geq 1$. Integrating both sides of the last inequality over $\epsilon \in \mathcal{B}_{n_{0}}$, we find that

$$
\begin{aligned}
\int_{\mathcal{B}_{n_{0}}} d \epsilon \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \leq & \frac{n_{0}}{n_{1}^{2}} \int_{R\left(\mathcal{B}_{n_{0}}\right)} \frac{d R_{1} d R_{2}}{\left|J_{R}\left(R^{-1}\left(R_{1}, R_{2}\right)\right)\right|} \frac{\partial^{2} f_{n}}{\partial R_{1}^{2}} \\
& +\frac{\rho_{1}\left(n_{0}\right) s_{n_{0}}}{4 n_{1}} \int_{s_{n_{0}}}^{2 s_{n_{0}}} \frac{d \epsilon_{1}}{\epsilon_{1}} \\
\leq & \frac{n_{0}}{n_{1}^{2}} \int_{R\left(\mathcal{B}_{n_{0}}\right)} d R_{1} d R_{2} \frac{\partial^{2} f_{n}}{\partial R_{1}^{2}} \\
& +\frac{\rho_{1}\left(n_{0}\right) s_{n_{0}}}{4 n_{1}} \int_{s_{n_{0}}}^{2 s_{s_{0}}} \frac{d \epsilon_{1}}{\epsilon_{1}}
\end{aligned}
$$

Clearly, $R\left(\mathcal{B}_{n_{0}}\right) \subseteq\left[s_{n_{0}}, 2 s_{n_{0}}+r_{\max }\right] \times\left[s_{n_{0}}, 2 s_{n_{0}}+\rho_{1}\left(n_{0}\right)\right]$. Hence,

$$
\begin{aligned}
& \int_{\mathcal{B}_{n_{0}}} d \epsilon \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \\
& \quad \leq \frac{n_{0}}{n_{1}^{2}} \int_{s_{n_{0}}}^{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right)} d R_{2}\left(\left.\frac{\partial f_{n}}{\partial R_{1}}\right|_{R_{1}=2 s_{n_{0}}+r_{\max }, R_{2}}-\left.\frac{\partial f_{n}}{\partial R_{1}}\right|_{R_{1}=s_{n_{0}}, R_{2}}\right)+\frac{\rho_{1}\left(n_{0}\right) s_{n_{0}}}{4 n_{1}} \ln 2 \\
& \quad \leq-\left.\frac{n_{0}}{n_{1}^{2}} \int_{s_{n_{0}}}^{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right)} \quad d R_{2} \frac{\partial f_{n}}{\partial R_{1}}\right|_{R_{1}=s_{n_{0}}, R_{2}}+\frac{\rho_{1}\left(n_{0}\right) s_{n_{0}}}{4 n_{1}} \ln 2 \\
& \quad \leq \frac{n_{0}}{n_{1}^{2}}\left(s_{n_{0}}+\rho_{1}\left(n_{0}\right)\right) \frac{n_{1}}{2 n_{0}} \rho_{1}\left(n_{0}\right)+\frac{\rho_{1}\left(n_{0}\right) s_{n_{0}}}{4 n_{1}} \ln 2 \\
& \quad=\frac{\rho_{1}\left(n_{0}\right)}{2 n_{1}}\left(s_{n_{0}}+s_{n_{0}} \frac{\ln 2}{2}+\rho_{1}\left(n_{0}\right)\right)
\end{aligned}
$$

For the second inequality, we use that $\partial f_{n} / \partial R_{1}$ is nonpositive, which is clear from 6.116). For the third inequality, we use that $0 \leq-\partial f_{n} / \partial R_{1} \leq \frac{n_{1}}{2 n_{0}} \rho_{1}\left(n_{0}\right)$ (see (6.116). Finally, $s_{n_{0}}+s_{n_{0}}^{\ln 2 / 2} \leq 1$ and $\rho_{1}\left(n_{0}\right) \rightarrow \rho_{1}$, so for $n_{0}$ large enough

$$
\int_{\mathcal{B}_{n_{0}}} d \epsilon \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}\right\rangle_{\boldsymbol{n}, t, \epsilon} \leq \frac{\rho_{1}\left(1+\rho_{1}\right)}{\alpha_{1} n_{0}}
$$

Lemma 6.24 (Concentration of $\langle\mathcal{L}\rangle_{n, t, \epsilon}$ on $\mathbb{E}\langle\mathcal{L}\rangle_{n, t, \epsilon}$ ). Assume that (H1), (H2), (H3) hold and $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}},\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n_{0}}}$ are regular. Then, there exists a constant $C$ such that $\forall t \in[0,1]$ :

$$
\begin{equation*}
\int_{\mathcal{B}_{n_{0}}} d \epsilon \mathbb{E}\left(\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}-\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2} \leq \frac{C}{n_{0}^{1 / 4}} \tag{6.119}
\end{equation*}
$$

Proof. We define the two following functions:

$$
\begin{aligned}
\widetilde{F}\left(R_{1}\right) & :=F_{\boldsymbol{n}}(t, \epsilon)-\frac{\sqrt{R_{1}}}{n_{0}}\left\|\varphi_{1}\right\|_{\infty} \sum_{i=1}^{n_{1}}\left|\widetilde{Z}_{i}\right| \\
\widetilde{f}\left(R_{1}\right) & :=f_{\boldsymbol{n}}(t, \epsilon)-\frac{\sqrt{R_{1}}}{n_{0}}\left\|\varphi_{1}\right\|_{\infty} \sum_{i=1}^{n_{1}} \mathbb{E}\left|\widetilde{Z}_{i}\right| .
\end{aligned}
$$

The addition of the second term makes $\widetilde{F}$ convex as its second derivative is positive (remember that the second derivative $\partial^{2} F_{n} / \partial R_{1}^{2}$ is given by 6.115). Therefore, $\tilde{f}$ is convex as well. Define $A:=\frac{1}{n_{1}} \sum_{i=1}^{n_{1}}\left(\left|\widetilde{Z}_{i}\right|-\mathbb{E}\left|\widetilde{Z}_{i}\right|\right)$. Note that

$$
\begin{equation*}
\widetilde{F}\left(R_{1}\right)-\widetilde{f}\left(R_{1}\right)=F_{\boldsymbol{n}}(t, \epsilon)-f_{\boldsymbol{n}}(t, \epsilon)-\sqrt{R_{1}}\left\|\varphi_{1}\right\|_{\infty} \frac{n_{1}}{n_{0}} A \tag{6.120}
\end{equation*}
$$

and the derivative of this difference (we use (6.114) and (6.116) reads

$$
\begin{align*}
\widetilde{F}^{\prime}\left(R_{1}\right)-\widetilde{f}^{\prime}\left(R_{1}\right)=-\frac{n_{1}}{n_{0}}\left(\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right. & \left.-\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)-\frac{n_{1}}{2 n_{0}}\left(\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right) \\
& -\frac{1}{2 n_{0} \sqrt{R_{1}}} \sum_{i=1}^{n_{1}} \widetilde{Z}_{i} X_{i}^{(1)}-\frac{\left\|\varphi_{1}\right\|_{\infty}}{2 \sqrt{R_{1}}} \frac{n_{1}}{n_{0}} A . \tag{6.121}
\end{align*}
$$

For $\delta \in\left(0, R_{1}\right)$, define $C_{\delta}\left(R_{1}\right):=\widetilde{f^{\prime}}\left(R_{1}+\delta\right)-\widetilde{f^{\prime}}\left(R_{1}-\delta\right) \geq 0$. By Lemma 2.8
(applied to $G=\widetilde{F}, g=\widetilde{f})$, and making use of (6.120), (6.121), $\forall \delta \in\left(0, R_{1}\right)$ :

$$
\begin{align*}
\mid\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}- & \mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon} \mid \\
\leq & C_{\delta}\left(R_{1}\right)+\delta^{-1} \sum_{u \in\{-\delta, 0, \delta\}} \frac{n_{0}}{n_{1}}\left|F_{\boldsymbol{n}}\left(R_{1}+u\right)-f_{\boldsymbol{n}}\left(R_{1}+u\right)\right|+\left\|\varphi_{1}\right\|_{\infty}|A| \sqrt{R_{1}+u} \\
& +\left|\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right|+\left|\frac{1}{2 n_{1} \sqrt{R_{1}}} \sum_{i=1}^{n_{1}} \widetilde{Z}_{i} X_{i}^{(1)}\right|+\frac{\left\|\varphi_{1}\right\|_{\infty}}{2 \sqrt{R_{1}}}|A| \\
\leq & C_{\delta}\left(R_{1}\right)+\delta^{-1} \sum_{u \in\{-\delta, 0, \delta\}} \frac{n_{0}}{n_{1}}\left|F_{\boldsymbol{n}}\left(R_{1}+u\right)-f_{\boldsymbol{n}}\left(R_{1}+u\right)\right|+\left|\frac{\left\|\mathbf{X}^{(1)}\right\|^{2}}{n_{1}}-\rho_{1}\left(n_{0}\right)\right| \\
& +\left|\frac{1}{2 n_{1} \sqrt{R_{1}}} \sum_{i=1}^{n_{1}} \widetilde{Z}_{i} X_{i}^{(1)}\right|+\left\|\varphi_{1}\right\|_{\infty}\left(\frac{1}{2 \sqrt{R_{1}}}+\frac{3 \sqrt{R_{1}}}{\delta}\right)|A| . \tag{6.122}
\end{align*}
$$

By Proposition 6.14 in Appendix 6.D there exists a constant $C$ such that $\forall(t, \epsilon) \in$ $[0,1] \times \mathcal{B}_{n_{0}}: \mathbb{E}\left[\left(F_{\boldsymbol{n}}(t, \epsilon)-f_{\boldsymbol{n}}(t, \epsilon)\right)^{2}\right] \leq C / n_{0}$. In the proof of Appendix 6.C we also prove the existence of a constant $C^{\prime}$ such that $\mathbb{E}\left[\left(\left\|\mathbf{X}^{(1)}\right\|^{2} / n_{1}-\rho_{1}\left(n_{0}\right)\right)^{2}\right] \leq C^{\prime} / n_{0}$. Besides,

$$
\begin{aligned}
\mathbb{E}\left[\left(\frac{1}{n_{1}} \sum_{i=1}^{n_{1}} \widetilde{Z}_{i} X_{i}^{(1)}\right)^{2}\right] & =\frac{1}{n_{1}^{2}} \sum_{i, j=1}^{n_{1}} \mathbb{E}\left[\widetilde{Z}_{i} \widetilde{Z}_{j} X_{i}^{(1)} X_{j}^{1}\right] \\
& =\frac{1}{n_{1}^{2}} \sum_{i=1}^{n_{1}} \mathbb{E}\left[\widetilde{Z}_{i}^{2}\left(X_{i}^{(1)}\right)^{2}\right]=\frac{\mathbb{E}\left[\left(X_{1}^{1}\right)^{2}\right]}{n_{1}}=\frac{\rho_{1}\left(n_{0}\right)}{n_{1}},
\end{aligned}
$$

and $\mathbb{E} A^{2}=\operatorname{Var}\left|\widetilde{Z}_{1}\right| n_{1} \leq 1 / n_{1}$. Taking the square and the expectation on both sides of (6.122), and then making use of the inequality $\left(\sum_{i=1}^{p} v_{i}\right)^{2} \leq p \sum_{i=1}^{p} v_{i}^{2}$ as well as the different upper bounds that we have just mentioned, yields $\forall \delta \in\left(0, R_{1}\right)$ :

$$
\begin{align*}
& \frac{\mathbb{E}\left(\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}-\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2}}{7} \\
& \quad \leq C_{\delta}\left(R_{1}\right)^{2}+\frac{3 C n_{0}}{\delta^{2} n_{1}^{2}}+\frac{C^{\prime}}{n_{0}}+\frac{\rho_{1}\left(n_{0}\right)}{4 R_{1} n_{1}}+\frac{\left\|\varphi_{1}\right\|_{\infty}^{2}}{n_{1}}\left(\frac{1}{2 R_{1}}+\frac{18 R_{1}}{\delta^{2}}\right) \\
& \quad \leq C_{\delta}\left(R_{1}\right)^{2}+\frac{12}{\delta^{2} n_{0}}\left(\frac{C}{\alpha_{1}^{2}}+\frac{3}{\alpha_{1}} K\left\|\varphi_{1}\right\|_{\infty}^{2}\right)+\frac{C^{\prime}}{n_{0}}+\frac{\left\|\varphi_{1}\right\|_{\infty}^{2}}{2 \alpha_{1} \epsilon_{1} n_{0}}+\frac{\left\|\varphi_{1}\right\|_{\infty}^{2}}{\alpha_{1} \epsilon_{1} n_{0}} \tag{6.123}
\end{align*}
$$

where the second inequality is due to $\rho_{1}\left(n_{0}\right) \leq\left\|\varphi_{1}\right\|_{\infty}^{2}, \epsilon_{1} \leq R_{1} \leq K$ with $K:=1+r_{\text {max }}$, and $\frac{n_{0}}{n_{1}} \leq \frac{2}{\alpha_{1}}$ for $n_{0}$ large enough. Note that

$$
\begin{align*}
\left|C_{\delta}\left(R_{1}\right)\right|=C_{\delta}\left(R_{1}\right) \leq-\widetilde{f^{\prime}}\left(R_{1}-\delta\right) & \leq \frac{n_{1}}{2 n_{0}}\left\|\varphi_{1}\right\|_{\infty}\left(\left\|\varphi_{1}\right\|_{\infty}+\frac{1}{\sqrt{R_{1}-\delta}}\right) \\
& \leq \alpha_{1}\left\|\varphi_{1}\right\|_{\infty}\left(\left\|\varphi_{1}\right\|_{\infty}+\frac{1}{\sqrt{s_{n_{0}}-\delta}}\right) \tag{6.124}
\end{align*}
$$

The first inequality follows from $\widetilde{f^{\prime}}$ being a nonpositive function, and the last one from $R_{1} \geq s_{n_{0}}$ and $n_{1} / n_{0} \leq 2 \alpha_{1}$ for $n_{0}$ large enough. We use (6.124) to obtain

$$
\begin{aligned}
\int_{\mathcal{B}_{n_{0}}} d \epsilon C_{\delta}\left(R_{1}(t, \epsilon)\right)^{2} \leq & \alpha_{1}\left(\left\|\varphi_{1}\right\|_{\infty}^{2}+\frac{\left\|\varphi_{1}\right\|_{\infty}}{\sqrt{s_{n_{0}}-\delta}}\right) \int_{\mathcal{B}_{n}} d \epsilon C_{\delta}\left(R_{1}(t, \epsilon)\right) \\
= & \alpha_{1}\left(\left\|\varphi_{1}\right\|_{\infty}^{2}+\frac{\left\|\varphi_{1}\right\|_{\infty}}{\sqrt{s_{n_{0}}-\delta}}\right) \int_{R\left(\mathcal{B}_{n_{0}}\right)} \frac{d R_{1} d R_{2} C_{\delta}\left(R_{1}\right)}{\left|J_{R}\left(R^{-1}\left(R_{1}, R_{2}\right)\right)\right|} \\
\leq & \alpha_{1}\left(\left\|\varphi_{1}\right\|_{\infty}^{2}+\frac{\left\|\varphi_{1}\right\|_{\infty}}{\sqrt{s_{n_{0}}-\delta}}\right) \int_{R\left(\mathcal{B}_{n_{0}}\right)} d R_{1} d R_{2} C_{\delta}\left(R_{1}\right) \\
= & \alpha_{1}\left(\left\|\varphi_{1}\right\|_{\infty}^{2}+\frac{\left\|\varphi_{1}\right\|_{\infty}}{\sqrt{s_{n_{0}}-\delta}}\right) \\
& \cdot \int_{s_{n_{0}}}^{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right)} d R_{2}\left(\widetilde{f}\left(2 s_{n_{0}}+r_{\max }+\delta\right)-\widetilde{f}\left(2 s_{n_{0}}+r_{\max }-\delta\right)\right. \\
& \left.\quad+\left(\widetilde{f}\left(s_{n_{0}}-\delta\right)-\widetilde{f}\left(s_{n_{0}}+\delta\right)\right)\right)
\end{aligned}
$$

By the mean value theorem, we finally have

$$
\begin{align*}
& \int_{\mathcal{B}_{n_{0}}} d \epsilon C_{\delta}\left(R_{1}(t, \epsilon)\right)^{2} \\
& \quad \leq \alpha_{1}\left(\left\|\varphi_{1}\right\|_{\infty}^{2}+\frac{\left\|\varphi_{1}\right\|_{\infty}}{\sqrt{s_{n_{0}}-\delta}}\right) \int_{s_{n_{0}}}^{2 s_{n_{0}}+\rho_{1}\left(n_{0}\right)} d R_{2} 4 \delta \alpha_{1}\left(\left\|\varphi_{1}\right\|_{\infty}^{2}+\frac{\left\|\varphi_{1}\right\|_{\infty}}{\sqrt{s_{n_{0}}-\delta}}\right) \\
& \quad=4 \delta \alpha_{1}^{2}\left(\left\|\varphi_{1}\right\|_{\infty}^{2}+\frac{\left\|\varphi_{1}\right\|_{\infty}}{\sqrt{s_{n_{0}}-\delta}}\right)^{2}\left(s_{n_{0}}+\left\|\varphi_{1}\right\|_{\infty}^{2}\right) . \tag{6.125}
\end{align*}
$$

Integrating (6.123) over $\epsilon \in \mathcal{B}_{n_{0}}$ and making use of 6.125 yields

$$
\begin{aligned}
\int_{\mathcal{B}_{n_{0}}} d \epsilon \mathbb{E}\left(\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}-\mathbb{E}\langle\mathcal{L}\rangle_{\boldsymbol{n}, t, \epsilon}\right)^{2} \leq 4 \delta \alpha_{1}^{2}( & \left.\left\|\varphi_{1}\right\|_{\infty}^{2}+\frac{\left\|\varphi_{1}\right\|_{\infty}}{\sqrt{s_{n_{0}}-\delta}}\right)^{2}\left(s_{n_{0}}+\left\|\varphi_{1}\right\|_{\infty}^{2}\right) \\
& +\frac{84 s_{n_{0}}^{2}}{\delta^{2} n_{0}}\left(\frac{C}{\alpha_{1}^{2}}+\frac{3}{\alpha_{1}} K\left\|\varphi_{1}\right\|_{\infty}^{2}\right) \\
& +\frac{C^{\prime} s_{n_{0}}^{2}}{n_{0}}+\frac{21\left\|\varphi_{1}\right\|_{\infty}^{2} \ln (2) s_{n_{0}}}{2 \alpha_{1} n_{0}} .
\end{aligned}
$$

Finally, we obtain the desired result from this last upper bound by using that $s_{n_{0}} \leq 1 / 2$ and choosing $\delta=s_{n_{0}} n_{0}^{-1 / 4}$.

## Information theoretic limits of learning a sparse rule

## 7

### 7.1 Introduction

Modern tasks in statistical analysis, signal processing and learning require solving high-dimensional inference problems with a very large number of parameters. This arises in areas as diverse as learning with neural networks [140], high-dimensional regression [141] or compressed sensing [77], [142]. In many situations, there appear barriers to what is possible to estimate or learn when the data becomes too scarce or too noisy. Such barriers can be of algorithmic nature, but they can also be intrinsic to the very nature of the problem. A celebrated example is the impossibility of reconstructing a noisy signal when the noise is beyond the so-called Shannon capacity of the communication channel [50]. A large amount of interdisciplinary work has shown that these intrinsic barriers can be understood as static phase transitions (in the sense of physics) when the system size tends to infinity (see [64, [143], 144]).

When the problem can be formulated as an (optimal) Bayesian inference problem the mathematically rigorous theory of these phase transitions is now quite well developed. Progress initially came from applications of the GuerraToninelli interpolation method (developed for the Sherrington-Kirkpatrick spinglass model [145]) to coding and communication theory [43], [146]-[150], and more recently to low-rank matrix and tensor estimation [34], [36], [46], [66], [69], [72], [89], [151], [152], compressive sensing and high-dimensional regression [44], [73, [74, [153], and generalized linear models [29]. In particular, for all these problems it has been possible to reduce the asymptotic mutual information to a low-dimensional variational expression, and deduce from its solution relevant error measures (e.g., minimum mean-square and generalization errors). All these works consider the traditional regime of statistical mechanics where the system size goes to infinity while relevant control parameters (such as signal sparsity, sampling rate, or signal-to-noise ratio) are kept fixed.

However, there exist other interesting regimes for which many of the above mentioned problems also display fundamental intrinsic limits akin to phase transi-
tions. Consider for example the problem of compressive sensing. An interesting regime is one where both the number of nonzero components and of samples scale in a sublinear manner as the system size tends to infinity. In this case we would like to identify the phase transition, if there is any, and its nature. This question has first been addressed recently in the framework of compressed sensing for binary Bernoulli signals by [32], [33], [154]. An all-or-nothing phenomenon is identified, that is, in an appropriate sparse regime, the minimum mean-square error (MMSE) sharply drops from its maximum possible value (no reconstruction) for "too small" sampling rates to zero (perfect reconstruction) for "large enough" sampling rates. The interest of such regime is not limited to estimation problems. It is also relevant from a learning point of view, e.g., it corresponds to learning scenarios where we have access to a high number of features but only a sublinear number of them - unknown to us - are relevant for the learning task at hand.

Examples abound where the "bet on sparsity principle" [155], [156] is of utmost importance for the interpretability of a high-dimensional model. Let us mention the MNIST handwritten digit database, where each digit can be seen as a $784=28 \times 28$-dimensional binary vector representing the pixels whereas the digits effectively live in a space of the order of tens of dimensions [157, [158]. Another example of effective sparsity comes from natural images which are often sparse in a wavelet basis [112]. Then, a fundamental question is "when is it possible to achieve a low estimation or generalization error with a sublinear amount of samples (sublinear with respect to the total number of features)?"

In this contribution we address this question for a mathematically simple, but precise and tractable, setting. We consider generalized linear models in the regime of vanishing sparsity and sample rate, or equivalently, of sublinear number of data samples and nonzero signal components. As explained below these models can be used for estimation as well as learning, and we uncover in the sublinear regime intrinsic statistical barriers to these tasks in the form of sharp phase transitions. These statistical barriers are computed exactly and thus provide precise benchmarks to which algorithmic performance can be compared.

Let us outline the mathematical setting (further detailed in Section 7.2). In a probabilistic setting the unknown signal vector $\mathbf{X}^{*} \in \mathbb{R}^{n}$ has entries drawn independently at random from a distribution $P_{X}^{(n)}:=\rho_{n} P_{0}+\left(1-\rho_{n}\right) \delta_{0}$ with $P_{0}$ a fixed distribution. The parameter $\rho_{n}$ controls the sparsity of the signal so that $\mathbf{X}^{*}$ has $k_{n}:=n \rho_{n}$ nonzero components on average. We observe the data $\mathbf{Y}=\varphi\left(\mathbf{W X}^{*} / \sqrt{k_{n}}\right) \in \mathbb{R}^{m_{n}}$ obtained by first multiplying the signal with a known $m_{n} \times n$ random matrix $\mathbf{W}$ whose entries are independent standard Gaussian random variables, and then applying $\varphi$ component-wise. The number of data points is controlled by the sampling rate $\alpha_{n}$, i.e., $m_{n}:=\alpha_{n} n$. We consider the regime $\left(\rho_{n}, \alpha_{n}\right) \rightarrow(0,0)$ as $n$ goes to infinity with $\alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$, for which sharp phase transitions appear when $P_{0}$ is discrete with finite support. Note that both $m_{n}$ and $k_{n}$ scale sublinearly as $n \rightarrow+\infty$.

The model can be interpreted as either an estimation problem or a learning problem:

- In the estimation interpretation, we assume a purely Bayesian (or optimal)
setting. We know the model, the activation function $\varphi$, the prior $P_{X}^{(n)}$ as well as the measurement matrix $\mathbf{W}$. Our goal is then to determine what is the lowest reconstruction error that we can achieve, i.e., what is the average minimum mean-square error $k_{n}^{-1} \mathbb{E}\left\|\mathbf{X}^{*}-\mathbb{E}\left[\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right]\right\|^{2}$ when $n$ gets large.
- In the learning interpretation, we consider a teacher-student scenario in which a teacher hands out training samples $\left\{\left(Y_{\mu},\left(W_{\mu i}\right)_{i=1}^{n}\right)\right\}_{\mu=1}^{m_{n}}$ to a student. The teacher produces the output label $Y_{\mu}$ by feeding the input $\left(W_{\mu i}\right)_{i=1}^{n}$ to its own one-layer neural network with activation function $\varphi$ and weights $\mathbf{X}^{*}=\left(X_{i}^{*}\right)_{i=1}^{n}$. The student - who is given the model and the prior - has to learn the weights $\mathrm{X}^{*}$ of the teacher's one-layer neural network by minimizing the empirical training error of the $m_{n}$ training samples. For example, the binary perceptron corresponds to $\varphi=\operatorname{sign}$ and $Y_{\mu} \in\{ \pm 1\}$. Of particular interest is the generalization error. Given a new - previously unseen - random pattern $\mathbf{W}_{\text {new }}:=\left(W_{\text {new },}\right)_{i=1}^{n}$ whose true label is $Y_{\text {new }}$ (generated by the teacher's neural network), the optimal generalization error is $\mathbb{E}\left[\left(Y_{\text {new }}-\mathbb{E}\left[\varphi\left(\mathbf{W}_{\text {new }}^{\top} \mathbf{X}^{*} / \sqrt{k_{n}}\right) \mid \mathbf{Y}, \mathbf{W}, \mathbf{W}_{\text {new }}\right]\right)^{2}\right]$; the error made when estimating $Y_{\text {new }}$ in a purely Bayesian way.

Let us summarize informally our results. We set $\alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ where $\gamma$ is fixed and $\rho_{n}$ vanishes as $n$ diverges. We first rigorously determine the mutual information $m_{n}^{-1} I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)$ in terms of a low-dimensional variational problem, see Theorem 7.1 which also provides a precise control of the finite size fluctuations. Remarkably, when $P_{0}$ is a discrete distribution with finite support, this variational problem simplifies to a minimization problem over a finite set of values, see Theorem 7.2. For such signals, using I-MMSE type formulas [51], we can deduce from the solution to this minimization problem the asymptotic MMSE and optimal generalization error, see Theorem 7.3. Our analysis shows that both errors are nonincreasing piecewise constant functions of $\gamma$. In particular, if the entries of $\left|\mathbf{X}^{*}\right|$ are either 0 or some $a>0$ then both errors display an all-or-nothing behavior as $n \rightarrow+\infty$, with a sharp transition at a threshold $\gamma=\gamma_{c}$ explicitly computed. These findings are illustrated, and their significance discussed, in Section 7.3.

In this chapter the generalized linear model is treated by entirely different methods than the linear model in [32], [33]. Importantly, the sparsity regime treated by our method requires the sparsity $\rho_{n}$ to go to zero slower than $n^{-1 / 9}$, while it has to go to zero faster than $n^{-1 / 2}$ in the results of $|33|$ for the linear case. From this angle, both results complement each other. Our proof technique for Theorem 7.1 exploits the adaptive interpolation method. We adapt the analysis of [29] in a non-trivial way in order to consider the new scaling regime of our problem where $\alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$, and $\rho_{n} \rightarrow 0$ as $n$ gets large instead of being fixed. We show that the adaptive interpolation can still be carried through, which requires a more refined control of the error terms compared to [29]. It is interesting, and not a priori obvious, that this can be done since this is not the usual extensive regime of statistical mechanics. For example, the mutual information has to be normalized by the subextensive quantity $m_{n}=o(n)$. Quite remarkably, with this suitable normalization, the asymptotic normalized mutual information, MMSE
and optimal generalization error have a similar form to those famously found in ordinary thermodynamic regimes in physics [159]-162.

In Section 7.2 we present the setting and state our theoretical results on the normalized mutual information and MMSE in the sublinear regime. We use these results in Section 7.3 to uncover the all-or-nothing phenomenon for general activation functions. In Section 7.4 we show how to simplify the formula for the asymptotic normalized mutual information of Theorem 7.1 when $\mathbf{X}^{*}$ is a sparse binary signal (i.e., its entries are zero or one with a vanishing probability $\rho_{n}$ of being one). This simplified formula corresponds to Theorem 7.2 and demonstrates the existence of the all-or-nothing phenomenon for such sparse binary signals. Besides, the proof in Section 7.4 presents all the main ideas in order to prove Theorem 7.2 for more general priors $P_{X}^{(n)}$ while being simpler due to the binary nature of the signal. The full proofs of our results are given in the appendices.

### 7.2 Problem setting and main results

### 7.2.1 Generalized linear estimation of low sparsity signals at low sampling rates

For all $n \in \mathbb{N}^{*}$, we define the probability distribution

$$
\begin{equation*}
P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}, \tag{7.1}
\end{equation*}
$$

where $\delta_{0}$ is the Dirac distribution (the distribution fully suported on $\{0\}$ ), $P_{0}$ is a probability distribution over $\mathbb{R}$ with finite second moment $\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]$, and $\left(\rho_{n}\right)_{n \in \mathbb{N}^{*}}$ is a decreasing sequence of real numbers in $(0,1)$. Let $\left(X_{i}^{*}\right)_{i=1}^{n} \stackrel{\text { i.i.d. }}{\sim} P_{X}^{(n)}$ be the components of a $n$-dimensional signal vector $\mathbf{X}^{*}$. This is also denoted by $\mathbf{X}^{*} \stackrel{\text { i.i.d. }}{\sim} P_{X}^{(n)}$. The parameter $\rho_{n} \in(0,1)$ controls the sparsity of the signal; the latter being made of $k_{n}:=\rho_{n} n$ nonzero components in expectation. We are interested in low sparsity regimes where $k_{n}=o(n)$.

Let $k_{A}$ be a nonnegative integer, $P_{A}$ a probability distribution over $\mathbb{R}^{k_{A}}$, and $\varphi: \mathbb{R} \times \mathbb{R}^{k_{A}} \rightarrow \mathbb{R}$ a measurable function. Let $m_{n}:=\alpha_{n} n$ where $\left(\alpha_{n}\right)_{n \in \mathbb{N}^{*}}$ is a decreasing sequence of positive sampling rates. We have access to $m_{n}$ data points $\mathbf{Y}:=\left(Y_{\mu}\right)_{\mu=1}^{m_{n}}$ generated as

$$
\begin{equation*}
Y_{\mu}:=\varphi\left(\frac{\left(\mathbf{W X}^{*}\right)_{\mu}}{\sqrt{k_{n}}}, \mathbf{A}_{\mu}\right)+\sqrt{\Delta} Z_{\mu}, \quad 1 \leq \mu \leq m_{n} \tag{7.2}
\end{equation*}
$$

where $\left.\left.\left(\mathbf{A}_{\mu}\right)\right)_{\mu=1}^{m_{n}} \stackrel{\text { i.i.d. }}{\sim} P_{A},\left(Z_{\mu}\right)\right)_{\mu=1}^{m} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ is an additive white Gaussian noise (AWGN), $\Delta>0$ is the noise variance, and $\mathbf{W}$ is a $m_{n} \times n$ measurement (or data) matrix with independent entries having zero mean and unit variance. Note that the noise $\left(Z_{\mu}\right)_{\mu=1}^{m}$ can be considered as part of the model, or as a "regularising" noise needed for the analysis but that can be set arbitrarily small by tuning $\Delta$. Typically, and as $n$ gets large, $\left(\mathbf{W} \mathbf{X}^{*}\right)_{\mu} / \sqrt{k_{n}}=\Theta(1)$. The estimation problem is to recover $\mathbf{X}^{*}$ from the knowledge of $\mathbf{Y}, \mathbf{W}, \Delta, \varphi, P_{X}^{(n)}$ and $P_{A}$. The realization of
the random stream $\left(\mathbf{A}_{\mu}\right)_{\mu=1}^{m_{n}}$ itself, if present in the model, is unknown. It will be helpful to think of the measurements as the outputs of a channel,

$$
\begin{equation*}
Y_{\mu} \sim P_{\mathrm{out}}\left(\cdot \left\lvert\, \frac{\left(\mathbf{W X}^{*}\right)_{\mu}}{\sqrt{k_{n}}}\right.\right), \quad 1 \leq \mu \leq m_{n} \tag{7.3}
\end{equation*}
$$

The transition kernel $P_{\text {out }}$ admits a transition density with respect to Lebesgue's measure given by

$$
\begin{equation*}
P_{\text {out }}(y \mid x):=\frac{1}{\sqrt{2 \pi \Delta}} \int d P_{A}(\mathbf{a}) e^{-\frac{1}{2 \Delta}(y-\varphi(x, \mathbf{a}))^{2}} . \tag{7.4}
\end{equation*}
$$

The random stream $\left(\mathbf{A}_{\mu}\right)_{\mu=1}^{m_{n}}$ represents any source of randomness in the model. For example, the logistic regression

$$
\mathbb{P}\left(Y_{\mu}=1\right)=f\left(\left(\mathbf{W} \mathbf{X}^{*}\right)_{\mu} / \sqrt{k_{n}}\right), \quad \mathbb{P}\left(Y_{\mu}=-1\right)=1-f\left(\left(\mathbf{W} \mathbf{x}^{*}\right)_{\mu} / \sqrt{k_{n}}\right),
$$

where $f(x):=\left(1+e^{-\lambda x}\right)^{-1}$, is modeled by considering a teacher that draws i.i.d. uniform numbers $A_{\mu} \sim \mathcal{U}(0,1)$ and then obtains the labels through

$$
Y_{\mu}=\mathbf{1}_{\left\{A_{\mu} \leq f\left(\left(\mathbf{w x}^{*}\right) \mu / \sqrt{k_{n}}\right)\right\}}-\mathbf{1}_{\left\{A_{\mu} \geq f\left(\left(\mathbf{w x}^{*}\right) \mu / \sqrt{k_{n}}\right)\right\}},
$$

where $\mathbf{1}_{\mathcal{E}}$ denotes the indicator function of an event $\mathcal{E}$. In the absence of such a randomness in the model, the activation $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ is deterministic, $k_{A}=0$ and the integral $\int d P_{A}(\mathbf{a})$ in (7.4) simply disappears. Our numerical experiments in Section 7.3 are for deterministic activations but all of our theoretical results hold for the broader setting.

We have presented the problem from an estimation point of view. In this case, the important quantity to assess the performance of an algorithm estimating $\mathbf{X}^{*}$ is the mean-square error. Another point of view is the learning one where each row $\mathbf{W}_{\mu}$, of the matrix $\mathbf{W}$ is the input to a one-layer neural network whose weights $\mathbf{X}^{*}$ have been sampled independently at random by a teacher. The student is given the input/output pairs $\left(\mathbf{W}_{\mu,}, Y_{\mu}\right)_{\mu=1}^{m_{n}}$ as well as the model used by the teacher. The student's role is then to learn the weights. In this case, more than the mean-square error, the important quantity is the generalization error.

### 7.2.2 Asymptotic normalized mutual information

The mutual information $I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)$ between the signal $\mathbf{X}^{*}$ and the data $\mathbf{Y}$ given the matrix $\mathbf{W}$ is the main quantity of interest in our work. Before stating Theorem 7.1 on the value of this mutual information, let us introduce two scalar denoising models that play a key role.

The first model is an additive Gaussian channel. Let $X^{*} \sim P_{X}^{(n)}$ be a scalar random variable. We observe $Y^{(r)}:=\sqrt{r} X^{*}+Z$ where $r \geq 0$ plays the role of a signal-to-noise ratio (SNR) and the noise $Z \sim \mathcal{N}(0,1)$ is independent of $X^{*}$. We denote by $I_{P_{X}^{(n)}}(r):=I\left(X^{*} ; Y^{(r)}\right)$ the mutual information between the signal of interest $X^{*}$ and $Y^{(r)}$. It depends on $\rho_{n}$ through the prior $P_{X}^{(n)}$ and reads

$$
\begin{equation*}
I_{P_{X}^{(n)}}(r)=\frac{r \rho_{n} \mathbb{E}\left[X_{0}^{2}\right]}{2}-\mathbb{E} \ln \int d P_{X}^{(n)}(x) e^{r X^{*} x+\sqrt{r} Z x-\frac{r x^{2}}{2}}, \tag{7.5}
\end{equation*}
$$

where $X_{0} \sim P_{0}$. The function $I_{P_{X}^{(n)}}$ is nondecreasing, convex and Lipschitz continuous with Lipschitz constant $\rho_{n} \mathbb{E}\left[X_{0}^{2}\right] / 2$ on $[0,+\infty)$, see Lemma 2.3 .

The second scalar channel is linked to the transition kernel $P_{\text {out }}$ defined in (7.4). Let $U$ and $V$ be two independent standard Gaussian random variables. In this scalar estimation problem, we want to infer $U$ from the knowledge of $V$ and the observation $\widetilde{Y}^{(q, \rho)} \sim P_{\text {out }}(\cdot \mid \sqrt{\rho-q} U+\sqrt{q} V)$ where $\rho>0$ and $q \in[0, \rho]$. Equivalently, $\widetilde{Y}^{(q, \rho)}:=\varphi(\sqrt{\rho-q} U+\sqrt{q} V)+\sqrt{\Delta} Z$ with $Z \sim \mathcal{N}(0,1)$ independent of $(U, V)$. We denote by $I_{P_{\text {out }}}(q, \rho):=I\left(U ; \widetilde{Y}^{(q, \rho)} \mid V\right)$ the conditional mutual information between $U$ and $\widetilde{Y}^{(q, \rho)}$ given $V$. Note that
$I_{P_{\text {out }}}(q, \rho)=\mathbb{E} \ln P_{\text {out }}\left(\widetilde{Y}^{(\rho, \rho)} \mid \sqrt{\rho} V\right)-\mathbb{E} \ln \int d u \frac{e^{-\frac{u^{2}}{2}}}{\sqrt{2 \pi}} P_{\text {out }}\left(\widetilde{Y}^{(q, \rho)} \mid \sqrt{\rho-q} u+\sqrt{q} V\right)$.
Like $I_{P_{X}^{(n)}}$, the function $I_{P_{\text {out }}}$ has nice monotonicity, concavity and Lipschitzness properties that are listed in Lemma 7.22 and are important for the proof of Theorem 7.1 stated below.

We use the mutual informations (7.5) and (7.6) to define the replica symmetric potential

$$
\begin{equation*}
i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right):=\frac{1}{\alpha_{n}} I_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)+I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)-\frac{r\left(\mathbb{E}\left[X_{0}^{2}\right]-q\right)}{2} \tag{7.7}
\end{equation*}
$$

where $X_{0} \sim P_{0}$. Our first result links the extrema of this potential to the mutual information of our original problem.

Theorem 7.1 (Normalized mutual information of the GLM at sublinear sparsity and sampling rate). Suppose that $\Delta>0$ and that the following hypotheses hold:
(H1) There exists $S>0$ such that the support of $P_{0}$ is included in $[-S, S]$.
(H2) $\varphi$ is bounded, and its first and second partial derivatives with respect to its first argument exist, are bounded and continuous. They are denoted $\partial_{x} \varphi$, $\partial_{x x} \varphi$.
(H3) $W_{\mu i} \stackrel{i . i . d .}{\sim} \mathcal{N}(0,1)$ for $(\mu, i) \in\left\{1, \ldots, m_{n}\right\} \times\{1, \ldots, n\}$.
Let $\rho_{n}=\Theta\left(n^{-\lambda}\right)$ with $\lambda \in[0,1 / 9)$ and $\alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ with $\gamma>0$. Then, for all $n \in \mathbb{N}^{*}$ :

$$
\left|\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}-\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)\right| \leq \frac{\sqrt{C}|\ln n|^{1 / 6}}{n^{\frac{1}{12}-\frac{31}{4}}}
$$

where $C$ is a polynomial in $\left(S,\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x}}{\sqrt{\Delta}}\right\|_{\infty}, \lambda, \gamma\right)$ with positive coefficients.

Hence, the asymptotic mutual information is given to leading order by the variational formula $\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$. Note that this variational formula depends on $\rho_{n}$. Theorem 7.1 does not say anything on its value in the asymptotic regime, e.g., does it converge or diverge? Our next theorem answers this question when $P_{0}$ is a discrete distribution with finite support.

### 7.2.3 Specialization to discrete priors: all-or-nothing phenomenon and its generalization

Theorem 7.2 (Specialization of Theorem 7.1 to discrete priors with finite support). Suppose that $\Delta>0$ and that $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ where $P_{0}$ is a discrete distribution with finite support

$$
\operatorname{supp}\left(P_{0}\right) \subseteq\left\{-v_{K},-v_{K-1}, \ldots,-v_{1}, v_{1}, v_{2}, \ldots, v_{K}\right\}
$$

where $0<v_{1}<v_{2}<\cdots<v_{K}<v_{K+1}:=+\infty$. Further assume that the hypotheses (H2) and (H3) in Theorem 7.1 hold. Let $\rho_{n}=\Theta\left(n^{-\lambda}\right)$ with $\lambda \in(0,1 / 9)$ and $\alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ with $\gamma>0$. Then,

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}=\min _{1 \leq k \leq K+1} I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} \tag{7.8}
\end{equation*}
$$

where $X_{0} \sim P_{0}$.
To prove Theorem 7.2 we compute the limit of $\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$ when $\rho_{n}$ vanishes. We show how to compute this limit for binary sparse signals, i.e., $P_{0}=\delta_{1}$, in Section 7.4. We give the proof for a general discrete distribution with finite support $P_{0}$ in Appendix 7.B

When doing estimation, one important metric to assess the quality of an estimator $\widehat{\mathbf{X}}(\mathbf{Y}, \mathbf{W})$ is its mean-square error $\mathbb{E}\left\|\mathbf{X}^{*}-\widehat{\mathbf{X}}(\mathbf{Y}, \mathbf{W})\right\|^{2} / k_{n}$. The latter is always lower bounded by the mean-square error of the Bayesian estimator $\mathbb{E}\left[\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right]$; the so-called minimum mean-square error (MMSE). Remarkably, once we have Theorem 7.2 , we can obtain the asymptotic MMSE with a little more work. First, we have to introduce a modified inference problem where in addition to the observations $\mathbf{Y}$ we are given $\widetilde{\mathbf{Y}}{ }^{(\tau)}=\sqrt{\alpha_{n} \tau / \rho_{n}} \mathbf{X}^{*}+\widetilde{\mathbf{Z}}$. When $\tau$ is close enough to 0 , the analysis yielding Theorem 7.2 can be adapted to obtain the limit

$$
\begin{aligned}
\lim _{n \rightarrow+\infty} & \frac{I\left(\mathbf{X}^{*} ; \mathbf{Y}, \tilde{\mathbf{Y}}^{(\tau)} \mid \mathbf{W}\right)}{m_{n}} \\
& =\min _{1 \leq k \leq K+1} I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}+\frac{\tau \mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k}\right\}}\right]}{2} .
\end{aligned}
$$

We can then apply the I-MMSE relationshir [18], [51] to obtain the asymptotic MMSE.

Theorem 7.3 (Asymptotic MMSE). Under the assumptions of Theorem 7.2, if the minimization problem on the right-hand side of (7.8) has a unique solution $k^{*} \in\{1, \ldots, K+1\}$ then

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{\mathbb{E}\left\|\mathbf{X}^{*}-\mathbb{E}\left[\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right]\right\|^{2}}{k_{n}}=\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k^{*}}\right\}}\right], \tag{7.9}
\end{equation*}
$$

where $X_{0} \sim P_{0}$.

[^19]We prove Theorem 7.3 in Appendix 7.C. We remark that it is possible with more technical work [29, Appendix C.2] to weaken (H2) in Theorems 7.2 and 7.3 to the assumption

There exists $\epsilon>0$ such that the sequence $\mathbb{E}\left|\varphi\left(\left(\mathbf{W X}^{*}\right)_{1} / \sqrt{k_{n}}, \mathbf{A}_{1}\right)\right|^{2+\epsilon}$ is bounded, and for almost all $\mathbf{a} \sim P_{A}$ the function $x \mapsto \varphi(x, \mathbf{a})$ is continuous almost everywhere.

Hence, Theorems 7.2 and 7.3 also apply to the linear activation $\varphi(x)=x$, the perceptron $\varphi(x)=\operatorname{sign}(x)$ and the $\operatorname{ReLU} \varphi(x)=\max (0, x)$.

### 7.3 The all-or-nothing phenomenon

We now highlight interesting consequences of our results regarding the MMSE of the estimation problem as well as the optimal generalization error of the learning problem in the teacher-student scenario. Reeves et al. [33] have proved the existence of an all-or-nothing phenomenon for the linear model when $\mathbf{X}^{*}$ is a 0-1 vector and here we extend their results in two ways: $i$ ) for the estimation error of a generalized linear model, and $i i$ ) for the generalization error of a perceptron neural network with general activation function $\varphi$.

We consider signals whose entries are either Bernoulli random variables, i.e., $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ with $P_{0}=\delta_{1}$, or Bernoulli-Rademacher random variables, i.e., $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ with $P_{0}=\left(\delta_{1}+\delta_{-1}\right) / 2$. In both cases, the second moment of a random variable $X_{0} \sim P_{0}$ is $\rrbracket^{2} \mathbb{E} X_{0}^{2}=1$. We place ourselves in the regime of Theorem 7.3 where $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for some fixed $\gamma>0$ and $\rho_{n} \rightarrow 0$ in the high-dimensional limit $n \rightarrow+\infty$.

MMSE In this regime, and for such signals, Theorem 7.3 states that the minimum mean-square error $\operatorname{MMSE}\left(\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right):=\frac{\mathbb{E}\left\|\mathbf{X}^{*}-\mathbb{E}\left[\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right]\right\|^{2}}{k_{n}}$ satisfies:

$$
\lim _{n \rightarrow+\infty} \operatorname{MMSE}\left(\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right)= \begin{cases}0 & \text { if } I_{P_{\text {out }}}(0,1)>\gamma^{-1}  \tag{7.10}\\ 1 & \text { if } I_{P_{\text {out }}}(0,1)<\gamma^{-1}\end{cases}
$$

Therefore, we locate an all-or-nothing phase transition at the threshold

$$
\begin{equation*}
\gamma_{c}:=\frac{1}{I_{P_{\text {out }}}(0,1)} . \tag{7.11}
\end{equation*}
$$

Remember that $\gamma$ controls the amount $m_{n}$ of training samples. In the highdimensional limit, perfect reconstruction is possible if $\gamma>\gamma_{c}$ (the asymptotic MMSE is zero) while it is impossible to do better than a random guess if $\gamma<\gamma_{c}$ (the asymptotic MMSE is equal to $\lim _{n \rightarrow+\infty} \mathbb{E}\left\|\mathbf{X}^{*}-\mathbb{E} \mathbf{X}^{*}\right\|^{2} / k_{n}=1$; the asymptotic MMSE in the absence of observations). As $I_{P_{\text {out }}}(0,1):=I(U ; \varphi(U, \mathbf{A})+\sqrt{\Delta} Z)$ where $U, Z \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1) \perp \mathbf{A} \sim P_{A}$, the threshold $\gamma_{c}$ is fully determined by

[^20]

Figure 1: Threshold $\gamma_{c}$ of the all-or-nothing phase transition for different activation functions as a function of the noise variance $\Delta$.

| Activation $\varphi(x)$ | $\gamma_{c}(\Delta=0)$ | $\gamma_{c}(\Delta)$ for $\Delta>0$ |
| :---: | :---: | :--- |
| $x$ | 0 | $2 / \ln \left(1+\Delta^{-1}\right)$ |
| $\operatorname{sign}(x)$ | $1 / \ln 2$ | $1 /\left(\ln 2-\mathbb{E}\left[\ln \left(1+e^{-2(1+\sqrt{\Delta} Z) / \Delta}\right)\right]\right)$ |
| $\max (0, x)$ | 0 | $4 \Delta /\left(1-4 \Delta \mathbb{E}\left[h_{\Delta}(Z) \ln h_{\Delta}(Z)\right]\right)$ |
|  |  | with $h_{\Delta}(Z):=\frac{1}{2}+\sqrt{\frac{\Delta}{1+\Delta}} e^{\frac{Z^{2}}{2(1+\Delta)}} \int_{-\infty}^{\frac{Z}{\sqrt{1+\Delta}}} \frac{d t}{\sqrt{2 \pi}} e^{-\frac{t^{2}}{2}}$ |

Table 1: Closed-formed formulas of $\gamma_{c}$ for different activation functions. We use a standard Gaussian random variable $Z \sim \mathcal{N}(0,1)$ in the formulas.
the activation function and the amount of noise, and it can be easily evaluated in a number of cases. In Figure 1 we draw $\gamma_{c}$ for $\varphi(x)=x, \varphi(x)=\operatorname{sign}(x)$, $\varphi(x)=\max (0, x)$ and noise variance $\Delta \in[0,0.5]$. We see that for $\Delta$ small enough the ReLU activation requires less training samples to learn the sparse rule than the linear one; it is the opposite once $\Delta$ becomes large enough. When $\Delta$ diverges both the linear and sign activations have the asymptote $\gamma_{c} \sim 2 \Delta$ while the ReLU activation has another steeper asymptote $\gamma_{c} \sim a \Delta, a \approx 5.87$. The corresponding formulas for $\gamma_{c}$ are given in Table 1. Note that for the random linear model $\varphi(x)=x$, the threshold $\alpha_{c}\left(\rho_{n}\right):=\gamma_{c} \rho_{n}\left|\ln \rho_{n}\right|=2 \rho_{n}\left|\ln \rho_{n}\right| / \ln \left(1+\Delta^{-1}\right)$ is in agreement with the sample rate $n^{*}$ for which [33] prove that weak recovery is impossible below it while strong recovery is possible above.

Optimal generalization error When learning in a (matched) teacher-student scenario, the components of $\mathbf{X}^{*}$ correspond to the unknown weights of the teacher's one-layer neural network. The student is given the model and training samples $\left\{\left(Y_{\mu},\left(W_{\mu i}\right)_{i=1}^{n}\right)\right\}_{\mu=1}^{m_{n}}$. Then, the optimal generalization error is the MMSE for predicting the output

$$
Y_{\text {new }} \sim P_{\text {out }}\left(\cdot \left\lvert\, \frac{\mathbf{W}_{\text {new }}^{\top} \mathbf{X}^{*}}{\sqrt{k_{n}}}\right.\right)
$$

generated by a new input $\mathbf{W}_{\text {new }}:=\left(W_{\text {new }, i}\right) \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. More precisely, the optimal generalization error is

$$
\operatorname{MMSE}\left(Y_{\text {new }} \mid \mathbf{Y}, \mathbf{W}, \mathbf{W}_{\text {new }}\right):=\mathbb{E}\left[\left(Y_{\text {new }}-\mathbb{E}\left[Y_{\text {new }} \mid \mathbf{Y}, \mathbf{W}, \mathbf{W}_{\text {new }}\right]\right)^{2}\right]
$$

Based on our proof of Theorem 7.3 and the formula for the optimal generalization error when $\rho_{n}=\Theta(1)$ (regime of linear sparsity and sampling rate) 29, Theorem 2] we conjecture that, under the assumptions of Theorem 7.3 ,

$$
\begin{align*}
& \lim _{n \rightarrow+\infty} \operatorname{MMSE}\left(Y_{\text {new }} \mid \mathbf{Y}, \mathbf{W}, \mathbf{W}_{\text {new }}\right) \\
& \quad=\Delta+\mathbb{E}\left[\left(\varphi(V, \mathbf{A})-\mathbb{E}\left[\varphi\left(\sqrt{\mathbb{E} X_{0}^{2}-q^{*}} U+\sqrt{q^{*}} V, \mathbf{A}\right) \mid V\right]\right)^{2}\right] \tag{7.12}
\end{align*}
$$

where $U, V \sim \mathcal{N}(0,1), \mathbf{A} \sim P_{A}$ are independent random variables and $q^{*}$ is such that $\mathbb{E} X_{0}^{2}-q^{*}=\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{|X|<v_{\left.k^{*}\right\}}\right\}}\right]$ is the asymptotic MMSE (7.9). For Bernoulli and Bernoulli-Rademacher signals (the ones considered in this section), it simplifies to

$$
\begin{align*}
& \lim _{n \rightarrow+\infty} \operatorname{MMSE}\left(Y_{\text {new }} \mid \mathbf{Y}, \mathbf{W}, \mathbf{W}_{\text {new }}\right) \\
& \quad= \begin{cases}\Delta+\mathbb{E}\left[(\varphi(V, \mathbf{A})-\mathbb{E}[\varphi(V, \mathbf{A}) \mid V])^{2}\right] & \text { if } \gamma>\gamma_{c} \\
\Delta+\operatorname{Var}(\varphi(V, \mathbf{A})) & \text { if } \gamma<\gamma_{c}\end{cases} \tag{7.13}
\end{align*}
$$

We thus find that the optimal generalization error also displays an all-or-nothing phase transition at $\gamma_{c}$. More precisely, if $\gamma<\gamma_{c}$ then the optimal generalization error equals $\Delta+\operatorname{Var}(\varphi(V, \mathbf{A}))$ when $n \rightarrow+\infty$. This is the same generalization error achieved by the dumb label estimator in the Bayesian sense; the one predicting the new label to be the output value averaged over all possible inputs, weights and noise. If instead $\gamma>\gamma_{c}$ then it is equal to $\Delta+\mathbb{E}[\operatorname{Var}(\varphi(V, \mathbf{A}) \mid V)]$; the irreducible error due to both the noise $\mathbf{Z}$ and the random stream $\left(\mathbf{A}_{\mu}\right)_{\mu=1}^{m_{n}}$.

Proving (7.12) entails introducing side observations in the original problem and differentiating with respect to the signal-to-noise ratio of this side channel to exploit the I-MMSE relationship, in a similar fashion to what we do in the proof of Theorem 7.3 (see Appendix 7.C). The side observations have the same form than the ones used in [29, Section 5 of SI Appendix] to determine the asymptotic optimal generalization error in the regime of linear sparsity and sampling rate.

Illustration of the all-or-nothing phenomenon In Figure 2 we use (7.10) to draw in solid black lines the asymptotic MMSE in the regime of sublinear sparsity and sampling rate, for both priors Bernoulli and Bernoulli-Rademacher and the activation functions $\varphi(x)=x, \varphi(x)=\operatorname{sign}(x), \varphi(x)=\max (0, x)$. For comparison we also draw in dashed colored lines the asymptotic MMSE in regimes of linear sparsity and sampling rate, that is, $\rho_{n}=\rho$ and $\alpha_{n}=\gamma \rho|\ln \rho|$ are constant with $n$. In this case, the asymptotic MMSE is given by [29, Theorem 2]

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \operatorname{MMSE}\left(\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right)=1-q^{*} \tag{7.14}
\end{equation*}
$$

whenever $\arg \min _{q \in[0,1]} \sup _{r \geq 0} i_{\mathrm{RS}}(q, r ; \gamma \rho|\ln \rho|, \rho)$ is a singleton $\left\{q^{*}\right\}$. In order to optimize the potential $i_{\mathrm{RS}}(q, r ; \gamma \rho|\ln \rho|, \rho)$, we initialize $q \in[0,1]$ at different values and iterate the fixed point equation (obtained from the stationary point equation $\left.\nabla i_{\mathrm{RS}}=0\right)$

$$
\begin{equation*}
r=-\left.2 \frac{\partial I_{P_{\text {out }}}}{\partial q}\right|_{q, 1}, \quad q=-\frac{2}{\rho_{n}} I_{P_{X}^{(n)}}^{\prime}\left(\frac{\alpha_{n}}{\rho_{n}} r\right) . \tag{7.15}
\end{equation*}
$$



Figure 2: Asymptotic MMSE as a function of $\gamma / \gamma_{c}$ in the regime of sublinear sparsity and sampling rate ( $\rho_{n}=\Theta\left(n^{-\lambda}\right)$ with $\lambda \in(0,1 / 9)$, solid black line), and in the regime of linear sparsity and sampling rate ( $\rho_{n}$ fixed, dashed colored lines). Dotted lines correspond to algorithmic performance in the regime of linear sparsity and sampling rate (iterating 7.15) from $q=10^{-10}$ ). Left panels: Bernoulli prior. Right panels: Bernoulli-Rademacher prior. From top to bottom: $\varphi(x)=x, \Delta=0.1 ; \varphi(x)=\operatorname{sign}(x), \Delta=0 ; \varphi(x)=\max (0, x), \Delta=0.5$.

Finally, the fixed point $q^{*}$ yielding the lowest potential $\sup _{r \geq 0} i_{\mathrm{RS}}\left(q^{*}, r ; \gamma \rho|\ln \rho|, \rho\right)$ is used to determine the MMSE thanks to (7.14). In all configurations the asymptotic MMSE jumps from a value close to 1 to approximately 0 as $\gamma$ increases past $\gamma_{c}$. As $\rho_{n}=\rho$ gets closer to 0 , this jump becomes sharper with the MMSE approaching 0 or 1 depending on which side of $\gamma_{c}$ we are. Though this jump becomes sharper, a pure all-or-nothing phase transition only occurs in the regime of sublinear sparsity and sampling rate (solid black lines).

In Figure 3 we use $(7.13$ ) to plot in solid black lines the asymptotic optimal generalization error for the Bernoulli prior and the same activation functions. The dashed colored lines again correspond to regimes of linear sparsity and sampling rate; they are obtained using the formula for the asymptotic optimal generalization error given by [29, Theorem 2],

$$
\begin{align*}
& \lim _{n \rightarrow+\infty} \operatorname{MMSE}\left(Y_{\text {new }} \mid \mathbf{Y}, \mathbf{W}, \mathbf{W}_{\text {new }}\right) \\
& \quad=\Delta+\mathbb{E}\left[\left(\varphi(V, \mathbf{A})-\mathbb{E}\left[\varphi\left(\sqrt{1-q^{*}} U+\sqrt{q^{*}} V, \mathbf{A}\right) \mid V\right]\right)^{2}\right] \tag{7.16}
\end{align*}
$$

In all configurations the optimal generalization error jumps from a value close to $\Delta+\operatorname{Var}(\varphi(V))$ to approximately $\Delta$ as $\gamma$ increases past $\gamma_{c}$ (note that the activations are deterministic so there is no contribution from $\mathbf{A}$ in the error). The value $\Delta$ is as good as the optimal generalization error can get, i.e., it is equal to the noise variance which is the squared error we would get if we were given the true weights $\mathbf{X}^{*}$. Again, the jump gets sharper as $\rho_{n}=\rho$ approaches 0 but a pure all-or-nothing phase transition only occurs in the regime of sublinear sparsity and sampling rate (solid black lines).


Figure 3: Asymptotic optimal generalization error as a function of $\gamma / \gamma_{c}$ in the regime of sublinear sparsity and sampling rate $\left(\rho_{n}=\Theta\left(n^{-\lambda}\right)\right.$ with $\lambda \in(0,1 / 9)$, solid black line $)$, and in the regime of linear sparsity and sampling rate ( $\rho_{n}$ is fixed, dashed colored lines). Dotted lines correspond to algorithmic performance in the regime of linear sparsity and sampling rate (iterating (7.15) from $q=10^{-10}$ ). Top left: random linear model $\varphi(x)=x, \Delta=0.1$. Top right: perceptron $\varphi(x)=\operatorname{sign}(x), \Delta=0$. Bottom: ReLU $\varphi(x)=\max (0, x), \Delta=0.5$.

The all-or-nothing behavior of the asymptotic MMSE and optimal generalization error is quite striking. Indeed, in the limit of vanishing sparsity and sampling rate either estimation or learning is as good as it can get or as bad as a random guess. This purely dichotomic behavior only occurs in the truly sparse limit, and is shown here to be pretty general in the sense that it occurs for a wide variety of activation functions. An important aspect of our results is to provide a definitive statistical benchmark allowing to measure the quality of algorithms with respect to the minimal amount of sparse data needed to estimate or learn. This benchmark is provided by non-trivial formulas (7.11) for the threshold $\gamma_{c}$ given for several examples in Table 1. We note that such precise benchmarks are quite rarely obtained in traditional machine learning approaches.

Further remarks Algorithmic aspects are beyond the scope of this paper. However, we make a few remarks about generalized approximate message passing (GAMP) algorithms. In the regime of linear sparsity and sampling rate, the state evolution equations precisely tracking the asymptotic performance of the algorithm are linked to the fixed point equation (7.15) [85]. The fixed point $q^{\text {alg }}$ reached by initializing (7.15) arbitrarily close to $q=0$ can be used in (7.14) and (7.16) - instead of $q^{*}$ - to obtain both the mean-square and generalization errors of GAMP algorithms. These errors are represented with dotted colored lines in Figures 2 and 3. We observe an algorithmic-to-statistical gap, that is, the dotted lines corresponding to the algorithmic performance do not drop to zero around $\gamma_{c}$ but at a higher algorithmic threshold. In this work we don't study the performance of GAMP algorithms in the regime of sublinear sparsity and sampling rate. However, reference 154 rigorously shows that in this regime the all-or-nothing behavior also occurs at an algorithmic level for GAMP algorithms. It would be highly desirable to extend their results to other activations and derive
the corresponding thresholds.

### 7.4 Proof of Theorem 7.2 for a Bernoulli prior

In this section, we assume that $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} \delta_{1}$, that is, each entry of the signal $\mathbf{X}^{*}$ is either 0 with probability $1-\rho_{n}$ or 1 with probability $\rho_{n}$. We prove Theorem 7.2 for this specific case. The proof contains all the main ideas needed to establish Theorem 7.2 while being technically simpler. The interested reader can find the proof of Theorem 7.2 for a general discrete prior with finite support in Appendix 7.B.

Let us introduce a few notations. For fixed $\rho_{n}, \alpha_{n}>0$, we denote the variational problem appearing in Theorem 7.1 by

$$
I\left(\rho_{n}, \alpha_{n}\right):=\inf _{q \in[0,1]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right),
$$

where the potential $i_{\mathrm{RS}}$ is defined in (7.7). Let $X^{*} \sim P_{X}^{(n)}$ and $Z \sim \mathcal{N}(0,1)$ be independent random variables. We define for all $r \geq 0$ :

$$
\begin{equation*}
\psi_{P_{X}^{(n)}}(r):=\mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} e^{-\frac{r}{2}+r X^{*}+\sqrt{r} Z}\right)\right] . \tag{7.17}
\end{equation*}
$$

Note that $I_{P_{X}^{(n)}}(r):=I\left(X^{*} ; \sqrt{r} X^{*}+Z\right)=\frac{r \rho_{n}}{2}-\psi_{P_{X}^{(n)}}(r)$ so

$$
\begin{equation*}
I\left(\rho_{n}, \alpha_{n}\right)=\inf _{q \in[0,1]} I_{P_{\text {out }}}(q, 1)+\sup _{r \geq 0}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\} . \tag{7.18}
\end{equation*}
$$

The latter expression for $I\left(\rho_{n}, \alpha_{n}\right)$ is easier to work with. We point out that $\psi_{P_{X}^{(n)}}$ is twice differentiable, nondecreasing, strictly convex and ( $\left.\rho_{n} / 2\right)$-Lipschitz continuous on $[0,+\infty)$ (see Lemma 2.3) while $I_{P_{\text {out }}}(\cdot, 1)$ is nonincreasing and concave on $[0,1]$ (see [29, Appendix B.2, Proposition 18]).

Our goal is now to compute the limit of $I\left(\rho_{n}, \alpha_{n}\right)$ when $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$ and $\rho_{n} \rightarrow 0$. Once we know this limit, we directly obtain Theorem 7.2 thanks to Theorem 7.1. We first show that, for all $q$ in a growing interval, the point at which the supremum over $r$ is achieved is located in an interval that shrinks on $r^{*}:=2 / \gamma$.

Lemma 7.4. Let $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} \delta_{1}$ and $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$. Define $g_{\rho_{n}}: r \in(0,+\infty) \mapsto \frac{2}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)$ and $\forall \rho_{n} \in\left(0, e^{-1}\right)$ :

$$
\begin{equation*}
a_{\rho_{n}}:=g_{\rho_{n}}\left(\frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}\right), \quad b_{\rho_{n}}:=g_{\rho_{n}}\left(\frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}\right) . \tag{7.19}
\end{equation*}
$$

Then, $\left[a_{\rho_{n}}, b_{\rho_{n}}\right] \subset\left(\rho_{n}, 1\right)$ and $\lim _{\rho_{n} \rightarrow 0} a_{\rho_{n}}=0, \lim _{\rho_{n} \rightarrow 0} b_{\rho_{n}}=1$. Besides, for every $q \in\left(\rho_{n}, 1\right)$ there exists a unique $r_{n}^{*}(q) \in(0,+\infty)$ such that

$$
\begin{equation*}
\frac{r_{n}^{*}(q) q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}(q)\right)=\sup _{r \geq 0} \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right), \tag{7.20}
\end{equation*}
$$

and

$$
\begin{align*}
& \forall q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]: \frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma} \leq r_{n}^{*}(q) \leq \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}  \tag{7.21}\\
& \forall q \in\left[b_{\rho_{n}}, 1\right): r_{n}^{*}(q) \geq \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma} \tag{7.22}
\end{align*}
$$

Proof. For every $q \in(0,1)$ we define $f_{\rho_{n}, q}: r \in[0,+\infty) \mapsto \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)$ whose supremum over $r$ we want to compute. The derivative of $f_{\rho_{n}, q}$ with respect to $r$ reads

$$
f_{\rho_{n}, q}^{\prime}(r)=\frac{q}{2}-\frac{1}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\frac{\alpha_{n}}{\rho_{n}} r\right) .
$$

The derivative $\psi_{P_{X}^{(n)}}^{\prime}$ is continuously increasing and thus one-to-one from $(0,+\infty)$ onto

$$
\left(\psi_{P_{X}^{(n)}}^{\prime}(0), \lim _{r \rightarrow+\infty} \psi_{P_{X}^{(n)}}^{\prime}(r)\right)=\left(\frac{\rho_{n}^{2}}{2}, \frac{\rho_{n}}{2}\right)
$$

Therefore, if $q \in\left(0, \rho_{n}\right], f_{\rho_{n}, q}^{\prime}$ is negative on the whole interval $(0,+\infty)$ and the supremum of $f_{\rho_{n}, q}$ is achieved at $r=0$. On the contrary, if $q \in\left(\rho_{n}, 1\right)$, there exists a unique solution $r_{n}^{*}(q) \in(0,+\infty)$ to the stationary point equation $f_{\rho_{n}, q}^{\prime}(r)=0$. As $f_{\rho_{n}, q}$ is concave (given that $\psi_{P_{0}, n}$ is convex), this solution $r_{n}^{*}(q)$ is the global maximum of $f_{\rho_{n}, q}$. We now transform the critical point equation,

$$
f_{\rho_{n}, q}(r)=0 \Leftrightarrow \frac{2}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)=q \Leftrightarrow g_{\rho_{n}}(r)=q
$$

where $g_{\rho_{n}}: r \mapsto \frac{2}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)$ is increasing and one-to-one from $(0,+\infty)$ to $\left(\rho_{n}, 1\right)$. For all $\rho_{n} \in\left(0, e^{-1}\right),\left|\ln \rho_{n}\right|^{-\frac{1}{4}} \in(0,1)$. By Lemma 7.5 (following the proof) applied to $\epsilon=\left|\ln \rho_{n}\right|^{-\frac{1}{4}}$, we have

$$
\begin{aligned}
& \rho_{n}<a_{\rho_{n}}:=g_{\rho_{n}}\left(\frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}\right) \leq \frac{\exp \left(-\frac{\left|\ln \rho_{n}\right|^{\frac{1}{2}}}{16\left(1-\left|\ln \rho_{n}\right|^{1 / 4}\right)}\right)}{2}+\frac{\exp \left(-\frac{\left|\ln \rho_{n}\right|^{\frac{3}{4}}}{2}\right)}{1-\rho_{n}} ; \\
& 1>b_{\rho_{n}}:=g_{\rho_{n}}\left(\frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}\right) \geq \frac{1-0.5 \exp \left(-\frac{\left|\ln \rho_{n}\right|^{1 / 2}}{16}\right)}{1+\exp \left(-\frac{\left|\ln \rho_{n}\right|^{3 / 4}}{2}\right)} .
\end{aligned}
$$

It directly follows from these two bounds that $\lim _{\rho_{n} \rightarrow 0} a_{\rho_{n}}=0$ and $\lim _{\rho_{n} \rightarrow 0} b_{\rho_{n}}=1$. We know that, for all $q \in\left(\rho_{n}, 1\right), g_{\rho_{n}}\left(r_{n}^{*}(q)\right)=q$. As $g_{\rho_{n}}$ is increasing, if $q=g_{\rho_{n}}\left(r_{n}^{*}(q)\right) \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]$ then

$$
\frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma} \leq r_{n}^{*}(q) \leq \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}
$$

while if $q=g_{\rho_{n}}\left(r_{n}^{*}(q)\right) \in\left[b_{\rho_{n}}, 1\right)$ then $r_{n}^{*}(q) \geq \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}$.

Lemma 7.5. Let $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} \delta_{1}$ and $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$. Define $g_{\rho_{n}}: r \in(0,+\infty) \mapsto \frac{2}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)$ For all $\left(\rho_{n}, \epsilon\right) \in(0,1)^{2}$ :

$$
\begin{aligned}
& g_{\rho_{n}}\left(\frac{2(1-\epsilon)}{\gamma}\right) \leq \frac{\exp \left(-\frac{\epsilon^{2}}{16} \frac{\left|\ln \rho_{n}\right|}{1-\epsilon}\right)}{2}+\frac{\exp \left(-\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)}{1-\rho_{n}} ; \\
& g_{\rho_{n}}\left(\frac{2(1+\epsilon)}{\gamma}\right) \geq \frac{1-0.5 \exp \left(-\frac{\epsilon^{2}}{16}\left|\ln \rho_{n}\right|\right)}{1+\exp \left(-\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)} .
\end{aligned}
$$

Proof. The derivative of $\psi_{P_{X}^{(n)}}$ reads

$$
\psi_{P_{X}^{(n)}}^{\prime}(r)=\frac{\rho_{n}}{2} \mathbb{E}\left[\left(1+\frac{1-\rho_{n}}{\rho_{n}} e^{-\frac{r}{2}-\sqrt{r} Z}\right)^{-1}\right]
$$

so

$$
g_{\rho_{n}}(r)=\mathbb{E}\left[\frac{1}{1+\left(1-\rho_{n}\right) \exp \left\{\left|\ln \rho_{n}\right|\left(1-\frac{\gamma r}{2}-\sqrt{\frac{\gamma r}{\left|\ln \rho_{n}\right|}} Z\right)\right\}}\right] \in(0,1) .
$$

Hence, for all $\epsilon \in(0,1)$ :

$$
\begin{equation*}
g_{\rho_{n}}\left(\frac{2(1 \pm \epsilon)}{\gamma}\right)=\mathbb{E}\left[\frac{1}{1+\left(1-\rho_{n}\right) \exp \left\{\left|\ln \rho_{n}\right|\left(\mp \epsilon-\sqrt{\frac{2(1 \pm \epsilon)}{\left|\ln \rho_{n}\right|}} Z\right)\right\}}\right] . \tag{7.23}
\end{equation*}
$$

We see directly that, by the dominated convergence theorem,

$$
\lim _{\rho_{n} \rightarrow 0} g_{\rho_{n}}(2(1+\epsilon) / \gamma)=1 \text { and } \lim _{\rho_{n} \rightarrow 0} g_{\rho_{n}}(2(1-\epsilon) / \gamma)=0 .
$$

We first lower bound $g_{\rho_{n}}(2(1+\epsilon) / \gamma)$. Using (7.23) and $-\epsilon-\sqrt{2(1+\epsilon) /\left|\ln \rho_{n}\right|} z \leq-\epsilon / 2$ for all $z \geq-\frac{\epsilon}{2} \sqrt{\frac{\left|\ln \rho_{n}\right|}{2(1+\epsilon)}}$, it comes

$$
\begin{aligned}
g_{\rho_{n}}\left(\frac{2(1+\epsilon)}{\gamma}\right) & =\int_{-\infty}^{+\infty} \frac{d z}{\sqrt{2 \pi}} \frac{e^{-\frac{z^{2}}{2}}}{1+\left(1-\rho_{n}\right) \exp \left\{\left|\ln \rho_{n}\right|\left(-\epsilon-\sqrt{\frac{2(1+\epsilon)}{\left|\ln \rho_{n}\right|}} z\right)\right\}} \\
& \geq \int_{-\frac{\epsilon}{2} \sqrt{\frac{\left|\ln \rho_{n}\right|}{2(1+\epsilon)}} \frac{d z}{\sqrt{2 \pi}} \frac{e^{-\frac{z^{2}}{2}}}{1+\left(1-\rho_{n}\right) \exp \left(-\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)}} \\
& =\frac{1-F\left(-\frac{\epsilon}{2} \sqrt{\frac{\left|\ln \rho_{n}\right|}{2(1+\epsilon)}}\right)}{1+\left(1-\rho_{n}\right) \exp \left(-\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)} \geq \frac{1-F\left(-\frac{\epsilon}{2} \sqrt{\frac{\left|\ln \rho_{n}\right|}{2}}\right)}{1+\exp \left(-\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)}
\end{aligned}
$$

where $F(x):=\int_{-\infty}^{x} \frac{d z}{\sqrt{2 \pi}} e^{-\frac{z^{2}}{2}}$ is the cumulative distribution function of the standard normal distribution. Making use of the upper bound $F(-x) \leq \frac{e^{-x^{2} / 2}}{2}$ for $x>0$ yields

$$
\begin{equation*}
g_{\rho_{n}}\left(\frac{2(1+\epsilon)}{\gamma}\right) \geq \frac{1-0.5 \exp \left(-\frac{\epsilon^{2}}{16}\left|\ln \rho_{n}\right|\right)}{1+\exp \left(-\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)} . \tag{7.24}
\end{equation*}
$$

Next we prove the upper bound on $g_{\rho_{n}}(2(1-\epsilon) / \gamma)$. We denote the indicator function of an event $\mathcal{E}$ by $\mathbf{1}_{\mathcal{E}}$. We have

$$
\begin{align*}
g_{\rho_{n}}\left(\frac{2(1-\epsilon)}{\gamma}\right) & =\mathbb{E}\left[\frac{1}{1+\left(1-\rho_{n}\right) \exp \left\{\left|\ln \rho_{n}\right|\left(\epsilon-\sqrt{\frac{2(1-\epsilon)}{\mid \ln \rho_{n}}} Z\right)\right\}}\right] \\
& \leq \mathbb{E}\left[\boldsymbol{1}_{\left\{Z \geq \frac{\epsilon}{2} \sqrt{\frac{\mid \ln \rho_{n}}{2(1-\epsilon)}}\right\}}+\frac{\left.\mathbf{1}_{\left\{Z<\frac{\epsilon}{2}\right.} \sqrt{\frac{\mid \ln \rho_{n}}{2(1-\epsilon \epsilon}}\right\}}{1+\left(1-\rho_{n}\right) \exp \left(\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)}\right] \\
& =F\left(-\frac{\epsilon}{2} \sqrt{\frac{\left|\ln \rho_{n}\right|}{2(1-\epsilon)}}\right)+\frac{1-F\left(-\frac{\epsilon}{2} \sqrt{\frac{\ln \rho_{n} \mid}{2(1-\epsilon)}}\right)}{1+\left(1-\rho_{n}\right) \exp \left(\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)} \\
& \leq F\left(-\frac{\epsilon}{2} \sqrt{\frac{\left|\ln \rho_{n}\right|}{2(1-\epsilon)}}\right)+\frac{\exp \left(-\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)}{1-\rho_{n}} \\
& \leq \frac{\exp \left(-\frac{\epsilon^{2}}{16} \frac{\left|\ln \rho_{n}\right|}{1-\epsilon}\right)}{2}+\frac{\exp \left(-\frac{\epsilon}{2}\left|\ln \rho_{n}\right|\right)}{1-\rho_{n}} . \tag{7.25}
\end{align*}
$$

The last inequality follows from the upper bound on $F(-x)$ that we have already used to obtain (7.24).

Lemma 7.4 essentially states that the global maximum of $r \mapsto \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)$ is located in a tight interval around $2 / \gamma$ when $q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]$. The next step is to use this knowledge to tightly bound the maximum value $\sup _{r \geq 0} \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)$ for all $q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]$. The following lemma gives a useful bound on $\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)$ when $0 \leq r \leq{ }^{2(1+\epsilon)} / \gamma$.

Lemma 7.6. Let $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} \delta_{1}$ and $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$. For every $\epsilon \in(0,1)$ and $r \in[0,2(1+\epsilon) / \gamma]$ we have

$$
\begin{equation*}
0 \leq \frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right) \leq \frac{\epsilon}{\gamma}+\frac{\ln 2}{\gamma\left|\ln \rho_{n}\right|}+\frac{1}{\gamma} \sqrt{\frac{2}{\pi\left|\ln \rho_{n}\right|}} \tag{7.26}
\end{equation*}
$$

Proof. The function $\psi_{P_{X}^{(n)}}$ is nondecreasing on $[0,+\infty)$ so $\forall r \in[0,2(1+\epsilon) / \gamma]$ :

$$
\begin{equation*}
0 \leq \frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right) \leq \frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2(1+\epsilon)}{\gamma}\right)=\frac{\psi_{P_{X}^{(n)}}\left(2(1+\epsilon)\left|\ln \rho_{n}\right|\right)}{\gamma \rho_{n}\left|\ln \rho_{n}\right|} \tag{7.27}
\end{equation*}
$$

By definition 7.17 of $\psi_{P_{X}^{(n)}}$, the upper bound on the right-hand side of (7.27) reads

$$
\begin{align*}
\frac{\psi_{P_{X}^{(n)}}\left(2(1+\epsilon)\left|\ln \rho_{n}\right|\right)}{\gamma \rho_{n}\left|\ln \rho_{n}\right|}= & \frac{1-\rho_{n}}{\gamma \rho_{n}\left|\ln \rho_{n}\right|} \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} e^{-(1+\epsilon)\left|\ln \rho_{n}\right|+\sqrt{2(1+\epsilon)\left|\ln \rho_{n}\right|} Z}\right)\right] \\
& +\frac{1}{\gamma\left|\ln \rho_{n}\right|} \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} e^{(1+\epsilon)\left|\ln \rho_{n}\right|+\sqrt{2(1+\epsilon)\left|\ln \rho_{n}\right|} Z}\right)\right] \\
= & \frac{1-\rho_{n}}{\gamma \rho_{n}\left|\ln \rho_{n}\right|} \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} e^{-(1+\epsilon)\left|\ln \rho_{n}\right|+\sqrt{2(1+\epsilon)\left|\ln \rho_{n}\right|} Z}\right)\right] \\
+ & \frac{1}{\gamma\left|\ln \rho_{n}\right|} \mathbb{E}\left[\ln \left(1-\rho_{n}+e^{\epsilon\left|\ln \rho_{n}\right|+\sqrt{2(1+\epsilon)\left|\ln \rho_{n}\right|} Z}\right)\right] . \tag{7.28}
\end{align*}
$$

To control the first term on the right-hand side of 7.28 we use that $\ln (1+x) \leq x$,

$$
\begin{align*}
& \frac{1-\rho_{n}}{\gamma \rho_{n}\left|\ln \rho_{n}\right|} \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} e^{-(1+\epsilon)\left|\ln \rho_{n}\right|+\sqrt{2(1+\epsilon)\left|\ln \rho_{n}\right|} Z}\right)\right] \\
& \leq \frac{1-\rho_{n}}{\gamma\left|\ln \rho_{n}\right|} \mathbb{E}\left[e^{-(1+\epsilon)\left|\ln \rho_{n}\right|+\sqrt{2(1+\epsilon)\left|\ln \rho_{n}\right|} Z}-1\right]=0 . \tag{7.29}
\end{align*}
$$

To control the second term on the right-hand side of $(7.28)$, we use
$\ln \left(1-\rho_{n}+e^{\epsilon\left|\ln \rho_{n}\right|+\sqrt{2(1+\epsilon)\left|\ln \rho_{n}\right| z}}\right) \leq\left\{\begin{array}{cl}\ln \left(1+e^{\epsilon\left|\ln \rho_{n}\right|}\right) \leq \ln \left(2 e^{\epsilon\left|\ln \rho_{n}\right|}\right) & \text { if } z \leq 0 ; \\ \ln \left(2 e^{\epsilon\left|\ln \rho_{n}\right|+\sqrt{2(1+\epsilon)\left|\ln \rho_{n}\right| z}}\right) & \text { if } z \geq 0 .\end{array}\right.$
It directly follows that

$$
\frac{1}{\gamma\left|\ln \rho_{n}\right|} \mathbb{E}\left[\ln \left(1-\rho_{n}+e^{\epsilon\left|\ln \rho_{n}\right|+\sqrt{2(1+\epsilon)\left|\ln \rho_{n}\right|} Z}\right)\right] \leq \frac{\epsilon}{\gamma}+\frac{\ln 2}{\gamma\left|\ln \rho_{n}\right|}+\frac{1}{\gamma} \sqrt{\frac{1+\epsilon}{\pi\left|\ln \rho_{n}\right|}}
$$

The latter combined with $(7.28$ and $(7.29)$ ends the proof.
We can now compute the limit of $I\left(\rho_{n}, \alpha_{n}\right)$ when $\rho_{n} \rightarrow 0$ and $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$.

Proposition 7.7. Let $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} \delta_{1}$ and $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$. Then, the quantity $I\left(\rho_{n}, \alpha_{n}\right):=\inf _{q \in[0,1]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$ converges when $\rho_{n} \rightarrow 0^{+}$and

$$
\lim _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}\right)=\min \left\{I_{P_{\text {out }}}(0,1), \frac{1}{\gamma}\right\} .
$$

Proof. Let $a_{\rho_{n}}$ and $b_{\rho_{n}}$ the quantities defined in Lemma 7.4 . By Lemmas 7.4 and 7.6 (applied to $\epsilon=\left|\ln \rho_{n}\right|^{-\frac{1}{4}}$ for $\rho_{n}$ small enough), we have $\forall q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]$ :

$$
\begin{align*}
\frac{\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right) q}{\gamma}- & \frac{1}{\gamma}\left(\frac{1}{\left|\ln \rho_{n}\right|^{\frac{1}{4}}}+\frac{\ln 2}{\left|\ln \rho_{n}\right|}+\sqrt{\frac{2}{\pi\left|\ln \rho_{n}\right|}}\right) \\
& \leq \frac{r_{n}^{*}(q) q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}(q)\right) \leq \frac{\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right) q}{\gamma} . \tag{7.30}
\end{align*}
$$

Therefore, $\forall q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]$ :

$$
\begin{aligned}
& I_{P_{\text {out }}}(q, 1)+\frac{q}{\gamma}-\frac{1}{\gamma}\left(\frac{2}{\left|\ln \rho_{n}\right|^{\frac{1}{4}}}+\frac{\ln 2}{\left|\ln \rho_{n}\right|}+\sqrt{\frac{2}{\pi\left|\ln \rho_{n}\right|}}\right) \\
& \quad \leq \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \leq I_{P_{\text {out }}}(q, 1)+\frac{q}{\gamma}+\frac{1}{\gamma\left|\ln \rho_{n}\right|^{\frac{1}{4}}} .
\end{aligned}
$$

It directly follows that

$$
\begin{align*}
& -\frac{1}{\gamma}\left(\frac{2}{\left|\ln \rho_{n}\right|^{\frac{1}{4}}}+\frac{\ln 2}{\left|\ln \rho_{n}\right|}+\sqrt{\frac{2}{\pi\left|\ln \rho_{n}\right|}}\right)+\inf _{q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]} I_{P_{\text {out }}}(q, 1)+\frac{q}{\gamma} \\
& \quad \leq \inf _{q \in\left[a_{\rho_{n}}, b_{\left.\rho_{n}\right]}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \leq \frac{1}{\gamma\left|\ln \rho_{n}\right|^{\frac{1}{4}}}+\inf _{q \in\left[\rho_{\rho_{n}}, b_{\rho_{n}}\right]} I_{P_{\text {out }}}(q, 1)+\frac{q}{\gamma} . \tag{7.31}
\end{align*}
$$

Note that $q \mapsto I_{P_{\text {out }}}(q, 1)+\frac{q}{\gamma}$ is concave on $[0,1]$ so

$$
\begin{align*}
\inf _{q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]} I_{P_{\text {out }}}(q, 1)+\frac{q}{\gamma} & =\min \left\{I_{P_{\text {out }}}\left(a_{\rho_{n}}, 1\right)+\frac{a_{\rho_{n}}}{\gamma}, I_{P_{\text {out }}}\left(b_{\rho_{n}}, 1\right)+\frac{b_{\rho_{n}}}{\gamma}\right\} \\
& \xrightarrow[\rho_{n} \rightarrow 0]{ } \min \left\{I_{P_{\text {out }}}(0,1), \frac{1}{\gamma}\right\} ; \tag{7.32}
\end{align*}
$$

where the limit is due to Lemma 7.4. Combining the bounds (7.31) with the limit (7.32) yields

$$
\begin{equation*}
\lim _{\rho_{n} \rightarrow 0} \inf _{q \in\left[\rho_{\rho_{n}}, b_{\rho_{n}}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)=\min \left\{I_{P_{\text {out }}}(0,1), \frac{1}{\gamma}\right\} . \tag{7.33}
\end{equation*}
$$

Upper bound on the limit superior of $I\left(\rho_{n}, \alpha_{n}\right)$ The upper bound on the limit superior of $I\left(\rho_{n}, \alpha_{n}\right):=\inf _{q \in[0,1]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$ follows from the simple upper bound $I\left(\rho_{n}, \alpha_{n}\right) \leq \inf _{q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]} \leq \operatorname{Sup}_{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$ and the limit (7.33),

$$
\begin{equation*}
\limsup _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}\right) \leq \min \left\{I_{P_{\text {out }}}(0,1), \frac{1}{\gamma}\right\} . \tag{7.34}
\end{equation*}
$$

Matching lower bound on the limit inferior of $I\left(\rho_{n}, \alpha_{n}\right)$ We first rewrite $I\left(\rho_{n}, \alpha_{n}\right)$ by splitting the segment $[0,1]=\left[0, a_{\rho_{n}}\right] \cup\left[a_{\rho_{n}}, b_{\rho_{n}}\right] \cup\left[b_{\rho_{n}}, 1\right]$ :

$$
\begin{align*}
I\left(\rho_{n}, \alpha_{n}\right)=\min \{ & \inf _{q \in\left[0, a_{\rho_{n}}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right), \inf _{q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]} \sup i_{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right), \\
& \left.\inf _{q \in\left[b_{\rho_{n}}, 1\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)\right\} . \tag{7.35}
\end{align*}
$$

For all $q \in\left[0, a_{\rho_{n}}\right]$, we have

$$
\begin{aligned}
\sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) & =I_{P_{\text {out }}}(q, 1)+\sup _{r \geq 0}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\} \\
& \geq I_{P_{\text {out }}}(q, 1)+\lim _{r \rightarrow 0^{+}}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\}=I_{P_{\text {out }}}(q, 1) .
\end{aligned}
$$

The function $q \mapsto I_{P_{\text {out }}}(q, 1)$ is decreasing and it follows that

$$
\begin{equation*}
\inf _{q \in\left[0, a_{\rho n}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \geq \inf _{q \in\left[0, a_{\rho}\right]} I_{P_{\text {out }}}(q, 1)=I_{P_{\text {out }}}\left(a_{\rho_{n}}, 1\right) . \tag{7.36}
\end{equation*}
$$

For all $q \in\left[b_{\rho_{n}}, 1\right)$, we have

$$
\begin{align*}
\sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) & =I_{P_{\text {out }}}(q, 1)+\sup _{r \geq 0}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\} \\
& \geq \frac{q\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}\right) \\
& \geq \frac{b_{\rho_{n}}}{\gamma}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}\right) \\
& \geq \frac{b_{\rho_{n}}}{\gamma}-\frac{1}{\gamma}\left(\frac{1}{\left|\ln \rho_{n}\right|^{\frac{1}{4}}}+\frac{\ln 2}{\left|\ln \rho_{n}\right|}+\sqrt{\frac{2}{\pi\left|\ln \rho_{n}\right|}}\right) . \tag{7.37}
\end{align*}
$$

The first inequality follows from the trivial lower bounds $I_{P_{\text {out }}}(q, 1) \geq 0$ and

$$
\sup _{r \geq 0} \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right) \geq \frac{\widetilde{r} q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \widetilde{r}\right) \quad \text { where } \quad \widetilde{r}:=\frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma} .
$$

To obtain the last inequality we apply Lemma 7.6 with $\epsilon=\left|\ln \rho_{n}\right|^{-\frac{1}{4}}$,

$$
\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma}\right) \leq \frac{1}{\gamma}\left(\frac{1}{\left|\ln \rho_{n}\right|^{\frac{1}{4}}}+\frac{\ln 2}{\left|\ln \rho_{n}\right|}+\sqrt{\frac{2}{\pi\left|\ln \rho_{n}\right|}}\right) .
$$

Note that the final lower bound (7.37) does not depend on $q \in\left[b_{\rho_{n}}, 1\right)$ so the same inequality holds for the infimum of $\sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$ over $q \in\left[b_{\rho_{n}}, 1\right]$. Combining (7.35), (7.36) and (7.37) yields

$$
\begin{aligned}
& I\left(\rho_{n}, \alpha_{n}\right) \geq \min \left\{I_{P_{\text {out }}}\left(a_{\rho_{n}}, 1\right), \inf _{q \in\left[a_{\rho_{n}}, b_{\rho_{n}}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right),\right. \\
& \\
& \left.\quad \frac{b_{\rho_{n}}}{\gamma}-\frac{1}{\gamma}\left(\frac{1}{\left|\ln \rho_{n}\right|^{\frac{1}{4}}}+\frac{\ln 2}{\left|\ln \rho_{n}\right|}+\sqrt{\frac{2}{\pi\left|\ln \rho_{n}\right|}}\right)\right\} .
\end{aligned}
$$

Hence (remember the limit (7.33) and that $a_{\rho_{n}} \rightarrow 0, b_{\rho_{n}} \rightarrow 1$ when $\rho_{n}$ vanishes),

$$
\begin{align*}
\liminf _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}\right) & \geq \min \left\{I_{P_{\text {out }}}(0,1) ; \min \left\{I_{P_{\text {out }}}(0,1) ; \frac{1}{\gamma}\right\} ; \frac{1}{\gamma}\right\} \\
& =\min \left\{I_{P_{\text {out }}}(0,1) ; \frac{1}{\gamma}\right\} . \tag{7.38}
\end{align*}
$$

We see thanks to (7.34) and (7.38) that the superior and inferior limits of $I\left(\rho_{n}, \alpha_{n}\right)$ match each other so $\lim _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}\right)=\min \left\{I_{P_{\text {out }}}(0,1), \frac{1}{\gamma}\right\}$.

We finally obtain Theorem 7.2 for the specific choice $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} \delta_{1}$ by combining Theorem 7.1 and Proposition 7.7 together,

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} \frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}=\min \left\{I_{P_{\text {out }}}(0,1) ; \frac{1}{\gamma}\right\} . \tag{7.39}
\end{equation*}
$$

## Appendices

## 7.A Proof of Theorem 7.1 with the adaptive interpolation method

Note that it is the same to observe (7.2) or their rescaled versions

$$
\frac{Y_{\mu}}{\sqrt{\Delta}} \frac{1}{\sqrt{\Delta}} \varphi\left(\frac{1}{\sqrt{k_{n}}}\left(\mathbf{W} \mathbf{X}^{*}\right)_{\mu}, \mathbf{A}_{\mu}\right)+Z_{\mu}
$$

Therefore, up to a rescaling of $\varphi$ by $1 / \sqrt{\Delta}$, we suppose $\Delta=1$ all along the proof of Theorem 7.1. For similar reason, we suppose that the second moment of $X_{0} \sim P_{0}$ is $\mathbb{E}\left[X_{0}^{2}\right]=1$.

## 7.A. 1 Interpolating estimation problem

Let $\left(s_{n}\right)_{n \in \mathbb{N}^{*}}$ be a sequence of real numbers in $(0,1 / 2]$ and define $\mathcal{B}_{n}:=\left[s_{n}, 2 s_{n}\right]^{2}$. We denote $r_{\text {max }}:=-\left.2\left(\partial I_{P_{\text {out }}} / \partial q\right)\right|_{q=1, \rho=1} ; r_{\text {max }}$ is a positive real number. For all $\epsilon:=\left(\epsilon_{1}, \epsilon_{2}\right) \in \mathcal{B}_{n}$, we define the interpolation functions

$$
R_{1}(\cdot, \epsilon): t \in[0,1] \mapsto \epsilon_{1}+\int_{0}^{t} r_{\epsilon}(v) d v
$$

and

$$
R_{2}(\cdot, \epsilon): t \in[0,1] \mapsto \epsilon_{2}+\int_{0}^{t} q_{\epsilon}(v) d v
$$

where $q_{\epsilon}:[0,1] \rightarrow[0,1]$ and $r_{\epsilon}:[0,1] \rightarrow\left[0, \frac{\alpha_{n}}{\rho_{n}} r_{\max }\right]$ are two continuous functions. We specify $q_{\epsilon}$ and $r_{\epsilon}$ more explicitly later in the proof. In particular, we will need the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ to satisfy the following notion of regularity.

Definition (Regular interpolation paths). We say that the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ are regular if, for all $t \in[0,1]$, the function

$$
\epsilon \mapsto\left(R_{1}(t, \epsilon), R_{2}(t, \epsilon)\right)
$$

is a $C^{1}$-diffeomorphism from $\mathcal{B}_{n}$ onto its image whose Jacobian determinant is greater than, or equal, to one.

Let $\mathbf{X}^{*}$ be a $n$-dimensional random vector with entries $X_{i}^{*} \stackrel{\text { i.i.d. }}{\sim} P_{X}^{(n)}, \mathbf{U}, \mathbf{V}$ two $m_{n}$-dimensional random vectors with entries $U_{\mu}, V_{\mu} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$, and $\mathbf{W}$ a $m_{n} \times n$ random matrix with entries $W_{\mu i} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. We denote by $\mathbf{S}^{(t, \epsilon)}$ the $m_{n}$-dimensional random vector whose entries are given for all $\mu \in\left\{1, \ldots, m_{n}\right\}$ by

$$
\begin{equation*}
S_{\mu}^{(t, \epsilon)}:=\sqrt{\frac{1-t}{k_{n}}}\left(\mathbf{W X}^{*}\right)_{\mu}+\sqrt{R_{2}(t, \epsilon)} V_{\mu}+\sqrt{t+2 s_{n}-R_{2}(t, \epsilon)} U_{\mu} \tag{7.40}
\end{equation*}
$$

Consider the following observations coming from two types of channels,

$$
\begin{cases}Y_{\mu}^{(t, \epsilon)} \sim P_{\text {out }}\left(\cdot \mid S_{\mu}^{(t, \epsilon)}\right) & , 1 \leq \mu \leq m_{n}  \tag{7.41}\\ \widetilde{Y}_{i}^{(t, \epsilon)}=\sqrt{R_{1}(t, \epsilon)} X_{i}^{*}+\widetilde{Z}_{i}, & 1 \leq i \leq n\end{cases}
$$

where $\left(\widetilde{Z}_{i}\right)_{i=1}^{n} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. For a fixed $t \in[0,1]$, the inference problem is to estimate ( $\mathbf{X}^{*}, \mathbf{U}$ ) from the knowledge of $\mathbf{V}, \mathbf{W}$ and the observations

$$
\mathbf{Y}^{(t, \epsilon)}:=\left(Y_{\mu}^{(t, \epsilon)}\right)_{\mu=1}^{m_{n}}, \quad \tilde{\mathbf{Y}}^{(t, \epsilon)}:=\left(\widetilde{Y}_{i}^{(t, \epsilon)}\right)_{i=1}^{n}
$$

The joint posterior density of $\left(\mathbf{X}^{*}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)$ is

$$
\begin{align*}
d P\left(\mathbf{x}, \mathbf{u} \mid \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right):=\frac{1}{\mathcal{Z}_{t, \epsilon}} & \left.\prod_{i=1}^{n} d P_{X}^{(n)}\left(x_{i}\right) e^{-\frac{1}{2}\left(\sqrt{R_{1}(t, \epsilon)} x_{i}-\widetilde{Y}_{i}^{(t, \epsilon)}\right.}\right)^{2} \\
& \cdot \prod_{\mu=1}^{m_{n}} \frac{d u_{\mu}}{\sqrt{2 \pi}} e^{-\frac{u_{\mu}^{2}}{2}} P_{\mathrm{out}}\left(Y_{\mu}^{(t, \epsilon)} \mid s_{\mu}^{(t, \epsilon)}\right) \tag{7.42}
\end{align*}
$$

where

$$
\begin{equation*}
s_{\mu}^{(t, \epsilon)}:=\sqrt{\frac{1-t}{k_{n}}}(\mathbf{W} \mathbf{x})_{\mu}+\sqrt{R_{2}(t, \epsilon)} V_{\mu}+\sqrt{t+2 s_{n}-R_{2}(t, \epsilon)} u_{\mu} \tag{7.43}
\end{equation*}
$$

and $\mathcal{Z}_{t, \epsilon} \equiv \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)$ is the normalization factor. The interpolating mutual information is the conditional mutual information between $\left(\mathbf{X}^{*}, \mathbf{U}\right)$ and $\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right)$ given $(\mathbf{W}, \mathbf{V})$ and is denoted by

$$
\begin{equation*}
i_{n, \epsilon}(t):=\frac{1}{m_{n}} I\left(\left(\mathbf{X}^{*}, \mathbf{U}\right) ;\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right) \mid \mathbf{W}, \mathbf{V}\right) . \tag{7.44}
\end{equation*}
$$

The perturbation $\epsilon$ only induces a small change in mutual information. In particular, we have the following result at $t=0$.

Lemma 7.8. Suppose that (H1), (H2), (H3) hold, $\Delta=\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$, and there exist real positive numbers $M_{\alpha}, M_{\rho / \alpha}$ such that $\forall n \in \mathbb{N}^{*}: \alpha_{n} \leq M_{\alpha}$ and $\rho_{n} / \alpha_{n} \leq M_{\rho / \alpha}$. Then, for all $\epsilon \in \mathcal{B}_{n}$ :

$$
\left|i_{n, \epsilon}(0)-\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}\right| \leq \sqrt{C} \frac{s_{n}}{\sqrt{\rho_{n}}},
$$

where $C$ is a polynomial in $\left(S,\|\varphi\|_{\infty},\left\|\partial_{x} \varphi\right\|_{\infty},\left\|\partial_{x x} \varphi\right\|_{\infty}, M_{\alpha}, M_{\rho / \alpha}\right)$ with positive coefficients.

We prove Lemma 7.8 in Appendix 7.E.2. At $t=1$, by the chain rule for the mutual information, we have for all $\epsilon \in \mathcal{B}_{n}$ :

$$
\begin{align*}
i_{n, \epsilon}(1) & =\frac{I\left(\mathbf{X}^{*} ; \widetilde{\mathbf{Y}}^{(1, \epsilon)} \mid \mathbf{W}\right)+I\left(\mathbf{U} ; \mathbf{Y}^{(1, \epsilon)} \mid \mathbf{W}, \mathbf{V}\right)}{m_{n}} \\
& =\frac{I_{P_{X}^{(n)}}\left(R_{1}(1, \epsilon)\right)}{\alpha_{n}}+I_{P_{\text {out }}}\left(R_{2}(1, \epsilon), 1+2 s_{n}\right) \\
& =\frac{1}{\alpha_{n}} I_{P_{X}^{(n)}}\left(\int_{0}^{1} r_{\epsilon}(t) d t\right)+I_{P_{\text {out }}}\left(\int_{0}^{1} q_{\epsilon}(t) d t, 1\right)+O\left(s_{n}\right), \tag{7.45}
\end{align*}
$$

where to obtain the last equality we use the Lipschitzness of $I_{P_{X}^{(n)}}$ (Lemma 2.3) and $I_{P_{\text {out }}}$ (Lemma 7.22 in Appendix 7.D, and assume that there exists $M_{\rho / \alpha}>0$ such that $\forall n \in \mathbb{N}^{*}: \rho_{n} \alpha_{n} \leq M_{\rho / \alpha}$. The "big O" $O\left(s_{n}\right)$ is a quantity whose absolute value is bounded by $C s_{n}$ where $C$ is a polynomial in $\left(S,\|\varphi\|_{\infty},\left\|\partial_{x} \varphi\right\|_{\infty},\left\|\partial_{x x} \varphi\right\|_{\infty}, M_{\rho / \alpha}\right)$ with positive coefficients.

## 7.A. 2 Fundamental sum rule

We want to compare the original model of interest (model at $t=0$ ) to the purely scalar one $(t=1)$. To do so, we use

$$
i_{n, \epsilon}(0)=i_{n, \epsilon}(1)-\int_{0}^{1} i_{n, \epsilon}^{\prime}(t) d t
$$

where $i_{n, \epsilon}^{\prime}(\cdot)$ is the derivative of $i_{n, \epsilon}(\cdot)$. Once combined with Lemma 7.8 and 7.45, the latter identity yields (note that $O\left(s_{n}\right)=O\left(s_{n} / \sqrt{\rho_{n}}\right)$ since $0<\rho_{n}<1$ )

$$
\begin{align*}
\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}=O\left(\frac{s_{n}}{\sqrt{\rho_{n}}}\right)+\frac{1}{\alpha_{n}} I_{P_{X}^{(n)}}\left(\int_{0}^{1} r_{\epsilon}(t) d t\right) & +I_{P_{\text {out }}}\left(\int_{0}^{1} q_{\epsilon}(t) d t, 1\right) \\
& -\int_{0}^{1} i_{n, \epsilon}^{\prime}(t) d t \tag{7.46}
\end{align*}
$$

From now on, we denote by $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^{n} \times \mathbb{R}^{m_{n}}$ a pair of random vectors sampled from the joint posterior distribution 7.A.1. We denote by angular brackets $\langle-\rangle_{n, t, \epsilon}$ the expectations w.r.t. this posterior, i.e., for a generic function $g$.

$$
\langle g(\mathbf{x}, \mathbf{u})\rangle_{n, t, \epsilon}:=\int g(\mathbf{x}, \mathbf{u}) d P\left(\mathbf{x}, \mathbf{u} \mid \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)
$$

We also define the scalar overlap $Q:=\frac{\mathbf{x}^{\top} \mathbf{X}^{*}}{k_{n}}=\frac{1}{k_{n}} \sum_{i=1}^{n} X_{i}^{*} x_{i}$, the inner product between the true signal $\mathbf{X}^{*}$ and the estimate $\mathbf{x}$ where $(\mathbf{x}, \mathbf{u})$ is sampled from the posterior distribution 7.A.1). We compute $i_{n, \epsilon}^{\prime}$ in Appendix 7.E.1.

Proposition 7.9. Suppose that (H1), (H2), (H3) hold and $\Delta=\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$. Further assume that there exist real positive numbers $M_{\alpha}$ and $M_{\rho / \alpha}$ such that
$\forall n \in \mathbb{N}^{*}: \alpha_{n} \leq M_{\alpha}, \rho_{n} / \alpha_{n} \leq M_{\rho / \alpha}$. For all $y \in \mathbb{R}, \ell_{y}(x):=\ln P_{\text {out }}(y \mid x)$ and $\ell_{y}^{\prime}(\cdot)$ denotes its derivative. Then, for all $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\begin{aligned}
i_{n, \epsilon}^{\prime}(t)=O & \left(\frac{1}{\rho_{n} \sqrt{n}}\right)+\frac{\rho_{n}}{2 \alpha_{n}} r_{\epsilon}(t)\left(1-q_{\epsilon}(t)\right) \\
& +\frac{1}{2} \mathbb{E}\left\langle\left(Q-q_{\epsilon}(t)\right)\left(\frac{1}{m_{n}} \sum_{\mu=1}^{m_{n}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-\frac{\rho_{n}}{\alpha_{n}} r_{\epsilon}(t)\right)\right\rangle_{n, t, \epsilon}
\end{aligned}
$$

where $\left|O\left(\frac{1}{\rho_{n} \sqrt{n}}\right)\right| \leq \frac{\sqrt{C}}{\rho_{n} \sqrt{n}}$ with $C$ a polynomial in $\left(S,\|\varphi\|_{\infty},\left\|\partial_{x} \varphi\right\|_{\infty},\left\|\partial_{x x} \varphi\right\|_{\infty}, M_{\alpha}\right.$, $\left.M_{\rho / \alpha}\right)$ that has positive coefficients and does not depend on $(t, \epsilon)$.

The next key result states that the overlap concentrates on its expectation. This behavior is called replica symmetric in statistical physics. Similar results have been obtained in the spin glass literature [47, [53]. In this work we use a formulation taylored to Bayesian inference problems as developed in the context of LDPC codes, random linear estimation [153] and Nishimori symmetric spin glasses [66], [147], [150].
Proposition 7.10 (Overlap concentration). Suppose that (H1), (H2), (H3) hold, $\Delta=\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$, and the family of functions $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}},\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ are regular. Further assume that there exist real positive numbers $M_{\alpha}, M_{\rho / \alpha}$ and $m_{\rho / \alpha}$ such that $\forall n \in \mathbb{N}^{*}: \alpha_{n} \leq M_{\alpha}$ and $\frac{m_{\rho / \alpha}}{n}<\frac{\rho_{n}}{\alpha_{n}} \leq M_{\rho / \alpha}$. Then, for all $t \in[0,1]$ :

$$
\int_{\mathcal{B}_{n}} \frac{d \epsilon}{s_{n}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(Q-\mathbb{E}\langle Q\rangle_{n, t, \epsilon}\right)^{2}\right\rangle_{n, t, \epsilon} \leq C M_{n}
$$

where

$$
M_{n}:=\frac{1}{s_{n}^{2} \rho_{n}^{2}\left(\frac{\rho_{n} n}{\alpha_{n} m_{\rho / \alpha}}\right)^{1 / 3}-s_{n}^{2} \rho_{n}^{2}}>0
$$

and $C$ is a polynomial in $\left(S,\|\varphi\|_{\infty},\left\|\partial_{x} \varphi\right\|_{\infty},\left\|\partial_{x x} \varphi\right\|_{\infty}, M_{\alpha}, M_{\rho / \alpha}, m_{\rho / \alpha}\right)$ with positive coefficients.

We prove Proposition 7.10 in Appendix 7.G. We can now prove the fundamental sum rule.

Proposition 7.11 (Fundamental sum rule). Suppose that $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ : $q_{\epsilon}(t)=\mathbb{E}\langle Q\rangle_{n, t, \epsilon}$. Under the assumptions of Proposition 7.10, we have

$$
\begin{aligned}
& \frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}=O\left(\sqrt{M_{n}}\right)+O\left(\frac{s_{n}}{\sqrt{\rho_{n}}}\right) \\
&+\int_{\mathcal{B}_{n}} \frac{d \epsilon}{s_{n}^{2}}\left\{\frac{1}{\alpha_{n}}\right. \\
& I_{P_{X}^{(n)}}\left(\int_{0}^{1} r_{\epsilon}(t) d t\right)+I_{P_{\text {out }}}\left(\int_{0}^{1} q_{\epsilon}(t) d t, 1\right) \\
&\left.\quad-\frac{\rho_{n}}{2 \alpha_{n}} \int_{0}^{1} r_{\epsilon}(t)\left(1-q_{\epsilon}(t)\right) d t\right\}
\end{aligned}
$$

where $\left|O\left(\sqrt{M_{n}}\right)\right| \leq \sqrt{C_{1}} M_{n}$ and $\left|O\left(s_{n} / \sqrt{\rho_{n}}\right)\right| \leq \sqrt{C_{2}} s_{n} / \sqrt{\rho_{n}}$ with $C_{1}, C_{2}$ polynomials in $\left(S,\|\varphi\|_{\infty},\left\|\partial_{x} \varphi\right\|_{\infty},\left\|\partial_{x x} \varphi\right\|_{\infty}, M_{\alpha}, M_{\rho / \alpha}, m_{\rho / \alpha}\right)$ having positive coefficients.

Proof. By Cauchy-Schwarz inequality,

$$
\begin{aligned}
& \left|\int_{\mathcal{B}_{n}} \frac{d \epsilon}{s_{n}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(Q-q_{\epsilon}(t)\right)\left(\frac{1}{m_{n}} \sum_{\mu=1}^{m_{n}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-\frac{\rho_{n}}{\alpha_{n}} r_{\epsilon}(t)\right)\right\rangle_{n, t, \epsilon}\right|^{2} \\
& \leq \int_{\mathcal{B}_{n}} \frac{d \epsilon}{s_{n}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(\frac{1}{m_{n}} \sum_{\mu=1}^{m_{n}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-\frac{\rho_{n}}{\alpha_{n}} r_{\epsilon}(t)\right)^{2}\right\rangle_{n, t, \epsilon} \\
& \quad \cdot \int_{\mathcal{B}_{n}} \frac{d \epsilon}{s_{n}^{2}} \int_{0}^{1} d t \mathbb{E}\left\langle\left(Q-q_{\epsilon}(t)\right)^{2}\right\rangle_{n, t, \epsilon}
\end{aligned}
$$

The first factor on the right-hand side of this inequality is bounded by a constant that depends polynomially on $\|\varphi\|_{\infty}$ and $\left\|\partial_{x} \varphi\right\|_{\infty}^{3}$. Since $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ : $q_{\epsilon}(t)=\mathbb{E}\langle Q\rangle_{n, t, \epsilon}$, the second term is in $O\left(M_{n}\right)$ (see Proposition 7.10). Therefore, by Proposition 7.9 ,

$$
\begin{equation*}
\int_{\mathcal{B}_{n}} \frac{d \epsilon}{s_{n}^{2}} \int_{0}^{1} d t i_{n, \epsilon}^{\prime}(t)=O\left(\sqrt{M_{n}}\right)+O\left(\frac{1}{\rho_{n} \sqrt{n}}\right)+\int_{\mathcal{B}_{n}} \frac{d \epsilon}{s_{n}^{2}} \int_{0}^{1} d t \frac{\rho_{n}}{2 \alpha_{n}} r_{\epsilon}(t)\left(1-q_{\epsilon}(t)\right) . \tag{7.47}
\end{equation*}
$$

Note that ${ }^{1} \rho_{n} \sqrt{n}=O\left(\sqrt{M_{n}}\right)$. Integrating (7.46) over $\epsilon \in \mathcal{B}_{n}$ and then making use of (7.47) gives the result.

## 7.A. 3 Matching bounds

To prove Theorem 7.1, we lower and upper bound $I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right) / m_{n}$ by the same quantity, up to a small error. To do so, we plug two different choices for $R(\cdot, \epsilon):=\left(R_{1}(\cdot, \epsilon), R_{2}(\cdot, \epsilon)\right)$ in the sum-rule of Proposition 7.11. In both cases, the function $R(\cdot, \epsilon)$ is the solution to a first-order ordinary differential equation (ODE). We now describe these ODEs.

Fix $t \in[0,1]$ and $R=\left(R_{1}, R_{2}\right) \in[0,+\infty) \times\left[0, t+2 s_{n}\right]$. We observe

$$
\begin{cases}Y_{\mu}^{\left(t, R_{2}\right)} \sim P_{\mathrm{out}}\left(\cdot \mid S_{\mu}^{\left(t, R_{2}\right)}\right), & 1 \leq \mu \leq m_{n} \\ \widetilde{Y}_{i}^{\left(t, R_{1}\right)}=\sqrt{R_{1}} X_{i}^{*}+\widetilde{Z}_{i}, \quad 1 \leq i \leq n\end{cases}
$$

where

$$
S_{\mu}^{\left(t, R_{2}\right)}:=\sqrt{\frac{1-t}{k_{n}}}\left(\mathbf{W X}^{*}\right)_{\mu}+\sqrt{R_{2}} V_{\mu}+\sqrt{t+2 s_{n}-R_{2}} U_{\mu}
$$

The joint posterior density of $\left(\mathbf{X}^{*}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{\left(t, R_{2}\right)}, \widetilde{\mathbf{Y}}^{\left(t, R_{1}\right)}, \mathbf{W}, \mathbf{V}\right)$ is

$$
\begin{aligned}
& d P\left(\mathbf{x}, \mathbf{u} \mid \mathbf{Y}^{\left(t, R_{2}\right)}, \widetilde{\mathbf{Y}}^{\left(t, R_{1}\right)}, \mathbf{W}, \mathbf{V}\right) \\
& \quad \propto \prod_{i=1}^{n} d P_{X}^{(n)}\left(x_{i}\right) e^{-\frac{1}{2}\left(\sqrt{R_{1}} x_{i}-\widetilde{Y}_{i}^{\left(t, R_{1}\right)}\right)^{2}} \prod_{\mu=1}^{m_{n}} \frac{d u_{\mu}}{\sqrt{2 \pi}} e^{-\frac{u_{\mu}^{2}}{2}} P_{\mathrm{out}}\left(Y_{\mu}^{\left(t, R_{2}\right)} \mid s_{\mu}^{\left(t, R_{2}\right)}\right),
\end{aligned}
$$

[^21]where
$$
s_{\mu}^{\left(t, R_{2}\right)}:=\sqrt{\frac{1-t}{k_{n}}}(\mathbf{W} \mathbf{x})_{\mu}+\sqrt{R_{2}} V_{\mu}+\sqrt{t+2 s_{n}-R_{2}} u_{\mu}
$$

We denote by the angular brackets $\langle-\rangle_{n, t, R}$ the expectation with respect to this posterior distribution and define

$$
F_{1}^{(n)}(t, R):=-\left.2 \frac{\alpha_{n}}{\rho_{n}} \frac{\partial I_{P_{\text {out }}}}{\partial q}\right|_{q=\mathbb{E}\langle Q\rangle_{n, t, R}, \rho=1}, \quad F_{2}^{(n)}(t, R):=\mathbb{E}\langle Q\rangle_{n, t, R}
$$

Let $r \in\left[0, r_{\max }\right]$. We consider the two following first-order ODEs:

$$
\begin{align*}
& \frac{d y}{d t}=\left(\frac{\alpha_{n}}{\rho_{n}} r, F_{2}^{(n)}(t, y)\right)  \tag{7.48}\\
& \frac{d y}{d t}=\left(F_{1}^{(n)}(t, y), F_{2}^{(n)}(t, y)\right) \tag{7.49}
\end{align*}
$$

The next proposition sums up useful properties of the solutions to these two ODEs, i.e., our two kinds of interpolation functions. The proof is given in Appendix 7.H.

Proposition 7.12. Suppose that (H1), (H2), (H3) hold and $\Delta=\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$. For every $\epsilon \in \mathcal{B}_{n}$, there exists a unique global solution $R(\cdot, \epsilon):[0,1] \rightarrow[0,+\infty)^{2}$ to the initial value problem

$$
\frac{d y}{d t}=\left(F_{1}^{(n)}(t, y), F_{2}^{(n)}(t, y)\right), \quad y(0)=\epsilon
$$

$R(\cdot, \epsilon)$ is continuously differentiable and the image of its derivative $R^{\prime}(\cdot, \epsilon)$ satisfies

$$
R^{\prime}([0,1], \epsilon) \subseteq\left[0, \frac{\alpha_{n}}{\rho_{n}} r_{\max }\right] \times[0,1]
$$

where $r_{\text {max }}:=-\left.2\left(\partial I_{P_{\text {out }}} / \partial q\right)\right|_{q=1, \rho=1} \geq 0$. Besides, for all $t \in[0,1], R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $\mathcal{B}_{n}$ onto its image whose Jacobian determinant is greater than, or equal to, one, i.e., $\forall \epsilon \in \mathcal{B}_{n}$ :

$$
\operatorname{det} J_{R(t, \cdot)}(\epsilon) \geq 1
$$

where $J_{R(t,)}$ denotes the Jacobian matrix of $R(t, \cdot)$. Finally, the same statement holds true if, for a fixed $r \in\left[0, r_{\max }\right]$, we instead consider the initial value problem

$$
\frac{d y}{d t}=\left(\frac{\alpha_{n}}{\rho_{n}} r, F_{2}^{(n)}(t, y)\right), y(0)=\epsilon
$$

Proposition 7.13 (Upper bound). Suppose that (H1), (H2), (H3) hold and $\Delta=\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$. Further assume that there exist positive numbers $M_{\alpha}, M_{\rho / \alpha}$ and $m_{\rho / \alpha}$ such that $\forall n \in \mathbb{N}^{*}: \alpha_{n} \leq M_{\alpha}, \frac{m_{\rho / \alpha}}{n}<\frac{\rho_{n}}{\alpha_{n}} \leq M_{\rho / \alpha}$. Then, $\forall n \in \mathbb{N}^{*}$ :

$$
\begin{equation*}
\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}} \leq \inf _{r \in\left[0, r_{\max }\right]} \sup _{q \in[0,1]} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)+O\left(\sqrt{M_{n}}\right)+O\left(\frac{s_{n}}{\sqrt{\rho_{n}}}\right) \tag{7.50}
\end{equation*}
$$

Proof. Fix $r \in\left[0, r_{\max }\right]$. For all $\epsilon \in \mathcal{B}_{n}$, we choose $R(\cdot, \epsilon):=\left(R_{1}(\cdot, \epsilon), R_{2}(\cdot, \epsilon)\right)$ to be the unique solution to the ODE (7.48) with initial condition $R(0, \epsilon)=\epsilon$ (see Proposition 7.12). Then, we define

$$
q_{\epsilon}(t):=R_{2}^{\prime}(t, \epsilon)=\mathbb{E}\langle Q\rangle_{n, t, \epsilon}, \quad r_{\epsilon}(t):=R_{1}^{\prime}(t, \epsilon)=\frac{\alpha_{n} r}{\rho_{n}}
$$

By Proposition 7.12, the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ are regular. We can now apply Proposition 7.11 to get

$$
\begin{align*}
\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}} & =\int_{\mathcal{B}_{n}} \frac{d \epsilon}{s_{n}^{2}} i_{\mathrm{RS}}\left(\int_{0}^{1} q_{\epsilon}(t) d t, r ; \alpha_{n}, \rho_{n}\right)+O\left(\sqrt{M_{n}}\right)+O\left(\frac{s_{n}}{\sqrt{\rho_{n}}}\right) \\
& \leq \sup _{q \in[0,1]} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)+O\left(\sqrt{M_{n}}\right)+O\left(\frac{s_{n}}{\sqrt{\rho_{n}}}\right) . \tag{7.51}
\end{align*}
$$

The inequality (7.51) holds for all $r \in\left[0, r_{\max }\right]$ and the constant factors in the quantities $O\left(\sqrt{M_{n}}\right), O\left(s_{n} / \sqrt{\rho_{n}}\right)$ are uniform in $r$. Hence the inequality (7.50) with the infimum over $r$.

Proposition 7.14 (Lower bound). Under the same hypotheses than Proposition 7.13. we have $\forall n \in \mathbb{N}^{*}$ :

$$
\begin{equation*}
\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}} \geq \inf _{r \in\left[0, r_{\max }\right]} \sup _{q \in[0,1]} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)+O\left(\sqrt{M_{n}}\right)+O\left(\frac{s_{n}}{\sqrt{\rho_{n}}}\right) \tag{7.52}
\end{equation*}
$$

Proof. For all $\epsilon \in \mathcal{B}_{n}$, we choose $R(\cdot, \epsilon):=\left(R_{1}(\cdot, \epsilon), R_{2}(\cdot, \epsilon)\right)$ to be the unique solution to the $\operatorname{ODE}(7.49)$ with initial condition $R(0, \epsilon)=\epsilon($ see Proposition 7.12). Then, we define

$$
q_{\epsilon}(t):=R_{2}^{\prime}(t, \epsilon)=\mathbb{E}\langle Q\rangle_{n, t, \epsilon}, \quad r_{\epsilon}(t):=R_{1}^{\prime}(t, \epsilon)=-\left.\frac{2 \alpha_{n}}{\rho_{n}} \frac{\partial I_{P_{\text {out }}}}{\partial q}\right|_{q=q_{\epsilon}(t), \rho=1}
$$

By Proposition 7.12, the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ are regular. Note that $\forall \epsilon \in \overline{\mathcal{B}_{n}}$ :

$$
\begin{align*}
& \frac{1}{\alpha_{n}} I_{P_{X}^{(n)}}\left(\int_{0}^{1} r_{\epsilon}(t) d t\right)+I_{P_{\text {out }}}\left(\int_{0}^{1} q_{\epsilon}(t) d t, 1\right)-\frac{\rho_{n}}{2 \alpha_{n}} \int_{0}^{1} r_{\epsilon}(t)\left(1-q_{\epsilon}(t)\right) d t \\
& \geq \int_{0}^{1}\left\{\frac{1}{\alpha_{n}} I_{P_{X}^{(n)}}\left(r_{\epsilon}(t)\right)+I_{P_{\text {out }}}\left(q_{\epsilon}(t), 1\right)-\frac{\rho_{n}}{2 \alpha_{n}} r_{\epsilon}(t)\left(1-q_{\epsilon}(t)\right)\right\} d t \\
&=\int_{0}^{1}\left\{\sup _{q \in[0,1]} \frac{1}{\alpha_{n}} I_{P_{X}^{(n)}}\left(r_{\epsilon}(t)\right)+I_{P_{\text {out }}}(q, 1)-\frac{\rho_{n}}{2 \alpha_{n}} r_{\epsilon}(t)(1-q)\right\} d t \\
&=\int_{0}^{1} \sup _{q \in[0,1]} i_{\mathrm{RS}}\left(q, \frac{\rho_{n}}{\alpha_{n}} r_{\epsilon}(t) ; \alpha_{n}, \rho_{n}\right) d t  \tag{7.53}\\
& \geq \inf _{r \in\left[0, r_{\max }\right]} \sup _{q \in[0,1]} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) . \tag{7.54}
\end{align*}
$$

The first inequality is an application of Jensen's inequality to the concave functions $I_{P_{X}^{(n)}}$ and $I_{P_{\text {out }}}(\cdot, 1)$ (see Lemmas 2.3 and 7.22 ). The subsequent equality is because the global maximum of the concave function

$$
h: q \in[0,1] \mapsto I_{P_{\text {out }}}(q, 1)-\frac{\rho_{n}}{2 \alpha_{n}} r_{\epsilon}(t)(1-q)
$$

is reached at $q_{\epsilon}(t)$ since $h^{\prime}\left(q_{\epsilon}(t)\right)=0$. To write (7.53) we simply use the definition (7.7) of $i_{\text {RS }}$. Finally, the inequality (7.54) is because $r_{\epsilon}(t) \in\left[0, \frac{\alpha_{n}}{\rho_{n}} r_{\text {max }}\right]$ so we lower bound the integrand in (7.53) by $\inf _{r \in\left[0, r_{\max }\right]} \sup _{q \in[0,1]} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$, a quantity independent of $t \in[0,1]$. We now apply Proposition 7.11 and make use of (7.54) to obtain the inequality 7.52).

## 7.A. 4 Combining the matching bounds

Proof of Theorem 7.1. We choose $\rho_{n}=\Theta\left(n^{-\lambda}\right), \alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ and $s_{n}=\Theta\left(n^{-\beta}\right)$ with $\lambda \in[0,1 / 9), \gamma>0, \beta \in(\lambda / 2,1 / 6-\lambda)$. Then, we apply Propositions 7.13 and 7.14, and use the identity (we refer to [29, Proposition 7 and Corollary 7 in SI] for the proof)

$$
\begin{aligned}
\inf _{r \in\left[0, r_{\max }\right]} \sup _{q \in[0,1]} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) & =\inf _{r \geq 0} \sup _{q \in[0,1]} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \\
& =\inf _{q \in[0,1]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) .
\end{aligned}
$$

We obtain the upper bound

$$
\left|\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}-\inf _{q \in[0,1]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)\right| \leq O\left(\sqrt{M_{n}}\right)+O\left(\frac{s_{n}}{\sqrt{\rho_{n}}}\right) .
$$

Optimizing over $\beta \in(\lambda / 2,1 / 6-\lambda)$ to maximize the convergence rate of

$$
O\left(\sqrt{M_{n}}\right)+O\left(\frac{s_{n}}{\sqrt{\rho_{n}}}\right)=O\left(\max \left\{\frac{1}{n^{\beta-\lambda / 2}}, \frac{|\ln n|^{1 / 6}}{n^{1 / 6-\lambda-\beta}}\right\}\right)
$$

yields Theorem 7.1.

## 7.B Proof of Theorem 7.2 for a general discrete prior

In the whole appendix we assume that $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ where $P_{0}$ is a discrete probability distribution with a finite support

$$
\operatorname{supp}\left(P_{0}\right) \subseteq\left\{-v_{K},-v_{K-1}, \ldots,-v_{1}, v_{1}, v_{2}, \ldots, v_{K}\right\},
$$

where $0<v_{1}<v_{2}<\cdots<v_{K}$. For all $i \in\{1, \ldots, K\}, P_{0}\left(v_{i}\right)=p_{i}^{+}$and $P_{0}\left(-v_{i}\right)=p_{i}^{-}$where $p_{i}^{+}, p_{i}^{-} \geq 0$ and $p_{i}:=p_{i}^{+}+p_{i}^{-}>0$. Of course, $\sum_{i=1}^{K} p_{i}=1$.

Note that the second moment of $X_{0} \sim P_{0}$ is $\mathbb{E} X_{0}^{2}=\sum_{j=1}^{K} p_{j} v_{j}^{2}$ and the support of $\left|X_{0}\right|$ equals $\left\{v_{1}, v_{2}, \ldots, v_{K}\right\}$.

For $\rho_{n}, \alpha_{n}>0$ we denote the variational problem appearing in Theorem 7.1 by

$$
I\left(\rho_{n}, \alpha_{n}\right):=\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right),
$$

where the potential $i_{\mathrm{RS}}$ is defined in (7.7). Let $X^{*} \sim P_{X}^{(n)}$ and $Z \sim \mathcal{N}(0,1)$ be independent random variables. We define for all $r \geq 0$ :

$$
\begin{align*}
\psi_{P_{X}^{(n)}}(r) & :=\mathbb{E}\left[\ln \int d P_{X}^{(n)}(x) e^{-\frac{r}{2} x^{2}+r X^{*} x+\sqrt{r} x Z}\right]  \tag{7.55}\\
& =\mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r X^{*} v_{i}+\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-r X^{*} v_{i}-\sqrt{r} Z v_{i}}\right)\right)\right]
\end{align*}
$$

Note that $I_{P_{X}^{(n)}}(r):=I\left(X^{*} ; \sqrt{r} X^{*}+Z\right)=r \rho_{n} \mathbb{E} X_{0}^{2} / 2-\psi_{P_{X}^{(n)}}(r)$ where $X_{0} \sim P_{0}$ so

$$
\begin{equation*}
I\left(\rho_{n}, \alpha_{n}\right)=\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\sup _{r \geq 0}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\} . \tag{7.56}
\end{equation*}
$$

The latter expression for $I\left(\rho_{n}, \alpha_{n}\right)$ is easier to work with. We point out that $\psi_{P_{X}^{(n)}}$ is twice differentiable, nondecreasing, strictly convex and ( $\rho_{n} \mathbb{E} X_{0}^{2} / 2$ )-Lipschitz on $[0,+\infty)$ (see Lemma 2.3) while $I_{P_{\text {out }}}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$ is nonincreasing concave on $\left[0, \mathbb{E} X_{0}^{2}\right]$ (see [29, Appendix B.2, Proposition 18]).

Our goal is now to compute the limit of $I\left(\rho_{n}, \alpha_{n}\right)$ when $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$ and $\rho_{n} \rightarrow 0$. We first look where the supremum over $r$ is reached depending on the value of $q \in\left[0, \mathbb{E} X_{0}^{2}\right]$.
Lemma 7.15. Let $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ where $P_{0}$ is a discrete distribution with finite support $\operatorname{supp}\left(P_{0}\right) \subseteq\left\{ \pm v_{1}, \pm v_{2}, \ldots, \pm v_{K}\right\}$ with $0<v_{1}<v_{2}<\cdots<v_{K}$. Let $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$. Define $g_{\rho_{n}}: r \in(0,+\infty) \mapsto \frac{2}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)$ and $\forall \rho_{n} \in\left(0, e^{-1}\right), \forall j \in\{1, \ldots, K\}$ :

$$
\begin{equation*}
a_{\rho_{n}}^{(j)}:=g_{\rho_{n}}\left(\frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{j}^{2}}\right) \quad, \quad b_{\rho_{n}}^{(j)}:=g_{\rho_{n}}\left(\frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{j}^{2}}\right) . \tag{7.57}
\end{equation*}
$$

Let $X_{0} \sim P_{0}$. For $\rho_{n}$ small enough, we have

$$
\begin{equation*}
\rho_{n} \mathbb{E}\left[X_{0}\right]^{2}<a_{\rho_{n}}^{(K)}<b_{\rho_{n}}^{(K)}<a_{\rho_{n}}^{(K-1)}<b_{\rho_{n}}^{(K-1)}<\cdots<a_{\rho_{n}}^{(1)}<b_{\rho_{n}}^{(1)}<\mathbb{E} X_{0}^{2} \tag{7.58}
\end{equation*}
$$

and for all $j \in\{1, \ldots, K\}$ :

$$
\begin{equation*}
\lim _{\rho_{n} \rightarrow 0} a_{\rho_{n}}^{(j)}=\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|>v_{j}\right\}}\right] \quad ; \quad \lim _{\rho_{n} \rightarrow 0} b_{\rho_{n}}^{(j)}=\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{j}\right\}}\right] \tag{7.59}
\end{equation*}
$$

Besides, for every $q \in\left(\rho_{n} \mathbb{E}[X]^{2}, \mathbb{E} X_{0}^{2}\right)$ there exists a unique $r_{n}^{*}(q) \in(0,+\infty)$ such that

$$
\begin{equation*}
\frac{r_{n}^{*}(q) q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}(q)\right)=\sup _{r \geq 0} \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right), \tag{7.60}
\end{equation*}
$$

and $\forall j \in\{1, \ldots, K\}, \forall q \in\left[a_{\rho_{n}}^{(j)}, b_{\rho_{n}}^{(j)}\right]:$

$$
\frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{j}^{2}} \leq r_{n}^{*}(q) \leq \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{j}^{2}}
$$

These bounds are tight, namely, $r_{n}^{*}\left(a_{\rho_{n}}^{(j)}\right)=\frac{2\left(1-\left|\ln \rho_{n}\right|^{\frac{1}{4}}\right)}{\gamma v_{j}^{2}}$ and $r_{n}^{*}\left(b_{\rho_{n}}^{(j)}\right)=\frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{j}^{2}}$.
Proof. For every $q \in(0,1)$ we define $f_{\rho_{n}, q}: r \in[0,+\infty) \mapsto \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)$ whose supremum over $r$ we want to compute. The derivative of $f_{\rho_{n}, q}$ with respect to $r$ reads

$$
f_{\rho_{n}, q}^{\prime}(r)=\frac{q}{2}-\frac{1}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\frac{\alpha_{n}}{\rho_{n}} r\right) .
$$

The derivative $\psi_{P_{X}^{(n)}}^{\prime}$ is continuously increasing and thus one-to-one from $(0,+\infty)$ onto $\left(\rho_{n}^{2} \mathbb{E}\left[X_{0}\right]^{2} / 2, \rho_{n} \mathbb{E} X_{0}^{2} / 2\right)$. Therefore, if $q \in\left(0, \rho_{n} \mathbb{E}\left[X_{0}\right]^{2}\right)$ then $f_{\rho_{n}, q}^{\prime} \leq 0$ and the supremum of $f_{\rho_{n}, q}$ is achieved at $r=0$. On the contrary, if $q \in\left(\rho_{n}, \mathbb{E} X_{0}^{2}\right)$ then there exists a unique solution $r_{n}^{*}(q) \in(0,+\infty)$ to the stationary point equation $f_{\rho_{n}, q}^{\prime}(r)=0$. As $f_{\rho_{n}, q}$ is concave ( $\psi_{P_{0}, n}$ is convex) this solution $r_{n}^{*}(q)$ is the global maximum of $f_{\rho_{n}, q}$. Let us transform the stationary point equation. We have

$$
\begin{equation*}
f_{\rho_{n}, q}(r)=0 \Leftrightarrow \frac{2}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)=q \Leftrightarrow g_{\rho_{n}}(r)=q \tag{7.61}
\end{equation*}
$$

where $g_{\rho_{n}}: r \mapsto \frac{2}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\alpha_{n} r / \rho_{n}\right)$ is continuously increasing and one-to-one from $(0,+\infty)$ to $\left(\rho_{n} \mathbb{E} X_{0}^{2}, \mathbb{E} X_{0}^{2}\right)$. Thus, by definition of $a_{\rho_{n}}^{(j)}$ and $b_{\rho_{n}}^{(j)}$, we have

$$
r_{n}^{*}\left(a_{\rho_{n}}^{(j)}\right)=\frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{j}^{2}}, \quad r_{n}^{*}\left(b_{\rho_{n}}^{(j)}\right)=2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right) / \gamma v_{j}^{2} .
$$

Besides, if $q=\in\left[a_{\rho_{n}}^{(j)}, b_{\rho_{n}}^{(j)}\right]$ then

$$
\frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{j}^{2}} \leq r_{n}^{*}(q) \leq \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{j}^{2}}
$$

since $g_{\rho_{n}}$ is increasing on $[0,+\infty), g_{\rho_{n}}\left(r_{n}^{*}(q)\right)=q$ and $a_{\rho_{n}}^{(j)}=g_{\rho_{n}}\left(2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right) / \gamma v_{j}^{2}\right)$, $b_{\rho_{n}}^{(j)}=g_{\rho_{n}}\left(2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right) / \gamma v_{j}^{2}\right)$. Because $g_{\rho_{n}}$ is increasing and $0<v_{1}<\cdots<v_{k}$, it is clear that we have the ordering (7.58) provided that $\rho_{n}$ is close enough to 0 .

We are left with proving the limits 7.59 . In order to so, we first rewrite the derivative of $\psi_{P_{X}^{(n)}}$. For all $r \geq 0$, we have

$$
\begin{aligned}
& \psi_{P_{X}^{(n)}}^{\prime}(r)=\frac{1}{2} \mathbb{E}\left[X^{*} \frac{\rho_{n} \sum_{i=1}^{K} v_{i} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r X^{*} v_{i}+\sqrt{r} Z v_{i}}-p_{i}^{-} e^{-r X^{*} v_{i}-\sqrt{r} Z v_{i}}\right)}{1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r X^{*} v_{i}+\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-r X^{*} v_{i}-\sqrt{r} Z v_{i}}\right)}\right] \\
& \quad=\frac{\rho_{n}^{2}}{2} \sum_{j=1}^{K} p_{j}^{+} v_{j} \mathbb{E}\left[\frac{\sum_{i=1}^{K} v_{i} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r v_{i} v_{j}+\sqrt{r} Z v_{i}}-p_{i}^{-} e^{-r v_{i} v_{j}-\sqrt{r} Z v_{i}}\right)}{1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r v_{i} v_{j}+\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-r v_{i} v_{j}-\sqrt{r} Z v_{i}}\right)}\right] \\
& \quad+\frac{\rho_{n}^{2}}{2} \sum_{j=1}^{K} p_{j}^{-} v_{j} \mathbb{E}\left[\frac{\sum_{i=1}^{K} v_{i} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{-} e^{r v_{i} v_{j}+\sqrt{r} Z v_{i}}-p_{i}^{+} e^{-r v_{i} v_{j}-\sqrt{r} Z v_{i}}\right)}{1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{-} e^{r v_{i} v_{j}+\sqrt{r} Z v_{i}}+p_{i}^{+} e^{-r v_{i} v_{j}-\sqrt{r} Z v_{i}}\right)}\right]
\end{aligned}
$$

and, after diving numerator and denominator by $\rho_{n}$, we get

$$
\begin{aligned}
\psi_{P_{X}^{(n)}}^{\prime}(r) & =\frac{\rho_{n}}{2} \sum_{j=1}^{K} p_{j}^{+} v_{j} \mathbb{E}\left[\frac{\sum_{i=1}^{K} v_{i} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r v_{i} v_{j}+\sqrt{r} Z v_{i}}-p_{i}^{-} e^{-r v_{i} v_{j}-\sqrt{r} Z v_{i}}\right)}{\frac{1-\rho_{n}}{\rho_{n}}+\sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r v_{i} v_{j}+\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-r v_{i} v_{j}-\sqrt{r} Z v_{i}}\right)}\right] \\
& +\frac{\rho_{n}}{2} \sum_{j=1}^{K} p_{j}^{-} v_{j} \mathbb{E}\left[\frac{\sum_{i=1}^{K} v_{i} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{-} e^{r v_{i} v_{j}+\sqrt{r} Z v_{i}}-p_{i}^{+} e^{-r v_{i} v_{j}-\sqrt{r} Z v_{i}}\right)}{\frac{1-\rho_{n}}{\rho_{n}}+\sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{-} e^{r v_{i} v_{j}+\sqrt{r} Z v_{i}}+p_{i}^{+} e^{-r v_{i} v_{j}-\sqrt{r} Z v_{i}}\right)}\right]
\end{aligned}
$$

This last expression shortens to

$$
\begin{align*}
& \psi_{P_{X}^{(n)}}^{\prime}(r)=\frac{\rho_{n}}{2} \sum_{j=1}^{K} p_{j}^{+} v_{j} \mathbb{E}\left[h\left(Z, r, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{+}, \mathbf{p}^{-}\right)\right] \\
&+\frac{\rho_{n}}{2} \sum_{j=1}^{K} p_{j}^{-} v_{j} \mathbb{E}\left[h\left(Z, r, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{-}, \mathbf{p}^{+}\right)\right] \tag{7.62}
\end{align*}
$$

where $\mathbf{v}:=\left(v_{1}, v_{2}, \ldots, v_{K}\right), \mathbf{p}^{+}:=\left(p_{1}^{+}, p_{2}^{+}, \ldots, p_{K}^{+}\right), \mathbf{p}^{-}:=\left(p_{1}^{-}, p_{2}^{-}, \ldots, p_{K}^{-}\right)$and we define $\forall(z, r, u) \in \mathbb{R} \times[0,+\infty) \times(0,+\infty)$ :

$$
\begin{aligned}
& h\left(z, r, u ; \rho_{n}, \mathbf{v}, \mathbf{p}^{ \pm}, \mathbf{p}^{\mp}\right) \\
&:=\frac{\sum_{i=1}^{K} v_{i} e^{-\frac{r\left(v_{i}-u\right)^{2}}{2}}+\sqrt{r} z\left(v_{i}-u\right)}{}\left(p_{i}^{ \pm}-p_{i}^{\mp} e^{-2 r v_{i} u-2 \sqrt{r} z v_{i}}\right) \\
& \frac{1-\rho_{n}}{\rho_{n}} e^{-\frac{r u^{2}}{2}-\sqrt{r} z u}+\sum_{i=1}^{K} e^{-\frac{r\left(v_{i}-u\right)^{2}}{2}+\sqrt{r} z\left(v_{i}-u\right)}\left(p_{i}^{ \pm}+p_{i}^{\mp} e^{-2 r v_{i} u-2 \sqrt{r} z v_{i}}\right)
\end{aligned} .
$$

Note that $\forall z \in \mathbb{R}$ :

$$
\begin{align*}
& h\left(z, \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{ \pm}, \mathbf{p}^{\mp}\right) \underset{\rho_{n} \rightarrow 0}{\longrightarrow} \begin{cases}0 & \text { if } j<k ; \\
v_{j} & \text { if } j \geq k\end{cases}  \tag{7.63}\\
& h\left(z, \frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{ \pm}, \mathbf{p}^{\mp}\right) \underset{\rho_{n} \rightarrow 0}{\longrightarrow}\left\{\begin{array}{l}
0 \text { if } j \leq k ; \\
v_{j} \text { if } j>k .
\end{array}\right. \tag{7.64}
\end{align*}
$$

By the dominated convergence theorem, making use of the identity (7.62) and the limit (7.63), we have $\forall k \in\{1, \ldots, K\}$ :

$$
\begin{aligned}
a_{\rho_{n}}^{(k)}: & =g_{\rho_{n}}\left(\frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{k}^{2}}\right) \\
= & \frac{2}{\rho_{n}} \psi^{\prime} P_{X}^{(n)}\left(\frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)\left|\ln \rho_{n}\right|}{v_{k}^{2}}\right) \\
= & \sum_{j=1}^{K} p_{j}^{+} v_{j} \mathbb{E}\left[h\left(z, \frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \mathbf{v}, \mathbf{p}^{+}, \mathbf{p}^{-}\right)\right] \\
& \quad+\sum_{j=1}^{K} p_{j}^{-} v_{j} \mathbb{E}\left[h\left(z, \frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \mathbf{v}, \mathbf{p}^{-}, \mathbf{p}^{+}\right)\right] \\
& \xrightarrow[\rho_{n} \rightarrow 0]{\longrightarrow} \sum_{j>k} p_{j}^{+} v_{j}^{2}+\sum_{j>k} p_{j}^{-} v_{j}^{2}=\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|>v_{k}\right\}}\right] .
\end{aligned}
$$

Similarly, using this time the limit (7.64), we have $\forall k \in\{1, \ldots, K\}$ :

$$
\begin{aligned}
b_{\rho_{n}}^{(k)}: & =g_{\rho_{n}}\left(\frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{k}^{2}}\right) \\
= & \frac{2}{\rho_{n}} \psi_{P_{X}^{(n)}}^{\prime}\left(\frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)\left|\ln \rho_{n}\right|}{v_{k}^{2}}\right) \\
= & \sum_{j=1}^{K} p_{j}^{+} v_{j} \mathbb{E}\left[h\left(z, \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \mathbf{v}, \mathbf{p}^{+}, \mathbf{p}^{-}\right)\right] \\
& \quad+\sum_{j=1}^{K} p_{j}^{-} v_{j} \mathbb{E}\left[h\left(z, \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \mathbf{v}, \mathbf{p}^{-}, \mathbf{p}^{+}\right)\right] \\
\xrightarrow[\rho_{n} \rightarrow 0]{\longrightarrow} & \sum_{j \geq k} p_{j}^{+} v_{j}^{2}+\sum_{j \geq k} p_{j}^{-} v_{j}^{2}=\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right] .
\end{aligned}
$$

Note that $\lim _{\rho_{n} \rightarrow 0} b_{\rho_{n}}^{(j)}=\lim _{\rho_{n} \rightarrow 0} a_{\rho_{n}}^{(j-1)}$. Thus, Lemma 7.15 essentially states that, in the limit $\rho_{n} \rightarrow 0$, the segment $\left[0, \mathbb{E} X_{0}^{2}\right]$ can be broken into $K$ subsegments $\left[a_{\rho_{n}}^{(j)}, b_{\rho_{n}}^{(j)}\right]$ such that the point where the supremum over $r$ is achieved is located in an interval shrinking on $r^{*}:=2 / \gamma v_{j}^{2}$ for all $q \in\left[a_{\rho_{n}}^{(j)}, b_{\rho_{n}}^{(j)}\right]$. The next step is then to determine what is the limit of $\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2}{\gamma v_{j}^{2}}\right)$.

Lemma 7.16. Let $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ where $P_{0}$ is a discrete distribution with finite support $\operatorname{supp}\left(P_{0}\right) \subseteq\left\{ \pm v_{1}, \pm v_{2}, \ldots, \pm v_{K}\right\}$ with $0<v_{1}<v_{2}<\cdots<v_{K}$. Let $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$. Then, for every $k \in\{1, \ldots, K\}$ :

$$
\begin{equation*}
\lim _{\rho_{n} \rightarrow 0} \frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1 \pm\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{k}^{2}}\right)=\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right]}{\gamma v_{k}^{2}}-\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} \tag{7.65}
\end{equation*}
$$

where $X_{0} \sim P_{0}$.
Proof. Fix $k \in\{1, \ldots, K\}$. The function $\psi_{P_{X}^{(n)}}$ is Lipschitz continuous with Lipschitz constant $\frac{\rho_{n} \mathbb{E} X_{0}^{2}}{2}$. Therefore,

$$
\begin{aligned}
\left\lvert\, \frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1 \pm\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{k}^{2}}\right)\right. & \left.-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2}{\gamma v_{k}^{2}}\right) \right\rvert\, \\
& \leq \frac{\rho_{n} \mathbb{E} X_{0}^{2}}{2 \alpha_{n}}\left|\frac{\alpha_{n}}{\rho_{n}} \frac{2\left|\ln \rho_{n}\right|^{-\frac{1}{4}}}{\gamma v_{k}^{2}}\right|=\frac{\mathbb{E} X_{0}^{2}}{\gamma v_{k}^{2}}\left|\ln \rho_{n}\right|^{-\frac{1}{4}} .
\end{aligned}
$$

The latter inequality shows that the limits of $\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1+\left|\ln \rho_{n}\right|^{1 / 4}\right)}{\gamma v_{k}^{2}}\right)$ and $\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1-\left|\ln \rho_{n}\right|^{-1 / 4}\right)}{\gamma v_{k}^{2}}\right)$ are the same and equal to the limit of $\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2}{\gamma v_{k}^{2}}\right)$.

To compute the latter, let us write $\psi_{P_{X}^{(n)}}(r)$ in a more explicit form. For all $r \geq 0$ :

$$
\begin{aligned}
& \psi_{P_{X}^{(n)}}(r):=\mathbb{E}\left[\ln \int d P_{X}^{(n)}(x) e^{-\frac{r}{2} x^{2}+r X^{*} x+\sqrt{r} x Z}\right] \\
&= \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r X^{*} v_{i}+\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-r X^{*} v_{i}-\sqrt{r} Z v_{i}}\right)\right)\right] \\
&=\left(1-\rho_{n}\right) \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-\sqrt{r} Z v_{i}}\right)\right)\right] \\
&+\rho_{n} \sum_{j=1}^{K} p_{j}^{+} \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r v_{j} v_{i}+\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-r v_{j} v_{i}-\sqrt{r} Z v_{i}}\right)\right)\right] \\
&+\rho_{n} \sum_{j=1}^{K} p_{j}^{-} \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{-r v_{j} v_{i}+\sqrt{r} Z v_{i}}+p_{i}^{-} e^{r v_{j} v_{i}-\sqrt{r} Z v_{i}}\right)\right)\right] \\
&=\left(1-\rho_{n}\right) \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-\sqrt{r} Z v_{i}}\right)\right)\right] \\
&+\rho_{n} \sum_{j=1}^{K} p_{j}^{+} \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{r v_{j} v_{i}+\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-r v_{j} v_{i}-\sqrt{r} Z v_{i}}\right)\right)\right] \\
&+\rho_{n} \sum_{j=1}^{K} p_{j}^{-} \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{-} e^{r v_{j} v_{i}+\sqrt{r} Z v_{i}}+p_{i}^{+} e^{-r v_{j} v_{i}-\sqrt{r} Z v_{i}}\right)\right)\right],
\end{aligned}
$$

where the last equality is obtained by replacing $Z$ by $-Z$ in the expectations of the last sum over $j$ without changing the values of these expectations since $Z \sim \mathcal{N}(0,1)$ has a symmetric distribution. Hence, we see that

$$
\begin{gather*}
\psi_{P_{X}^{(n)}}(r)=\left(1-\rho_{n}\right) \mathbb{E}\left[\ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{r v_{i}^{2}}{2}}\left(p_{i}^{+} e^{\sqrt{r} Z v_{i}}+p_{i}^{-} e^{-\sqrt{r} Z v_{i}}\right)\right)\right] \\
+\frac{\rho_{n} r \mathbb{E} X_{0}^{2}}{2}+\rho_{n} \ln \rho_{n}+\rho_{n} \sum_{j=1}^{K} p_{j}^{+} \mathbb{E}\left[\ln \widetilde{h}\left(Z, r, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{+}, \mathbf{p}^{-}\right)\right] \\
\quad+\rho_{n} \sum_{j=1}^{K} p_{j}^{-} \mathbb{E}\left[\ln \widetilde{h}\left(Z, r, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{-}, \mathbf{p}^{+}\right)\right] \tag{7.66}
\end{gather*}
$$

where $\mathbf{v}:=\left(v_{1}, v_{2}, \ldots, v_{K}\right), \mathbf{p}^{+}:=\left(p_{1}^{+}, p_{2}^{+}, \ldots, p_{K}^{+}\right), \mathbf{p}^{-}:=\left(p_{1}^{-}, p_{2}^{-}, \ldots, p_{K}^{-}\right)$and we define $\forall(z, r, u) \in \mathbb{R} \times[0,+\infty) \times(0,+\infty)$ :

$$
\begin{aligned}
& \widetilde{h}\left(z, r, u ; \rho_{n}, \mathbf{v}, \mathbf{p}^{ \pm}, \mathbf{p}^{\mp}\right) \\
& \quad:=\frac{1-\rho_{n}}{\rho_{n}} e^{-\frac{r u^{2}}{2}-\sqrt{r} z u}+\sum_{i=1}^{K} e^{-\frac{r\left(v_{i}-u\right)^{2}}{2}+\sqrt{r} z\left(v_{i}-u\right)}\left(p_{i}^{ \pm}+p_{i}^{\mp} e^{-2 r v_{i} u-2 \sqrt{r} z v_{i}}\right) .
\end{aligned}
$$

It follows directly from (7.66) that

$$
\begin{align*}
\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2}{\gamma v_{k}^{2}}\right)= & \frac{A_{\rho_{n}}}{\gamma}+\frac{\mathbb{E} X_{0}^{2}}{\gamma v_{k}^{2}}-\frac{1}{\gamma} \\
& +\frac{1}{\gamma} \sum_{j=1}^{K} p_{j}^{+} \mathbb{E}\left[\frac{\ln \widetilde{h}\left(Z, \frac{2\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{+}, \mathbf{p}^{-}\right)}{\left|\ln \rho_{n}\right|}\right] \\
& +\frac{1}{\gamma} \sum_{j=1}^{K} p_{j}^{-} \mathbb{E}\left[\frac{\ln \widetilde{h}\left(Z, \frac{2\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{-}, \mathbf{p}^{+}\right)}{\left|\ln \rho_{n}\right|}\right], \tag{7.67}
\end{align*}
$$

where

$$
\begin{aligned}
A_{\rho_{n}}= & \frac{1-\rho_{n}}{\rho_{n}\left|\ln \rho_{n}\right|} \\
& \cdot \mathbb{E} \ln \left(1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{v_{i}^{2}}{v_{k}^{2}}\left|\ln \rho_{n}\right|}\left(p_{i}^{+} e^{\left|\frac{2 v_{i}^{2}\left|\ln \rho_{n}\right|}{v_{k}^{2}}\right|^{\frac{1}{2}} Z}+p_{i}^{-} e^{-\left|\frac{2 v_{i}^{2}\left|\ln \rho_{n}\right|}{v_{k}^{2}}\right|^{\frac{1}{2}} Z}\right)\right) .
\end{aligned}
$$

Next we show that $A_{\rho_{n}}$ vanishes when $\rho_{n} \rightarrow 0$. We can use the inequalities $\frac{x}{1+x} \leq \ln (1+x) \leq x$, valid for all $x>-1$, to get the following bounds on $A_{\rho_{n}}$ :

$$
\begin{aligned}
A_{\rho_{n}} & \leq \frac{1-\rho_{n}}{\left|\ln \rho_{n}\right|}\left(\mathbb{E}\left[\sum_{i=1}^{K} e^{-\frac{v_{i}^{2}}{v_{k}^{2}}\left|\ln \rho_{n}\right|}\left(p_{i}^{+} e^{\left|\frac{2 v_{i}^{2}\left|\ln \rho_{n}\right|}{v_{k}^{2}}\right|^{\frac{1}{2}} Z}+p_{i}^{-} e^{-\left|\frac{2 v_{i}^{2}\left|\ln \rho_{n}\right|}{v_{k}^{2}}\right|^{\frac{1}{2}} Z}\right)\right]-1\right) \\
& =\frac{1-\rho_{n}}{\left|\ln \rho_{n}\right|}\left(\sum_{i=1}^{K} p_{i} e^{-2 \frac{v_{i}^{2}}{v_{k}^{2}}\left|\ln \rho_{n}\right|}-1\right) \leq-\frac{1-\rho_{n}}{\left|\ln \rho_{n}\right|} ; \\
A_{\rho_{n}} & \geq \frac{1-\rho_{n}}{\left|\ln \rho_{n}\right|} \mathbb{E}\left[\frac{\sum_{i=1}^{K} e^{-\frac{v_{i}^{2}}{v_{k}^{2}}\left|\ln \rho_{n}\right|}\left(p_{i}^{+} e^{\left|\frac{2 v_{i}^{2} \mid \ln \rho_{n}}{v_{k}^{2}}\right|^{\frac{1}{2}} Z}+p_{i}^{-} e^{-\left|\frac{2 v_{i}^{2}\left|\ln \rho_{n}\right|}{v_{k}^{2}}\right|^{\frac{1}{2}} Z}\right)-1}{\left.1-\rho_{n}+\rho_{n} \sum_{i=1}^{K} e^{-\frac{v_{i}^{2}}{v_{k}^{2}}\left|\ln \rho_{n}\right|}\left(p_{i}^{+} e^{\left|\frac{2 v_{i}^{2}\left|\ln \rho_{n}\right|}{v_{k}^{2}}\right|^{\frac{1}{2}} Z}+p_{i}^{-} e^{-\left|\frac{2 v_{i}^{2}\left|\ln \rho_{n}\right|}{v_{k}^{2}}\right|^{\frac{1}{2}} Z}\right)\right]}\right. \\
& \geq-\frac{1}{\left|\ln \rho_{n}\right|} .
\end{aligned}
$$

The last inequality is because ${ }^{(x-1)} /\left(1-\rho_{n}+\rho_{n} x\right) \geq-1 /\left(1-\rho_{n}\right)$ for $x>0$. Together the upper bound and lower bound imply that $\left|A_{\rho_{n}}\right| \leq 1 /\left|\ln \rho_{n}\right| \xrightarrow[\rho_{n} \rightarrow 0]{ } 0$. To conclude the proof, we need to compute the limits of each summand in both sums over $j \in\{1, \ldots, K\}$ on the right-hand side of (7.67). Note that $\forall z \in \mathbb{R}$ :

$$
\begin{aligned}
& \widetilde{h}\left(z, \frac{2\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{ \pm}, \mathbf{p}^{\mp}\right)=\left(1-\rho_{n}\right) e^{\left|\ln \rho_{n}\right|\left(1-\frac{v_{j}^{2}}{v_{k}^{2}}-\sqrt{\frac{2 v_{j}^{2}}{v_{k}^{2}\left|\ln \rho_{n}\right|}} z\right)} \\
& \quad+\sum_{i=1}^{K} e^{-\left|\ln \rho_{n}\right|\left(\frac{\left(v_{i}-v_{j}\right)^{2}}{v_{k}^{2}}-\sqrt{\frac{2}{\left|\ln \rho_{n}\right|} \frac{v_{i}-v_{j}}{v_{k}}} z\right)}\left(p_{i}^{ \pm}+p_{i}^{\mp} e^{-4\left|\ln \rho_{n}\right| \frac{v_{i}}{v_{k}}\left(\frac{v_{j}}{v_{k}}+\frac{z}{\sqrt{2\left|\ln \rho_{n}\right|}}\right)}\right)
\end{aligned}
$$

From this last expression we easily deduce the following pointwise limits for all $z \in \mathbb{R}$ :

$$
\frac{\ln \widetilde{h}\left(z, \frac{2\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{ \pm}, \mathbf{p}^{\mp}\right)}{\left|\ln \rho_{n}\right|} \underset{\rho_{n} \rightarrow 0}{ } \begin{cases}1-\frac{v_{j}^{2}}{v_{k}^{2}} & \text { if } j<k ;  \tag{7.68}\\ 0 & \text { if } j \geq k .\end{cases}
$$

By the dominated convergence theorem, making use of the pointwise limits 7.68,

$$
\begin{align*}
& \sum_{j=1}^{K} p_{j}^{+} \mathbb{E}\left[\frac{\ln \widetilde{h}\left(Z, \frac{2\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{+}, \mathbf{p}^{-}\right)}{\left|\ln \rho_{n}\right|}\right]+p_{j}^{-} \mathbb{E}\left[\frac{\ln \widetilde{h}\left(Z, \frac{2\left|\ln \rho_{n}\right|}{v_{k}^{2}}, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{-}, \mathbf{p}^{+}\right)}{\left|\ln \rho_{n}\right|}\right] \\
& \xrightarrow[\rho_{n} \rightarrow 0]{ } \sum_{j<k}\left(p_{j}^{+}+p_{j}^{-}\right)\left(1-\frac{v_{j}^{2}}{v_{k}^{2}}\right)=\mathbb{P}\left(\left|X_{0}\right|<v_{k}\right)-\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k}\right\}}\right]}{v_{k}^{2}} \tag{7.69}
\end{align*}
$$

Combining the identity (7.67), $\lim _{\rho_{n} \rightarrow 0} A_{\rho_{n}}=0$ and the limit (7.69) yields

$$
\begin{aligned}
\lim _{\rho_{n} \rightarrow 0} \frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2}{\gamma v_{k}^{2}}\right) & =\frac{\mathbb{E} X_{0}^{2}}{\gamma v_{k}^{2}}-\frac{1}{\gamma}+\frac{\mathbb{P}\left(\left|X_{0}\right|<v_{k}\right)}{\gamma}-\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k}\right\}}\right]}{\gamma v_{k}^{2}} \\
& =\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right]}{\gamma v_{k}^{2}}-\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma},
\end{aligned}
$$

thus ending the proof of the proposition.
We can now use Lemmas 7.15 and 7.16 to determine the limits when $\rho_{n} \rightarrow 0$ of the infimum of $\sup _{r \geq 0} i_{\text {RS }}\left(q, r ; \alpha_{n}, \rho_{n}\right)$ over $q$ restrained to different subsegments of $\left[0, \mathbb{E} X_{0}^{2}\right]$.
Proposition 7.17. Let $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ where $P_{0}$ is a discrete distribution with finite support $\operatorname{supp}\left(P_{0}\right) \subseteq\left\{ \pm v_{1}, \pm v_{2}, \ldots, \pm v_{K}\right\}$ with $0<v_{1}<v_{2}<\cdots<v_{K}$. Let $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$. Then, $\forall k \in\{1, \ldots, K\}$ :

$$
\begin{align*}
& \lim _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in\left[a_{\rho n}^{k,}, b_{\rho n}^{(k)}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \\
&=\min \{ \left\{I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|>v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right|>v_{k}\right)}{\gamma},\right. \\
&\left.I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}\right\}, \tag{7.70}
\end{align*}
$$

while $\forall k \in\{2, \ldots, K\}$ :

$$
\begin{align*}
\lim _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in\left[b_{\rho n}^{k n}, a_{\rho n}^{(k-1)}\right]} & \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \\
& =I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} . \tag{7.71}
\end{align*}
$$

Finally,

$$
\begin{align*}
& \lim _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in\left[0, a_{\rho n}^{(K)}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)=I_{P_{\text {out }}}\left(0, \mathbb{E} X_{0}^{2}\right),  \tag{7.72}\\
& \liminf _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in\left[b_{\rho n}^{(1)}, 1\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \geq \frac{1}{\gamma} . \tag{7.73}
\end{align*}
$$

Proof. In the whole proof $\rho_{n}$ is close enough to 0 for the ordering $(7.58)$ to hold. First we prove (7.70). Fix $k \in\{1, \ldots, K\}$. By Lemma 7.15, for all $q \in\left[a_{\rho_{n}}^{(k)}, b_{\rho_{n}}^{(k)}\right]$ :

$$
\sup _{r \geq 0} \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)=\frac{r_{n}^{*}(q) q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}(q)\right),
$$

 increasing imply that $\forall q \in\left[a_{\rho_{n}}^{(k)}, b_{\rho_{n}}^{(k)}\right]$ :

$$
\begin{align*}
& I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{q}{\gamma v_{k}^{2}}\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{k}^{2}}\right) \\
& \quad \leq \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \\
& \leq I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{q}{\gamma v_{k}^{2}}\left(1+\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{k}^{2}}\right) . \tag{7.74}
\end{align*}
$$

These inequalities are valid for every $q \in\left[a_{\rho_{n}}^{(k)}, b_{\rho_{n}}^{(k)}\right]$ so the same inequalities will hold if we take the infimum over $q \in\left[a_{\rho_{n}}^{(k)}, b_{\rho_{n}}^{(k)}\right]$ in (7.74). Note that

$$
q \mapsto I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{q}{\gamma v_{k}^{2}}\left(1 \mp\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)
$$

are concave functions on $\left[a_{\rho_{n}}^{(k)}, b_{\rho_{n}}^{(k)}\right]$ so the minimum of each function is achieved at either endpoint $a_{\rho_{n}}^{(k)}$ or $b_{\rho_{n}}^{(k)}$. Thus,

$$
\begin{align*}
& \inf _{q \in\left[a_{\rho_{n},}^{(k)}, b_{\rho_{n}}^{(k)}\right]} I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{q}{\gamma v_{k}^{2}}\left(1 \pm\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1 \mp\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{k}^{2}}\right) \\
& =-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \frac{2\left(1 \mp\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right)}{\gamma v_{k}^{2}}\right) \\
& +\min _{q \in\left\{a_{\left.\rho_{n}, b_{p n}^{(k)}\right\}}^{(k)}\right.} I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{q}{\gamma v_{k}^{2}}\left(1 \pm\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right) \\
& \xrightarrow[\rho_{n} \rightarrow 0]{\longrightarrow} \frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}-\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right]}{\gamma v_{k}^{2}}+\min _{q \in\left\{\mathbb{E}\left[X_{2}^{2} 1_{\left\{\left|X_{0}\right|>v_{k}\right\}}\right\},\right\}} I_{\mathbb{E}_{\text {out }}\left[X_{0}^{2} 1_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right\}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{q}{\gamma v_{k}^{2}} \\
& =\min \left\{I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|>v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right|>v_{k}\right)}{\gamma},\right. \\
& \left.I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}\right\} . \tag{7.75}
\end{align*}
$$

The limit when $\rho_{n} \rightarrow 0$ follows from (7.59) in Lemma 7.15 and (7.65) in Lemma 7.16. Taking the infimum over $q \in\left[a_{\rho_{n}}^{(k)}, b_{\rho_{n}}^{(k)}\right]$ in 7.74) and using the fact that the upper and lower bounds have the same limit (7.75) ends the proof of 7.70 .

We now turn to the proof of the limit 7.71. Fix $k \in\{2, \ldots, K\}$. Since it is the supremum of nondecreasing functions, the function

$$
\widetilde{\psi}_{P_{X}^{(n)}}: q \in\left[0, \mathbb{E} X_{0}^{2}\right] \mapsto \sup _{r \geq 0} \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)
$$

is nondecreasing. The fact that $I_{P_{\text {out }}}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$ and $\widetilde{\psi}_{P_{X}^{(n)}}$ are respectively nonincreasing and nondecreasing imply that

$$
\begin{align*}
& I_{P_{\text {out }}}\left(a_{\rho_{n}}^{(k-1)}, \mathbb{E} X_{0}^{2}\right)+\widetilde{\psi}_{P_{X}^{(n)}}\left(b_{\rho_{n}}^{(k)}\right) \\
& \quad \leq \inf _{q \in\left[b_{\rho_{n},}^{(k)}, a_{\rho_{n}}^{(k-1)}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \leq I_{P_{\text {out }}}\left(b_{\rho_{n}}^{(k)}, \mathbb{E} X_{0}^{2}\right)+\tilde{\psi}_{P_{X}^{(n)}}\left(a_{\rho_{n}}^{(k-1)}\right) . \tag{7.76}
\end{align*}
$$

By Lemma 7.15, we have

$$
\begin{aligned}
\widetilde{\psi}_{P_{X}^{(n)}}\left(b_{\rho_{n}}^{(k)}\right) & =\frac{r_{n}^{*}\left(b_{\rho_{n}}^{(k)}\right) b_{\rho_{n}}^{(k)}}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}\left(b_{\rho_{n}}^{(k)}\right)\right), \\
\widetilde{\psi}_{P_{X}^{(n)}}\left(a_{\rho_{n}}^{(k-1)}\right) & =\frac{r_{n}^{*}\left(a_{\rho_{n}}^{(k-1)}\right) a_{\rho_{n}}^{(k-1)}}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}\left(a_{\rho_{n}}^{(k-1)}\right)\right),
\end{aligned}
$$

where $r_{n}^{*}\left(b_{\rho_{n}}^{(k)}\right)=2\left(1+\left|\ln \rho_{n}\right|^{-1 / 4}\right) / \gamma v_{k}^{2}$ and $r_{n}^{*}\left(a_{\rho_{n}}^{(k-1)}\right)=2\left(1-\left|\ln \rho_{n}\right|^{-1 / 4}\right) / \gamma v_{k-1}^{2}$. Making use of the limits (7.59) in Lemma 7.15 and (7.65) in Lemma 7.16 yields

$$
\begin{aligned}
\lim _{\rho_{n} \rightarrow 0^{+}} \widetilde{\psi}_{P_{X}^{(x)}}\left(b_{\rho_{n}}^{(k)}\right) & =\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right]}{\gamma v_{k}^{2}}-\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right]}{\gamma v_{k}^{2}}+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} \\
& =\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} ; \\
\lim _{\rho_{n} \rightarrow 0^{+}} \widetilde{\psi}_{P_{X}^{(n)}}\left(a_{\rho_{n}}^{(k-1)}\right) & =\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|>v_{k-1}\right\}}\right]}{\gamma v_{k-1}^{2}}-\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k-1}\right\}}\right]}{\gamma v_{k-1}^{2}}+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k-1}\right)}{\gamma} \\
& =\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} .
\end{aligned}
$$

Besides, $\lim _{\rho_{n} \rightarrow 0^{+}} b_{\rho_{n}}^{(k)}=\lim _{\rho_{n} \rightarrow 0^{+}} a_{\rho_{n}}^{(k-1)}=\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right]$ and $I_{P_{\text {out }}}$ is continuous so

$$
\lim _{\rho_{n} \rightarrow 0^{+}} I_{P_{\text {out }}}\left(b_{\rho_{n}}^{(k)}, \mathbb{E} X_{0}^{2}\right)=\lim _{\rho_{n} \rightarrow 0^{+}} I_{P_{\text {out }}}\left(a_{\rho_{n}}^{(k-1)}, \mathbb{E} X_{0}^{2}\right)=I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right) .
$$

Hence, the lower and upper bounds in 7.76 have the same limit. It ends the proof of (7.71).

The proof of 7.72 is similar to the one of (7.71). We have the bounds

$$
\begin{align*}
I_{P_{\text {out }}}\left(a_{\rho_{n}}^{(K)},\right. & \left.\mathbb{E} X_{0}^{2}\right)+\widetilde{\psi}_{P_{X}^{(n)}}(0) \\
& \leq \inf _{q \in\left[0, a_{\rho_{n}}^{(K)}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \leq I_{P_{\text {out }}}\left(0, \mathbb{E} X_{0}^{2}\right)+\widetilde{\psi}_{P_{X}^{(n)}}\left(a_{\rho_{n}}^{(K)}\right) . \tag{7.77}
\end{align*}
$$

The function $I_{P_{\text {out }}}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$ is continuous and $\lim _{\rho_{n} \rightarrow 0^{+}} a_{\rho_{n}}^{(K)}=0$ so

$$
\lim _{\rho_{n} \rightarrow 0^{+}} I_{P_{\text {out }}}\left(a_{\rho_{n}}^{(K)}, \mathbb{E} X_{0}^{2}\right)=I_{P_{\text {out }}}\left(0, \mathbb{E} X_{0}^{2}\right)
$$

Clearly, $\widetilde{\psi}_{P_{X}^{(n)}}(0)=0$. By Lemma 7.15 .

$$
\widetilde{\psi}_{P_{X}^{(n)}}\left(a_{\rho_{n}}^{(K)}\right)=\frac{r_{n}^{*}\left(a_{\rho_{n}}^{(K)}\right) a_{\rho_{n}}^{(K)}}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}\left(a_{\rho_{n}}^{(K)}\right)\right),
$$

where $r_{n}^{*}\left(a_{\rho_{n}}^{(K)}\right)=2\left(1-\left|\ln \rho_{n}\right|^{-1 / 4}\right) / \gamma v_{K}^{2}$. It follows from the limits 7.59) (Lemma 7.15) and 7.65) (Lemma 7.16 that $\lim _{\rho_{n} \rightarrow 0^{+}} \widetilde{\psi}_{P_{X}^{(n)}}\left(a_{\rho_{n}}^{(K)}\right)=0$. Thus, the lower and upper bounds in 7.77) have the same limit. It ends the proof of 7.72.

We are left with proving (7.73). The fact that $I_{P_{\text {out }}}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$ and $\widetilde{\psi}_{P_{X}^{(n)}}$ are respectively nonincreasing and nondecreasing imply that

$$
\begin{equation*}
\inf _{q \in\left[b_{\rho_{n},}^{(k)}, \mathbb{E} X_{0}^{2}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \geq I_{P_{\text {out }}}\left(\mathbb{E} X_{0}^{2}, \mathbb{E} X_{0}^{2}\right)+\widetilde{\psi}_{P_{X}^{(n)}}\left(b_{\rho_{n}}^{(1)}\right)=\widetilde{\psi}_{P_{X}^{(n)}}\left(b_{\rho_{n}}^{(1)}\right) . \tag{7.78}
\end{equation*}
$$

Note that the right-hand side of (7.78) has a limit. More precisely,

$$
\begin{aligned}
\widetilde{\psi}_{P_{X}^{(n)}}\left(b_{\rho_{n}}^{(1)}\right) & =\frac{r_{n}^{*}\left(b_{\rho_{n}}^{(1)}\right) b_{\rho_{n}}^{(1)}}{2}-\frac{\psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}\left(b_{\rho_{n}}^{(1)}\right)\right)}{\alpha_{n}} \\
& \xrightarrow[\rho_{n} \rightarrow 0^{+}]{\longrightarrow} \frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{1}\right\}}\right]}{\gamma v_{1}^{2}}-\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{1}\right\}}\right]}{\gamma v_{1}^{2}}+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{1}\right)}{\gamma}=\frac{1}{\gamma} .
\end{aligned}
$$

Taking the limit inferior on both sides of $(7.78)$ and using the latter limit proves the inequality (7.73).

Proposition 7.18. Let $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ where $P_{0}$ is a discrete distribution with finite support

$$
\operatorname{supp}\left(P_{0}\right) \subseteq\left\{-v_{K},-v_{K-1}, \ldots,-v_{1}, v_{1}, v_{2}, \ldots, v_{K}\right\}
$$

where $0<v_{1}<\cdots<v_{K}<v_{K+1}=+\infty$. Let $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$ and $X_{0} \sim P_{0}$. Then, the quantity $I\left(\rho_{n}, \alpha_{n}\right):=\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$ has a limit when $\rho_{n} \rightarrow 0^{+}$and

$$
\lim _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}\right)=\min _{1 \leq k \leq K+1}\left\{I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}\right\} .
$$

Proof. The proof goes in two steps. We first prove a upper bound on the limit superior of $I\left(\rho_{n}, \alpha_{n}\right)$, and then prove a lower bound on the limit inferior thats turns out to match the limit superior.

Upper bound on the limit superior Note the following trivial upper bound:

$$
\begin{equation*}
I\left(\rho_{n}, \alpha_{n}\right) \leq \min _{1 \leq k \leq K}\left\{\inf _{q \in\left[a_{p n}^{(k)}, b_{p n}^{(k)}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)\right\} . \tag{7.79}
\end{equation*}
$$

The upper bound on the limit superior of $I\left(\rho_{n}, \alpha_{n}\right)$ thus directly follows from (7.79) and Proposition 7.17 on the limits of the infimums over $q \in\left[a_{\rho_{n}}^{(k)}, b_{\rho_{n}}^{(k)}\right]$

$$
\begin{align*}
& \limsup _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}\right) \leq \min _{1 \leq k \leq K} \min \{ \\
& I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|>v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right|>v_{k}\right)}{\gamma}, \\
&\left.I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}\right\}  \tag{7.80}\\
&= \min _{1 \leq k \leq K+1}\left\{I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}\right\} .
\end{align*}
$$

Matching lower bound on the limit inferior The lower bound on the limit inferior is obtained by studying the infimum on each segment of the following partition:

$$
\begin{equation*}
\left[0, \mathbb{E} X_{0}^{2}\right]=\left[0, a_{\rho_{n}}^{(K)}\right] \cup\left(\bigcup_{k=1}^{K}\left[a_{\rho_{n}}^{(k)}, b_{\rho_{n}}^{(k)}\right]\right) \cup\left(\bigcup_{k=2}^{K}\left[b_{\rho_{n}}^{(k)}, a_{\rho_{n}}^{(k-1)}\right]\right) \cup\left[b_{\rho_{n}}^{(1)}, \mathbb{E} X_{0}^{2}\right] \tag{7.81}
\end{equation*}
$$

By Proposition 7.17, we directly have:

$$
\begin{aligned}
& \liminf _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in \mathrm{U}_{k=1}^{K}\left[a_{\rho n}^{(k)}, b_{\rho n}^{(k)}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \\
& =\min _{1 \leq k \leq K+1}\left\{I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} 1_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}\right\} ; \\
& \begin{aligned}
& \liminf _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in \cup_{k=2}^{K}\left[b_{\left.\rho_{n}^{(k)}, a_{\rho_{n}}^{(k-1)}\right]} \sup _{r \geq 0}\right.} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \\
& \quad=\min _{2 \leq k \leq K}\left\{I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}\right\} ;
\end{aligned} \\
& \liminf _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in\left[0, a_{\rho_{n}}^{(K)}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right) \\
& =I_{P_{\text {out }}}\left(0, \mathbb{E} X_{0}^{2}\right)=I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq+\infty\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq+\infty\right)}{\gamma} ;
\end{aligned}
$$

$\liminf _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in\left[b_{\rho_{n}}^{(1)}, 1\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$

$$
\geq \frac{1}{\gamma}=I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{1}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{1}\right)}{\gamma}
$$

Following the partition (7.81), the limit inferior of $\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r ; \alpha_{n}, \rho_{n}\right)$ is equal to the minimum of the above four limits inferior. Hence,

$$
\begin{equation*}
\liminf _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}\right) \geq \min _{1 \leq k \leq K+1}\left\{I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}\right\} \tag{7.82}
\end{equation*}
$$

We see that the lower bound (7.82) on the limit inferior matches the upper bound (7.80) on the limit superior, thus ending the proof.

Proof of Theorem 7.2 Combining Theorem 7.1 together with Proposition 7.18 gives Theorem 7.2

## 7.C Asymptotic minimum mean-square error: proof of Theorem 7.3

Let $\widehat{\mathbf{X}}=\widehat{\mathbf{X}}(\mathbf{Y}, \mathbf{W})$ be an estimator of $\mathbf{X}^{*}$ that is a function of the observations $\mathbf{Y}$ and the measurement matrix $\mathbf{W}$. Then the mean-square error of this estimator is $\mathbb{E}\left\|\mathbf{X}^{*}-\widehat{\mathbf{X}}\right\|^{2} / k_{n} \in\left[0, \mathbb{E}_{X_{0} \sim P_{0}} X_{0}^{2}\right]$ where the normalization factor $k_{n}:=n \rho_{n}$ is the expected sparsity of $\mathbf{X}^{*}$. It is well-known that the Bayes estimator $\mathbb{E}\left[\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right]$ achieves the minimum mean-square error (MMSE) among all estimators of the form $\widehat{\mathbf{X}}(\mathbf{Y}, \mathbf{W})$. We denote the mean-square error of the Bayes estimator by

$$
\begin{equation*}
\operatorname{MMSE}\left(\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right):=\frac{\mathbb{E}\left\|\mathbf{X}^{*}-\mathbb{E}\left[\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right]\right\|^{2}}{k_{n}} \tag{7.83}
\end{equation*}
$$

The MMSE is therefore a tight lower bound on the error that we achieve when estimating $\mathbf{X}^{*}$ from the observations $\mathbf{Y}$ and the known measurement matrix $\mathbf{W}$. For this reason a result on the MMSE is easier to interprete than a result on the normalized mutual information $I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right) / m_{n}$. In this section, we prove Theorem 7.3 , that is, a formula for the asymptotic MMSE when $n$ diverges to infinity while $\rho_{n}=\Theta\left(n^{-\lambda}\right)$ with $\lambda \in(0,1 / 9)$ and $\alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ with $\gamma>0$. The proof of this theorem is given at the end of this section. The proof relies on the I-MMSE relation [51] that links the MMSE to the derivative of the mutual information with respect to the signal-to-noise ratio of some well-chosen observation channel. For this reason, we first have to determine the asymptotic mutual information of a modified inference problem in which, in addition to the observations (7.2), we have access to the side information $\widetilde{\mathbf{Y}}^{(\tau)}=\sqrt{\alpha_{n} \tau / \rho_{n}} \mathbf{X}^{*}+\widetilde{\mathbf{Z}}$ with $\tau>0$ and $\widetilde{\mathbf{Z}}$ an additive white Gaussian noise. Indeed, the parameter $\tau$ is akin to a signal-to-noise ratio and the derivative of the mutual information $I\left(\mathbf{X}^{*} ; \mathbf{Y}, \widetilde{\mathbf{Y}}^{(\tau)} \mid \mathbf{W}\right) / m_{n}$ with respect to $\tau$ yields half the MMSE [51]:

$$
\frac{\partial}{\partial \tau}\left(\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y}, \tilde{\mathbf{Y}}^{(\tau)} \mid \mathbf{W}\right)}{m_{n}}\right)=\frac{\operatorname{MMSE}\left(\mathbf{X}^{*} \mid \mathbf{Y}, \widetilde{\mathbf{Y}}^{(\tau)}, \mathbf{W}\right)}{2} \underset{\tau \rightarrow 0^{+}}{ } \frac{\operatorname{MMSE}\left(\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right)}{2}
$$

## 7.C.1 Generalized linear estimation with side information

Let $\left(X_{i}^{*}\right)_{i=1}^{n} \stackrel{\text { i.i.d. }}{\sim} P_{X}^{(n)}$ be the components of the signal vector $\mathbf{X}^{*}$. We now have access to the observations

$$
\left\{\begin{array}{lll}
Y_{\mu} & \sim P_{\mathrm{out}}\left(\cdot \left\lvert\, \frac{\left(\mathbf{W} \mathbf{X}^{*}\right)_{\mu}}{\sqrt{k_{n}}}\right.\right), & 1 \leq \mu \leq m_{n}  \tag{7.84}\\
\widetilde{Y}_{i}^{(\tau)} & =\sqrt{\frac{\alpha_{n}}{\rho_{n}} \tau} X_{i}^{*}+\widetilde{Z}_{i} & , \\
1 \leq i \leq n
\end{array}\right.
$$

where $\tau \geq 0$. Remember that the transition kernel $P_{\text {out }}$ is defined in (7.4) using the activation function $\varphi$ and the probability distribution $P_{A}$. The side information only induces a small change in the RS potential whose extremization gives the asymptotic normalized mutual information. More precisely, the potential now reads

$$
\begin{equation*}
i_{\mathrm{RS}}\left(q, r, \tau ; \alpha_{n}, \rho_{n}\right):=\frac{1}{\alpha_{n}} I_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}}(r+\tau)\right)+I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)-\frac{r\left(\mathbb{E} X_{0}^{2}-q\right)}{2}, \tag{7.85}
\end{equation*}
$$

where $X_{0} \sim P_{0}$. We then have the following generalization of Theorem 7.1.
Theorem 7.19 (Normalized mutual information of the GLM with side information at sublinear sparsity and sampling rate). Suppose that $\Delta>0$ and the following hypotheses hold:
(H1) There exists $S>0$ such that the support of $P_{0}$ is included in $[-S, S]$.
(H2) $\varphi$ is bounded, and its first and second partial derivatives with respect to its first argument exist, are bounded and continuous. They are denoted $\partial_{x} \varphi$, $\partial_{x x} \varphi$.
(H3) $W_{\mu i} \stackrel{i . i . d .}{\sim} \mathcal{N}(0,1)$ for $(\mu, i) \in\left\{1, \ldots, m_{n}\right\} \times\{1, \ldots, n\}$.
Let $\rho_{n}=\Theta\left(n^{-\lambda}\right)$ with $\lambda \in[0,1 / 9)$ and $\alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ with $\gamma>0$. Then, $\forall n \in \mathbb{N}^{*}$ :

$$
\begin{equation*}
\left|\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y}, \widetilde{\mathbf{Y}}^{(\tau)} \mid \mathbf{W}\right)}{m_{n}}-\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r, \tau ; \alpha_{n}, \rho_{n}\right)\right| \leq \frac{\sqrt{C}|\ln n|^{1 / 6}}{n^{\frac{1}{12}-\frac{3 \lambda}{4}}} \tag{7.86}
\end{equation*}
$$

where $X_{0} \sim P_{0}$ and $C$ is a polynomial in $\left(\tau, S,\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x \varphi}}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x \varphi}}{\sqrt{\Delta}}\right\|_{\infty}, \lambda, \gamma\right)$ with positive coefficients.

Proof. The proof is similar to the proof of Theorem 7.1 except for a small change in the adaptive interpolation method due to the side information. More precisely, at $t \in[0,1]$, we have access to the observations

$$
\begin{cases}Y_{\mu}^{(t, \epsilon)} & \sim \quad P_{\mathrm{out}}\left(\cdot \mid S_{\mu}^{(t, \epsilon)}\right)  \tag{7.87}\\ \widetilde{Y}_{i}^{(t, \epsilon, \tau)}=\sqrt{\frac{\alpha_{n}}{\rho_{n}} \tau+R_{1}(t, \epsilon)} X_{i}^{*}+\widetilde{Z}_{i}, 1 \leq i \leq n \leq m_{n} \\ \hline\end{cases}
$$

where $X_{i}^{*} \stackrel{\text { i.i.d. }}{\sim} P_{X}^{(n)}, \widetilde{Z}_{i} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ and

$$
S_{\mu}^{(t, \epsilon)}:=\sqrt{\frac{1-t}{k_{n}}} \sum_{i=1}^{n} W_{\mu i} X_{i}^{*}+\sqrt{R_{2}(t, \epsilon)} V_{\mu}+\sqrt{t+2 s_{n}-R_{2}(t, \epsilon)} U_{\mu}
$$

with $W_{\mu i}, V_{\mu}, U_{\mu} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. The proof then goes by looking to the interpolating mutual information $I\left(\left(\mathbf{X}^{*}, \mathbf{U}\right) ;\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon, \tau)}\right) \mid \mathbf{W}\right) / m_{n}$, and follows exactly the same lines than the proof of Theorem 7.1. In particular, the interpolation functions $\left(R_{1}, R_{2}\right)$ are chosen a posteriori as the solutions to the same first-order ordinary differential equations than for Theorem 7.1.

Let $X^{*} \sim P_{X}^{(n)}$ and $Z \sim \mathcal{N}(0,1)$ be independent random variables. We define for all $r \geq 0$ :

$$
\psi_{P_{X}^{(n)}}(r):=\mathbb{E}\left[\ln \int d P_{X}^{(n)}(x) e^{-\frac{r}{2} X_{0}^{2}+r X^{*} x+\sqrt{r} x Z}\right]
$$

Note that $I_{P_{X}^{(n)}}(r):=I\left(X^{*} ; \sqrt{r} X^{*}+Z\right)=\frac{r \rho_{n} \mathbb{E}\left[X_{0}^{2}\right]}{2}-\psi_{P_{X}^{(n)}}(r)$ where $X_{0} \sim P_{0}$. For fixed $\rho_{n}, \alpha_{n}>0$ and $\tau \geq 0$, we denote the variational problem appearing in Theorem 7.1 by

$$
\begin{align*}
I\left(\rho_{n}, \alpha_{n}, \tau\right) & :=\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} \sup _{r \geq 0} i_{\mathrm{RS}}\left(q, r, \tau ; \alpha_{n}, \rho_{n}\right) \\
& =\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{\tau \mathbb{E} X_{0}^{2}}{2}+\sup _{r \geq 0}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}}(r+\tau)\right)\right\} \\
& =\inf _{q \in\left[0, \mathbb{E} X_{0}^{2}\right]} I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{\tau\left(\mathbb{E} X_{0}^{2}-q\right)}{2}+\sup _{r \geq \tau}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\}, \tag{7.88}
\end{align*}
$$

where $X_{0} \sim P_{0}$. Similarly to what is done in Section 7.4 when $P_{0}=\delta_{1}$, we compute the limit of $I\left(\rho_{n}, \alpha_{n}, \tau\right)$ when $P_{0}$ is a discrete distribution with finite support.
Proposition 7.20. Let $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ where $P_{0}$ is a discrete probability distribution with finite support

$$
\operatorname{supp}\left(P_{0}\right) \subseteq\left\{-v_{K},-v_{K-1}, \ldots,-v_{1}, v_{1}, v_{2}, \ldots, v_{K}\right\}
$$

where $0<v_{1}<\cdots<v_{K}<v_{K+1}=+\infty$. Let $\alpha_{n}:=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ for a fixed $\gamma>0$. For every $\tau \in\left[0,2 / \gamma v_{K}^{2}\right)$, the limit of $I\left(\rho_{n}, \alpha_{n}, \tau\right)$ defined in (7.88) exists when $\rho_{n} \rightarrow 0^{+}$and (in what follows $X_{0} \sim P_{0}$ )

$$
\begin{aligned}
\lim _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}, \tau\right)=\min _{1 \leq k \leq K+1}\left\{I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)\right. & +\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} \\
& \left.+\frac{\tau \mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k}\right\}}\right]}{2}\right\}
\end{aligned}
$$

Proof. Fix $\tau \in\left[0,2 / \gamma v_{K}^{2}\right)$ and define $\widetilde{I}_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right):=I_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{\tau\left(\mathbb{E} X_{0}^{2}-q\right)}{2}$. From (7.88), we have

$$
\begin{equation*}
I\left(\rho_{n}, \alpha_{n}, \tau\right)=\inf _{q \in\left[0 \mathbb{E} X_{0}^{2}\right]} \widetilde{I}_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\sup _{r \geq \tau}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\} \tag{7.89}
\end{equation*}
$$

Note that (7.89) has a form similar to $I\left(\rho_{n}, \alpha_{n}\right)$ defined by (7.56) in Appendix 7.B The only differences are that $\widetilde{I}_{P_{\text {out }}}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$ replaces $I_{P_{\text {out }}}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$ and the supremum is over $r \in[\tau,+\infty)$ instead of $r \in[0,+\infty)$. Crucially, $\widetilde{I}_{P_{\text {out }}}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$ is concave nonincreasing on $\left[0, \mathbb{E} X_{0}^{2}\right]$ exactly like $I_{P_{\text {out }}}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$. For these reasons, we can reproduce most of the analysis of Appendix $7 . \mathrm{B}$ - where we compute
the limit $\lim _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}\right)$ - in order to obtain the limit of $I\left(\rho_{n}, \alpha_{n}, \tau\right)$ when $\rho_{n}$ vanishes positively. We just need to be careful with the fact that the supremum is now over $r \in[\tau,+\infty)$.

By Lemma 7.15 in Appendix $7 . \mathrm{B}$, for every $q \in\left(\rho_{n} \mathbb{E}\left[X_{0}\right]^{2}, \mathbb{E} X_{0}^{2}\right)$ there exists a unique $r_{n}^{*}(q) \in(0,+\infty)$ such that

$$
\begin{equation*}
\frac{r_{n}^{*}(q) q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}(q)\right)=\sup _{r \geq 0} \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right), \tag{7.90}
\end{equation*}
$$

and $r_{n}^{*}(q) \geq{ }^{2\left(1-\left|\ln \rho_{n}\right|^{-\frac{1}{4}}\right) / \gamma v_{K}^{2}}$ for all $q \in\left[a_{\rho_{n}}^{(K)}, \mathbb{E} X_{0}^{2}\right)$, where $a_{\rho_{n}}^{(K)}$ is defined in the same lemma by (7.57). By assumption $\tau<2 / \gamma v_{K}^{2}$ so $r_{n}^{*}(q)>\tau$ for all $q \in\left[a_{\rho_{n}}^{(K)}, \mathbb{E} X_{0}^{2}\right)$ when $\rho_{n}$ is close enough to 0 . It follows that $\forall q \in\left[a_{\rho_{n}}^{(K)}, \mathbb{E} X_{0}^{2}\right)$ :

$$
\begin{equation*}
\frac{r_{n}^{*}(q) q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r_{n}^{*}(q)\right)=\sup _{r \geq \tau} \frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right), \tag{7.91}
\end{equation*}
$$

where we have replaced the supremum over $r \in[0,+\infty)$ in (7.90) by a supremum over $r \in[\tau,+\infty)$. Thanks to the identity (7.91) we can repeat the analysis leading
 no difference as we only need for $\widetilde{I}_{\text {out }}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$ to be concave nonincreasing), and we obtain the limit

$$
\begin{align*}
\lim _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in\left[a_{\rho_{n}}^{(K)}, \mathbb{E} X_{o}^{2}\right]} & \sup _{r \geq \tau} \widetilde{I}_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right) \\
& =\min _{1 \leq k \leq K+1} \widetilde{I}_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} . \tag{7.92}
\end{align*}
$$

Note that the limit $(7.92)$ is for the infimum over $q \in\left[a_{\rho_{n}}^{(K)}, \mathbb{E} X_{0}^{2}\right]$, not the infimum over $q \in\left[0, \mathbb{E} X_{0}^{2}\right]$. This is because, for $q \in\left(\rho_{n} \mathbb{E} X_{0}^{2}, a_{\rho_{n}}^{(K)}\right), r_{n}^{*}(q)$ does not necessarily satisfy (7.91). However, the limit (7.92) directly implies the following upper bound on the limit superior,

$$
\begin{equation*}
\limsup _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}, \tau\right) \leq \min _{1 \leq k \leq K+1} \widetilde{I}_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} \tag{7.93}
\end{equation*}
$$

In order to lower bound the limit inferior, we have to lower bound the infimum over $q \in\left[0, a_{\rho_{n}}^{(K)}\right]$ of $\widetilde{I}_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\sup _{r \geq \tau}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\alpha_{n} r / \rho_{n}\right)\right\}$. Because $\widetilde{I}_{P_{\text {out }}}\left(\cdot, \mathbb{E} X_{0}^{2}\right)$ is nonincreasing and $q \mapsto \sup _{r \geq \tau}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\alpha_{n} r / \rho_{n}\right)\right\}$ is nondecreasing (it is the supremum of nondecreasing functions), we have

$$
\begin{align*}
\inf _{q \in\left[0, a_{\rho_{n}}^{(K)}\right]} \widetilde{I}_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\sup _{r \geq \tau}\left\{\frac{r q}{2}\right. & \left.-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\} \\
& \geq \widetilde{I}_{P_{\text {out }}}\left(a_{\rho_{n}}^{(K)}, \mathbb{E} X_{0}^{2}\right)+\sup _{r \geq \tau}\left\{-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\} \\
& =\widetilde{I}_{P_{\text {out }}}\left(a_{\rho_{n}}^{(K)}, \mathbb{E} X_{0}^{2}\right)-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \tau\right) . \tag{7.94}
\end{align*}
$$

The last equality follows from $\psi_{P_{X}^{(n)}}$ being nondecreasing (see Lemma 2.3). We can use the computations in the proof of Lemma 7.16 to write $\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}\left(\frac{\alpha_{n} \tau}{\rho_{n}}\right)}$ more explicitly.We have

$$
\begin{align*}
\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n} \tau}{\rho_{n}}\right)= & \frac{B_{\rho_{n}}}{\gamma}+\frac{\tau \mathbb{E} X_{0}^{2}}{2}-\frac{1}{\gamma} \\
& +\frac{1}{\gamma} \sum_{j=1}^{K} p_{j}^{+} \mathbb{E}\left[\frac{\ln \widetilde{h}\left(Z, \gamma \tau\left|\ln \rho_{n}\right|, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{+}, \mathbf{p}^{-}\right)}{\left|\ln \rho_{n}\right|}\right] \\
& +\frac{1}{\gamma} \sum_{j=1}^{K} p_{j}^{-} \mathbb{E}\left[\frac{\ln \widetilde{h}\left(Z, \gamma \tau\left|\ln \rho_{n}\right|, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{-}, \mathbf{p}^{+}\right)}{\left|\ln \rho_{n}\right|}\right] \tag{7.95}
\end{align*}
$$

where

$$
\begin{aligned}
B_{\rho_{n}}:=\frac{1-\rho_{n}}{\rho_{n}\left|\ln \rho_{n}\right|} \mathbb{E}[ & \ln \left(1-\rho_{n}\right. \\
& +\rho_{n} \sum_{i=1}^{K} e^{\left.\left.-\frac{\gamma \tau}{2 v_{k}^{2}\left|\ln \rho_{n}\right|}\left(p_{i}^{+} e^{\sqrt{\gamma \tau\left|\ln \rho_{n}\right| v_{i}^{2}} Z}+p_{i}^{-} e^{-\sqrt{\gamma \tau\left|\ln \rho_{n}\right| v_{i}^{2}} Z}\right)\right)\right]}
\end{aligned}
$$

and $\forall z \in \mathbb{R}$ :

$$
\begin{align*}
& \widetilde{h}\left(z, \gamma \tau\left|\ln \rho_{n}\right|, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{ \pm}, \mathbf{p}^{\mp}\right):=\left(1-\rho_{n}\right) e^{\left|\ln \rho_{n}\right|\left(1-\frac{\gamma \tau v_{j}^{2}}{2}-\sqrt{\frac{\gamma \tau v v_{j}^{2}}{\left|\ln \rho_{n}\right|}} z\right)} \\
& \quad+\sum_{i=1}^{K} e^{-\left|\ln \rho_{n}\right|\left(\frac{\gamma \tau\left(v_{i}-v_{j}\right)^{2}}{2}-\sqrt{\frac{\gamma \tau}{\ln \rho_{n} \mid}\left(v_{i}-v_{j}\right) z}\right)\left(p_{i}^{ \pm}+p_{i}^{\mp} e^{-2\left|\ln \rho_{n}\right| v_{i}\left(\gamma \tau v_{j}+z \sqrt{\gamma \tau /\left|\ln \rho_{n}\right|}\right.}\right) .} . \tag{7.96}
\end{align*}
$$

We can show, exactly as it is done for $A_{\rho_{n}}$ in the proof of Lemma 7.16, that $\left|B_{\rho_{n}}\right| \leq 1 /\left|\ln \rho_{n}\right|$. Besides, as $\tau<2 / \gamma v_{K}^{2}$, we have $1-\gamma \tau v_{j}^{2} / 2>0$ for all $j \in\{1, \ldots, K\}$ and we easily deduce from (7.96) that $\forall j \in\{1, \ldots, K\}, \forall z \in \mathbb{R}$ :

$$
\begin{equation*}
\lim _{\rho_{n} \rightarrow 0^{+}} \frac{\ln \widetilde{h}\left(z, \gamma \tau\left|\ln \rho_{n}\right|, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{ \pm}, \mathbf{p}^{\mp}\right)}{\left|\ln \rho_{n}\right|}=1-\frac{\gamma \tau v_{j}^{2}}{2} . \tag{7.97}
\end{equation*}
$$

By the dominated convergence theorem, making use of the pointwise limits (7.97),

$$
\begin{aligned}
& \lim _{\rho_{n} \rightarrow 0^{+}} \sum_{j=1}^{K} p_{j}^{+} \mathbb{E}\left[\frac{\ln \widetilde{h}\left(Z, \gamma \tau\left|\ln \rho_{n}\right|, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{+}, \mathbf{p}^{-}\right)}{\left|\ln \rho_{n}\right|}\right] \\
& +p_{j}^{-} \mathbb{E}\left[\frac{\ln \widetilde{h}\left(Z, \gamma \tau\left|\ln \rho_{n}\right|, v_{j} ; \rho_{n}, \mathbf{v}, \mathbf{p}^{-}, \mathbf{p}^{+}\right)}{\left|\ln \rho_{n}\right|}\right] \\
& \quad=\sum_{j=1}^{K}\left(p_{j}^{+}+p_{j}^{-}\right)\left(1-\frac{\gamma \tau v_{j}^{2}}{2}\right)=1-\frac{\gamma \tau \mathbb{E} X_{0}^{2}}{2} .
\end{aligned}
$$

Combining the identity (7.95) with the latter limit and $\lim _{\rho_{n} \rightarrow 0^{+}} B_{\rho_{n}}=0$ yields

$$
\begin{equation*}
\lim _{\rho_{n} \rightarrow 0} \frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} \tau\right)=\frac{\tau \mathbb{E} X_{0}^{2}}{2}-\frac{1}{\gamma}+\frac{1}{\gamma}\left(1-\frac{\gamma \tau \mathbb{E} X_{0}^{2}}{2}\right)=0 . \tag{7.98}
\end{equation*}
$$

Then, the lower bound (7.94) together with (7.98) and $\lim _{\rho_{n} \rightarrow 0^{+}} a_{\rho_{n}}^{(K)}=0$ (see Lemma 7.15 implies that

$$
\liminf _{\rho_{n} \rightarrow 0^{+}} \inf _{q \in\left[0, a_{\rho n}^{(K)}\right]} \widetilde{I}_{P_{\text {out }}}\left(q, \mathbb{E} X_{0}^{2}\right)+\sup _{r \geq \tau}\left\{\frac{r q}{2}-\frac{1}{\alpha_{n}} \psi_{P_{X}^{(n)}}\left(\frac{\alpha_{n}}{\rho_{n}} r\right)\right\} \geq \widetilde{I}_{P_{\text {out }}}\left(0, \mathbb{E} X_{0}^{2}\right) .
$$

Finally, we combine the latter inequality with the limit (7.92) to obtain

$$
\liminf _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}, \tau\right) \geq \min _{1 \leq k \leq K+1} \widetilde{I}_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} .
$$

This lower bound on the limit inferior matches the upper bound on the limit superior in (7.93). Hence,

$$
\begin{aligned}
& \lim _{\rho_{n} \rightarrow 0^{+}} I\left(\rho_{n}, \alpha_{n}, \tau\right)=\min _{1 \leq k \leq K+1} \widetilde{I}_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} \\
& \quad=\min _{1 \leq k \leq K+1} I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\tau \mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k}\right\}}\right]}{2}+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma},
\end{aligned}
$$

where the last equality follows simply from the definition of $\widetilde{I}_{P_{\text {out }}}$.
The next theorem is a direct corollary of Theorem 7.19 and Proposition 7.20.
Theorem 7.21. Suppose that $\Delta>0$ and $P_{X}^{(n)}:=\left(1-\rho_{n}\right) \delta_{0}+\rho_{n} P_{0}$ where $P_{0}$ is a discrete probability distribution with finite support

$$
\operatorname{supp}\left(P_{0}\right) \subseteq\left\{-v_{K},-v_{K-1}, \ldots,-v_{2},-v_{1}, v_{1}, v_{2}, \ldots, v_{K-1}, v_{K}\right\}
$$

where $0<v_{1}<v_{2}<\cdots<v_{K}<v_{K+1}=+\infty$. Further assume that the following hypotheses hold:
(H2) $\varphi$ is bounded, and its first and second partial derivatives with respect to its first argument exist, are bounded and continuous. They are denoted $\partial_{x} \varphi$, $\partial_{x x} \varphi$.
(H3) $W_{\mu i} \stackrel{i . i . d .}{\sim} \mathcal{N}(0,1)$ for $(\mu, i) \in\left\{1, \ldots, m_{n}\right\} \times\{1, \ldots, n\}$.
Let $\rho_{n}=\Theta\left(n^{-\lambda}\right)$ with $\lambda \in(0,1 / 9)$ and $\alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ with $\gamma>0$. Then, $\forall \tau \in\left[0,2 / \gamma v_{K}^{2}\right):$

$$
\begin{aligned}
\lim _{n \rightarrow+\infty} & \frac{I\left(\mathbf{X}^{*} ; \mathbf{Y}, \tilde{\mathbf{Y}}^{(\tau)} \mid \mathbf{W}\right)}{m_{n}} \\
& =\min _{1 \leq k \leq K+1} I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\tau \mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k}\right\}}\right]}{2}+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} .
\end{aligned}
$$

## 7.C. 2 Proof of Theorem 7.3

For all $n \in \mathbb{N}^{*}$ and $\tau \in[0,+\infty)$ we denote by $i_{n}(\tau)$ the normalized conditional mutual information between $\mathbf{X}^{*}$ and the observations $\mathbf{Y}, \widetilde{\mathbf{Y}}^{(\tau)}$ (defined in 7.84 ) $)$ given $\mathbf{W}$,

$$
i_{n}(\tau):=\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y}, \widetilde{\mathbf{Y}}^{(\tau)} \mid \mathbf{W}\right)}{m_{n}}
$$

We place ourselves in the regime of Theorem 7.3, that is, $\rho_{n}=\Theta\left(n^{-\lambda}\right)$ with $\lambda \in[0,1 / 9)$ and $\alpha_{n}=\gamma \rho_{n}\left|\ln \rho_{n}\right|$ with $\gamma>0$. By Theorem 7.21, if the side information is low enough, that is, $\tau<2 / \gamma v_{K}^{2}$, then $\lim _{n \rightarrow+\infty} i_{n}(\tau)=i(\tau)$ where

$$
\begin{equation*}
i(\tau):=\min _{1 \leq k \leq K+1} I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\tau \mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k}\right\}}\right]}{2}+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} . \tag{7.99}
\end{equation*}
$$

We first establish a few properties of the function $i_{n}$. The posterior density of $\mathbf{X}^{*}$ given $\mathbf{W}$ and the observations ( $\mathbf{Y}, \widetilde{\mathbf{Y}}^{(\tau)}$ ) is

$$
\begin{array}{r}
d P\left(\mathbf{x} \mid \mathbf{Y}, \widetilde{\mathbf{Y}}^{(\tau)}, \mathbf{W}\right):=\frac{1}{\mathcal{Z}\left(\mathbf{Y}, \tilde{\mathbf{Y}}^{(\tau)}, \mathbf{W}\right)} \prod_{i=1}^{n} d P_{X}^{(n)}\left(x_{i}\right) e^{-\frac{1}{2}\left(\tilde{Y}_{i}^{(\tau)}-\sqrt{\frac{\alpha_{n} \tau}{\rho_{n}}} x_{i}\right)^{2}} \\
\cdot \prod_{\mu=1}^{m_{n}} P_{\text {out }}\left(Y_{\mu} \left\lvert\, \frac{(\mathbf{W} \mathbf{x})_{\mu}}{\sqrt{k_{n}}}\right.\right), \tag{7.100}
\end{array}
$$

where $\mathcal{Z}\left(\mathbf{Y}, \tilde{\mathbf{Y}}^{(\tau)}, \mathbf{W}\right)$ is a normalization factor. In what follows, $\mathbf{x}$ denotes a $n$-dimensional random vector distributed with respect to the posterior distribution (7.100). We denote by angular brackets $\langle-\rangle_{n, \tau}$ the expectation with respect to this posterior. By definition of the mutual information we have

$$
\left.\left.\left.\begin{array}{rl}
i_{n}(\tau)= & -\frac{\mathbb{E} \ln \mathcal{Z}\left(\mathbf{Y}, \tilde{\mathbf{Y}}^{(\tau)}, \mathbf{W}\right)}{m_{n}} \\
& +\frac{1}{m_{n}} \mathbb{E}\left[\ln \prod_{i=1}^{n} e^{-\frac{1}{2}\left(\widetilde{Y}_{i}^{(\tau)}\right.}-\sqrt{\frac{\alpha_{n} \tau}{\rho n}} X_{i}^{*}\right.
\end{array}\right)^{2} \prod_{\mu=1}^{m_{n}} P_{\text {out }}\left(Y_{\mu} \left\lvert\, \frac{\left(\mathbf{W X}^{*}\right)_{\mu}}{\sqrt{k_{n}}}\right.\right)\right] .\right]\left(\frac{\mathbb{E} \ln \mathcal{Z}\left(\mathbf{Y}, \widetilde{\mathbf{Y}}^{(\tau)}, \mathbf{W}\right)}{m_{n}}-\frac{1}{2 \alpha_{n}}+\mathbb{E}\left[\ln P_{\text {out }}\left(Y_{1} \left\lvert\, \frac{\left(\mathbf{W} \mathbf{X}^{*}\right)_{1}}{\sqrt{k_{n}}}\right.\right)\right] .\right.
$$

Derivation under the expectation sign, justified by the dominated convergence theorem, yields the first derivative of $i_{n}$. We obtain

$$
\begin{align*}
i_{n}^{\prime}(\tau) & =\frac{1}{m_{n}} \sum_{i=1}^{n} \mathbb{E}\left[\left\langle\left(\widetilde{Y}_{i}^{(\tau)}-\sqrt{\frac{\alpha_{n} \tau}{\rho_{n} \tau}} x_{i}\right) \frac{1}{2} \sqrt{\frac{\alpha_{n}}{\rho_{n} \tau}}\left(X_{i}^{*}-x_{i}\right)\right\rangle_{n, \tau}\right] \\
& =\frac{1}{m_{n}} \sum_{i=1}^{n} \mathbb{E}\left[\left\langle\left(\widetilde{Y}_{i}^{(\tau)}-\sqrt{\frac{\alpha_{n} \tau}{\rho_{n} \tau}} X_{i}^{*}\right) \frac{1}{2} \sqrt{\frac{\alpha_{n}}{\rho_{n} \tau}}\left(x_{i}-X_{i}^{*}\right)\right\rangle_{n, \tau}\right] \\
& =\frac{1}{2 m_{n}} \sqrt{\frac{\alpha_{n}}{\rho_{n} \tau}} \mathbb{E}\left[\widetilde{Z}_{i}\left(\left\langle x_{i}\right\rangle_{n, \tau}-X_{i}^{*}\right)\right] \\
& =\frac{1}{2 m_{n}} \sqrt{\frac{\alpha_{n}}{\rho_{n} \tau}} \sum_{i=1}^{n} \mathbb{E}\left[\widetilde{Z}_{i}\left\langle x_{i}\right\rangle_{n, \tau}\right] \\
& =\frac{1}{2 m_{n}} \frac{\alpha_{n}}{\rho_{n}} \sum_{i=1}^{n} \mathbb{E}\left[\left\langle x_{i}^{2}\right\rangle_{n, \tau}-\left\langle x_{i}\right\rangle_{n, \tau}^{2}\right] \\
& =\frac{\mathbb{E}\left\|\mathbf{X}^{*}-\mathbb{E}\left[\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{Y}^{(\tau)}, \mathbf{W}\right]\right\|^{2}}{2 k_{n}} . \tag{7.101}
\end{align*}
$$

The second equality is due to the Nishimori identity, the third to the definition of $\widetilde{Y}_{i}^{(\tau)}$, and the second-to-last to a Gaussian integration by parts with respect to $\widetilde{Z}_{i}$. The final identity $(7.101)$ is the I-MMSE relationship given without proof at the beginning of this appendix. Further differentiating with respect to $\tau$, and integrating by parts with respect to the standard Gaussian random variables $\widetilde{Z}_{i}$, gives

$$
i_{n}^{\prime \prime}(\tau)=-\frac{1}{2 k_{n}} \sum_{i=1}^{n} \mathbb{E}\left[\left\langle\left(x_{i}-\left\langle x_{i}\right\rangle_{n, \tau}\right)^{2}\right\rangle_{n, \tau}^{2}\right]
$$

The latter shows that $i_{n}$ is concave as its second derivative is nonpositive. We have shown that $\left(i_{n}\right)_{n \in \mathbb{N}^{*}}$ is a sequence of continuously differentiable concave functions on $\left[0,{ }^{2} / \gamma v_{K}^{2}\right)$ that converges pointwise to $i$ defined by (7.99). By Griffiths' lemma [52. Appendix A], if the pointwise limit (7.99) is differentiable at $\tau \in\left[0,2 / \gamma v_{K}^{2}\right)$ then

$$
\begin{equation*}
\lim _{n \rightarrow+\infty} i_{n}^{\prime}(\tau)=i^{\prime}(\tau) \tag{7.102}
\end{equation*}
$$

The final step is to determine $i^{\prime}(\tau)$. Suppose that the minimization problem

$$
\begin{equation*}
\min _{1 \leq k \leq K+1} I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma} \tag{7.103}
\end{equation*}
$$

has a unique solution $k^{*} \in\{1, \ldots, K+1\}$. Then, there exists $\epsilon \in\left[0,2 / \gamma v_{K}^{2}\right)$ such that, for all $\tau \in[0, \epsilon), k^{*}$ is the unique solution to the minimization problem

$$
\min _{1 \leq k \leq K+1} I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\tau \mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k}\right\}}\right]}{2}+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)}{\gamma}
$$

Therefore, $\forall \tau \in[0, \epsilon)$ :

$$
\begin{aligned}
i(\tau) & =I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \geq v_{k^{*}}\right\}}\right], \mathbb{E} X_{0}^{2}\right)+\frac{\tau \mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k^{*}}\right\}}\right]}{2}+\frac{\mathbb{P}\left(\left|X_{0}\right| \geq v_{k^{*}}\right)}{\gamma}, \\
i^{\prime}(\tau) & =\frac{\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{\left.k^{*}\right\}}\right\}}\right]}{2} .
\end{aligned}
$$

Combining the latter with (7.101) and (7.102) yields $\forall \tau \in[0, \epsilon)$ :

$$
\lim _{n \rightarrow+\infty} \frac{\mathbb{E}\left\|\mathbf{X}^{*}-\mathbb{E}\left[\mathbf{X}^{*} \mid \mathbf{Y}, \tilde{\mathbf{Y}}^{(\tau)}, \mathbf{W}\right]\right\|^{2}}{k_{n}}=\lim _{n \rightarrow+\infty} 2 i_{n}^{\prime}(\tau)=2 i^{\prime}(\tau)=\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k^{*}}\right\}}\right]
$$

In particular, at $\tau=0$, we have

$$
\lim _{n \rightarrow+\infty} \frac{\mathbb{E}\left\|\mathbf{X}^{*}-\mathbb{E}\left[\mathbf{X}^{*} \mid \mathbf{Y}, \mathbf{W}\right]\right\|^{2}}{k_{n}}=\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k^{*}}\right\}}\right]
$$

whenever the minimization problem (7.103) has a unique solution $k^{*}$.

## 7.C. 3 All-or-nothing phenomenon and its generalization

We now look at the asymptotic MMSE as a function of the number of measurements, i.e., as a function of the parameter $\gamma$ that controls the number of measurements $m_{n}=\gamma \cdot n \rho_{n}\left|\log \rho_{n}\right|$. Let $X_{0} \sim P_{0}$ and assume that supp $\left|X_{0}\right|=K$. We place ourselves under the assumptions of Theorem 7.3. The functions $k \mapsto I_{P_{\text {out }}}\left(\mathbb{E}\left[X_{0}^{2} 1_{\left\{\left|X_{0}\right| \geq v_{k}\right\}}\right], \mathbb{E}\left[X_{0}^{2}\right]\right)$ and $k \mapsto \mathbb{P}\left(\left|X_{0}\right| \geq v_{k}\right)$ are nondecreasing and increasing on $\{1,2, \ldots, K+1\}$, respectively. Hence, the minimization problem on the right-hand side of (7.8) has a unique solution denoted $k^{*}(\gamma)$ for all but $K$ or less values of $\gamma \in(0,+\infty)$, and $\gamma_{1}<\gamma_{2} \Rightarrow k^{*}\left(\gamma_{1}\right) \geq k^{*}\left(\gamma_{2}\right)$ (assuming $k^{*}\left(\gamma_{1}\right), k^{*}\left(\gamma_{2}\right)$ are well-defined). By Theorem 7.3, it implies that the asymptotic MMSE as a function of $\gamma$ is nonincreasing and piecewise constant; its image is included in $\left\{\mathbb{E} X_{0}^{2}, \mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \leq v_{K-1}\right\}}\right], \ldots, \mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right| \leq v_{1}\right\}}\right], 0\right\}$. The asymptotic MMSE has at most $K$ discontinuities. As $\gamma$ increases past a discontinuity, the asymptotic MMSE jumps from $\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k_{1}^{*}}\right\}}\right]$ for some $k_{1}^{*} \in\{2, \ldots, K+1\}$ down to a lower value $\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k_{2}^{*}}\right\}}\right]$ where $k_{2}^{*} \in\left\{1, \ldots, k_{1}^{*}-1\right\}$.

Therefore, when $K=1$, the asymptotic MMSE has one discontinuity at $\gamma_{c}:=1 / I_{P_{\text {out }}}\left(0, \mathbb{E} X_{0}^{2}\right)$ where it jumps down from $\mathbb{E} X_{0}^{2}$ to 0 : this is the all-ornothing phenomenon previously observed in [32], [33], [154] for a linear activation function $\varphi(x)=x$ and a deterministic distribution $P_{0}$. Theorem 7.3 generalizes this all-or-nothing phenomenon to activation functions satisfying mild conditions and any discrete distribution $P_{0}$ whose support is included in $\{-v, v\}$ for some $v>0$.

When $K>1$, the phenomenology is more complex. The asymptotic MMSE exhibits intermediate plateaus between the plateaus "MMSE $=\mathbb{E} X_{0}^{2}$ " (no reconstruction at all) for low values of $\gamma$ and "MMSE $=0$ " (perfect reconstruction) for large values of $\gamma$. For illustration purposes, we now define the following three discrete distributions with support size $K \geq 1$ and unit second moment:

- $P_{\text {unif }}^{(K)}$ is the uniform distribution on $\{\sqrt{a}, 2 \sqrt{a}, \ldots, K \sqrt{a}\}$ with

$$
a:=\frac{6}{(K+1)(2 K+1)} .
$$

- $P_{\text {linear }}^{(K)}$ is the distribution on $\{\sqrt{b}, 2 \sqrt{b}, \ldots, K \sqrt{b}\}$ with

$$
b:=\sum_{j=1}^{K} \frac{1}{K j^{2}} \quad \text { and } \quad P_{\text {linear }}^{(K)}(i \sqrt{b}):=\frac{1}{K i^{2} b},
$$

so that $\mathbb{E} X_{0}^{2}=1$ and $\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<k \sqrt{b}\right\}}\right]={ }_{(k-1) / K}$ where $X_{0} \sim P_{\text {linear }}^{(K)}$, i.e., the quantity $\mathbb{E}\left[X_{0}^{2} \mathbf{1}_{\left\{\left|X_{0}\right|<v_{k}\right\}}\right]$ increases linearly with $k$.

- $P_{\text {binom }}^{(K, p)}$ is the binomial distribution on $\{\sqrt{c}, 2 \sqrt{c}, \ldots, K \sqrt{c}\}$ with

$$
c:=\frac{1}{(K-1)(K-2) p^{2}+3(K-1) p+1}
$$



Figure 7.C.1: Minimum mean-square error in the asymptotic regime of Theorem 7.3 for $\Delta \in[0,4]$ and $\gamma \in(0,10.5]$. From left to right: the activation function is linear $\varphi(x)=x$, the $\operatorname{ReLU} \varphi(x)=\max (0, x)$ and the sign function $\varphi(x)=\operatorname{sign}(x)$. Top to bottom: the prior distribution $P_{0}$ of the nonzero elements of $\mathbf{X}^{*}$ is $P_{\text {unif }}^{(5)}, P_{\text {linear }}^{(5)}$ and $P_{\text {binom }}^{(5,0.2)}$.
and

$$
P_{\text {binom }}^{(K, p)}(i \sqrt{c}):=\binom{K-1}{i-1} p^{i-1}(1-p)^{K-i} .
$$

In Figure 7.C.1 we plot the asymptotic MMSE (using Theorem 7.3) as a function of the noise variance $\Delta$ and the parameter $\gamma$ for three different activation functions and $P_{0} \in\left\{P_{\text {unif }}^{(5)}, P_{\text {linear }}^{(5)}, P_{\text {binom }}^{(5,0.2)}\right\}$.

## 7.D Properties of the mutual information $I_{P_{\text {out }}}$

Lemma 7.22. Let $\Delta$ be a positive real number, $k_{A}$ a nonnegative integer, $P_{A} a$ probability distribution over $\mathbb{R}^{k_{A}}$ and $\varphi: \mathbb{R} \times \mathbb{R}^{k_{A}} \rightarrow \mathbb{R}$ be a bounded measurable function. Further assume that the first and second partial derivatives of $\varphi$ with respect to its first argument, denoted $\partial_{x} \varphi$ and $\partial_{x x} \varphi$, exist and are bounded.
Let $U, V, Z \sim \mathcal{N}(0,1)$ and $\mathbf{A} \sim P_{A}$ be independent random variables. Define $I_{P_{\text {out }}}(q, \rho):=I\left(U ; \widetilde{Y}^{(q, \rho)} \mid V\right)$ the conditional mutual information between $U$ and $\tilde{Y}^{(q, \rho)}:=\varphi(\sqrt{\rho-q} U+\sqrt{q} V, \mathbf{A})+\sqrt{\Delta} Z$ given $V$. Then,

- $\forall \rho \in(0,+\infty): q \mapsto I_{P_{\text {out }}}(q, \rho)$ is continuously twice differentiable, nonincreasing, concave and Lipschitz continuous on $[0, \rho]$ with Lipschitz constant $C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)$, where

$$
C_{1}(a, b):=\left(4 a^{2}+1\right) b^{2} .
$$

- $\forall q \in[0,+\infty)=\rho \mapsto I_{P_{\text {out }}}(q, \rho)$ is Lipschitz continuous on $[q,+\infty)$ with Lipschitz constant $C_{2}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x}}{\sqrt{\Delta}}\right\|_{\infty}\right)$, where

$$
C_{2}(a, b, c):=b^{2}\left(128 a^{4}+12 a^{2}+27\right)+c\left(16 a^{3}+4 \sqrt{2 / \pi}\right) .
$$

Proof. Define $P_{\text {out }}(y \mid x):=\int \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}(y-\varphi(x, \mathbf{a}))^{2}}$. The posterior density function of $U$ given $\left(V, \widetilde{Y}^{(q, \rho)}\right)$ is

$$
\begin{equation*}
d P\left(w \mid V, \widetilde{Y}^{(q, \rho)}\right):=\frac{1}{\mathcal{Z}(q, \rho)} \frac{d u}{\sqrt{2 \pi}} e^{-\frac{u^{2}}{2}} P_{\mathrm{out}}\left(\widetilde{Y}^{(q, \rho)} \mid \sqrt{\rho-q} u+\sqrt{q} V\right) \tag{7.104}
\end{equation*}
$$

where $\mathcal{Z}(q, \rho):=\int \frac{d u}{\sqrt{2 \pi}} e^{-\frac{u^{2}}{2}} P_{\text {out }}\left(\widetilde{Y}^{(q, \rho)} \mid \sqrt{\rho-q} u+\sqrt{q} V\right)$ is the normalization factor. Then,

$$
\begin{align*}
I_{P_{\text {out }}}(q, \rho) & =\mathbb{E}\left[\ln P_{\text {out }}\left(\widetilde{Y}^{(q, \rho)} \mid \sqrt{\rho-q} U+\sqrt{q} V\right)\right]-\mathbb{E} \ln \mathcal{Z}(q, \rho) \\
& =\mathbb{E} \ln \mathcal{Z}(\rho, \rho)-\mathbb{E} \ln \mathcal{Z}(q, \rho) \tag{7.105}
\end{align*}
$$

It is shown in 29, Appendix B.2, Proposition 18] that, for all $\rho \in(0,+\infty)$, $q \mapsto \mathbb{E} \ln \mathcal{Z}(q, \rho)$ is continuously twice differentiable, convex and nondecreasing on $[0, \rho]$, i.e., $q \mapsto I_{P_{\text {out }}}(q, \rho)$ is continuously twice differentiable, concave and nonincreasing on $[0, \rho]$.

We now prove the Lipschitzness of $I_{P_{\text {out }}}(\cdot, \rho)$ by upper bounding the partial derivative of $I_{P_{\text {out }}}$ with respect to $q$. We denote by the angular brackets $\langle-\rangle_{q, \rho}$ the expectation with respect to the posterior distribution (7.104), i.e.,

$$
\langle g(w)\rangle_{q, \rho}:=\int g(w) d P\left(w \mid V, \widetilde{Y}^{(q, \rho)}\right)
$$

Let $\ell_{y}(x):=\ln P_{\text {out }}(y \mid x)$. We know from [29, Appendix B.2, Proposition 18] that $\forall \rho \in(0,+\infty), \forall q \in[0, \rho]:$

$$
\left.\frac{\partial I_{P_{\text {out }}}}{\partial q}\right|_{q, \rho}=-\left.\frac{\partial \mathbb{E} \ln \mathcal{Z}}{\partial q}\right|_{q, \rho}=-\frac{1}{2} \mathbb{E}\left[\left\langle\ell_{\tilde{Y}(q, \rho)}^{\prime}(\sqrt{\rho-q} u+\sqrt{q} V)\right\rangle_{q, \rho}^{2}\right]
$$

By Jensen's inequality and the Nishimory identity, it directly follows that

$$
\begin{align*}
\left.\left|\frac{\partial I_{P_{\text {out }}}}{\partial q}\right|_{q, \rho} \right\rvert\, & \leq \frac{1}{2} \mathbb{E}\left[\left\langle\ell_{\tilde{Y}(q, \rho)}^{\prime}(\sqrt{\rho-q} u+\sqrt{q} V)^{2}\right\rangle_{q, \rho}\right] \\
& =\frac{1}{2} \mathbb{E}\left[\ell_{\tilde{Y}^{(q, \rho)}}^{\prime}(\sqrt{\rho-q} U+\sqrt{q} V)^{2}\right] . \tag{7.106}
\end{align*}
$$

Remember that $\partial_{x} \varphi, \partial_{x x} \varphi$ denote the first and second partial derivatives of $\varphi$ with respect to its first coordinate. The infinity norms $\|\varphi\|_{\infty}$ and $\left\|\partial_{x} \varphi\right\|_{\infty}$ are finite by assumptions. Note that $\forall x \in \mathbb{R}$ :

$$
\begin{align*}
\ell_{y}^{\prime}(x) & =\frac{\int \frac{y-\varphi(x, \mathbf{a})}{\Delta} \partial_{x} \varphi(x, \mathbf{a}) \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}(y-\varphi(x, \mathbf{a}))^{2}}}{\int \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}(y-\varphi(x, \mathbf{a}))^{2}}} ;  \tag{7.107}\\
\left|\ell_{y}^{\prime}(x)\right| & \leq \frac{|y|+\|\varphi\|_{\infty}}{\Delta}\left\|\partial_{x} \varphi\right\|_{\infty} \tag{7.108}
\end{align*}
$$

Thus, $\left|\ell_{\tilde{Y}(q, \rho)}^{\prime}(x)\right| \leq \frac{2\|\varphi\|_{\infty}+\sqrt{\Delta}|Z|}{\Delta}\left\|\partial_{x} \varphi\right\|_{\infty}$. This upper bound combined with 7.106) yields

$$
\begin{equation*}
\left|\frac{\partial I_{P_{\text {out }}}}{\partial q}\right|_{q, \rho} \left\lvert\, \leq \frac{4\|\varphi\|_{\infty}^{2}+\Delta}{\Delta^{2}}\left\|\partial_{x} \varphi\right\|_{\infty}^{2}\right., \tag{7.109}
\end{equation*}
$$

hence $I_{P_{\text {out }}}(\cdot, \rho)$ is Lipschitz continuous with Lipschitz constant ${ }^{4 \| \varphi}\left\|_{\infty}^{2}+\Delta / \Delta^{2}\right\| \partial_{x} \varphi \|_{\infty}^{2}$.
To prove the second point of the lemma, we upper bound the partial derivative of $I_{P_{\text {out }}}$ with respect to $\rho$. Note that

$$
\mathbb{E} \ln \mathcal{Z}(q, \rho)=\mathbb{E}\left[\int d y e^{\ell_{y}(\sqrt{\rho-q} U+\sqrt{q} V)} \ln \left(\int \frac{d u}{\sqrt{2 \pi}} e^{\ell_{y}(\sqrt{\rho-q} u+\sqrt{q} V)-\frac{u^{2}}{2}}\right)\right]
$$

Differentiating the right-hand side under the expectation and integral signs, we get

$$
\begin{aligned}
& \left.\frac{\partial \mathbb{E} \ln \mathcal{Z}}{\partial \rho}\right|_{q, \rho} \\
& =\mathbb{E}\left[\left.\frac{U}{2 \sqrt{\rho-q}} \int d y\left(\ell_{y}^{\prime}(x) e^{\ell_{y}(x)}\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V} \ln \int \frac{d u}{\sqrt{2 \pi}} e^{\ell_{y}(\sqrt{\rho-q} u+\sqrt{q} V)-\frac{u^{2}}{2}}\right] \\
& \quad+\mathbb{E}\left[\left\langle\frac{w}{2 \sqrt{\rho-q}} l_{\tilde{Y}(q, \rho)}^{\prime}(\sqrt{\rho-q} u+\sqrt{q} V)\right\rangle_{q, \rho}\right] \\
& =\mathbb{E}\left[\left.\frac{U}{2 \sqrt{\rho-q}} \int d y\left(\ell_{y}^{\prime}(x) e^{\ell_{y}(x)}\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V} \ln \int \frac{d u}{\sqrt{2 \pi}} e^{\ell_{y}(\sqrt{\rho-q} u+\sqrt{q} V)-\frac{u^{2}}{2}}\right] \\
& \quad+\mathbb{E}\left[\frac{U}{2 \sqrt{\rho-q}} \ell_{\tilde{Y}(q, \rho)}^{\prime}(\sqrt{\rho-q} U+\sqrt{q} V)\right] \\
& =\frac{1}{2} \mathbb{E}\left[\left.\left(\ell_{\widetilde{Y}(q, \rho)}^{\prime \prime}(x)+\ell_{\tilde{Y}(q, \rho)}^{\prime}(x)^{2}\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V} \ln \mathcal{Z}(q, \rho)\right] \\
& \quad+\frac{1}{2} \mathbb{E}\left[\ell_{\tilde{Y}(q, \rho)}^{\prime \prime}(\sqrt{\rho-q} U+\sqrt{q} V)\right]
\end{aligned}
$$

$$
\begin{align*}
&=\frac{1}{2} \mathbb{E}\left[\left.\left(\ell_{\tilde{Y}(q, \rho)}^{\prime \prime}(x)+\ell_{\tilde{Y}(q, \rho)}^{\prime}(x)^{2}\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V}(\ln \mathcal{Z}(q, \rho)+1)\right] \\
&-\frac{1}{2} \mathbb{E}\left[\ell_{\tilde{Y}(q, \rho)}^{\prime}(\sqrt{\rho-q} U+\sqrt{q} V)^{2}\right] . \tag{7.110}
\end{align*}
$$

The second equality is due to the Nishimori identity and the third to a Gaussian integration by parts with respect to $U$. Let us define $\forall \rho \in[0,+\infty)$ :

$$
h(\rho):=\mathbb{E} \ln \mathcal{Z}(\rho, \rho)=\mathbb{E}\left[\int d y e^{\ell_{y}(\sqrt{\rho} V)} \ell_{y}(\sqrt{\rho} V)\right] .
$$

Its derivative is

$$
\begin{align*}
& h^{\prime}(\rho)= \mathbb{E}\left[\frac{V}{2 \sqrt{\rho}} \int d y e^{\ell_{y}(\sqrt{\rho} V)}\left(\ell_{y}(\sqrt{\rho} V)+1\right) \ell_{y}^{\prime}(\sqrt{\rho} V)\right] \\
&=\frac{1}{2} \mathbb{E}\left[\int d y e^{\ell_{y}(\sqrt{\rho} V)}\left(\ell_{y}^{\prime \prime}(\sqrt{\rho} V)+\ell_{y}^{\prime}(\sqrt{\rho} V)^{2}\right)\left(\ell_{y}(\sqrt{\rho} V)+1\right)\right] \\
&+\frac{1}{2} \mathbb{E}\left[\int d y e^{\ell_{y}(\sqrt{\rho} V)} \ell_{y}^{\prime}(\sqrt{\rho} V)^{2}\right] \\
&=\frac{1}{2} \mathbb{E}\left[\left.\left(\ell_{\tilde{Y}^{(\rho, \rho)}}^{\prime \prime}(x)+\ell_{\tilde{Y}^{(\rho, \rho)}}^{\prime}(x)^{2}\right)\right|_{x=\sqrt{\rho} V}(\ln \mathcal{Z}(\rho, \rho)+1)\right] \\
&+\frac{1}{2} \mathbb{E}\left[\ell_{\tilde{Y}^{(\rho, \rho)}}^{\prime}(\sqrt{\rho} V)^{2}\right] . \tag{7.111}
\end{align*}
$$

Differentiating both sides of (7.105 with respect to $\rho$, and using 7.110 and (7.111), yields

$$
\begin{align*}
\left.\frac{\partial I_{P_{\text {out }}}}{\partial \rho}\right|_{q, \rho}= & \frac{1}{2} \mathbb{E}\left[\left.\left(\ell_{\tilde{Y}(\rho, \rho)}^{\prime \prime}(x)+\ell_{\tilde{Y}(\rho, \rho)}^{\prime}(x)^{2}\right)\right|_{x=\sqrt{\rho} V}(\ln \mathcal{Z}(\rho, \rho)+1)\right] \\
& -\frac{1}{2} \mathbb{E}\left[\left.\left(\ell_{\tilde{Y}_{(q, \rho)}^{\prime \prime}}^{\prime \prime}(x)+\ell_{\tilde{Y}^{(q, \rho)}}^{\prime}(x)^{2}\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V}(\ln \mathcal{Z}(q, \rho)+1)\right] \\
& +\frac{1}{2} \mathbb{E}\left[\ell_{\tilde{Y}^{(q, \rho)}}^{\prime}(\sqrt{\rho} V)^{2}\right]+\frac{1}{2} \mathbb{E}\left[\ell_{\tilde{Y}^{(\rho, \rho)}}^{\prime}(\sqrt{\rho-q} U+\sqrt{q} V)^{2}\right] . \tag{7.112}
\end{align*}
$$

The last two summands on the right-hand side of (7.112) are upper bounded by $\frac{4\|\varphi\|_{\infty}^{2}+\Delta}{\Delta^{2}}\left\|\partial_{x} \varphi\right\|_{\infty}^{2}$ (see the proof of the Lipschitzness of $I_{P_{\text {out }}}(\cdot, \rho)$ earlier in this proof). The first two summands on the right-hand side of (7.112) involve the function $(x, y) \mapsto \ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}$. We have

$$
\begin{equation*}
\ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}=\frac{\int \frac{\left((y-\varphi(x, \mathbf{a}))^{2}-\Delta\right) \partial_{x} \varphi(x, \mathbf{a})^{2}+\Delta \partial_{x x} \varphi(x, \mathbf{a})(y-\varphi(x, \mathbf{a}))}{\Delta^{2}} \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}(y-\varphi(x, \mathbf{a}))^{2}}}{\int \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}(y-\varphi(x, \mathbf{a}))^{2}}} . \tag{7.113}
\end{equation*}
$$

Then, by a direct computation, we obtain

$$
\begin{align*}
& \int_{-\infty}^{+\infty}\left(\ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}\right) e^{\ell_{y}(x)} d y \\
& =\int d P_{A}(\mathbf{a}) \int_{-\infty}^{+\infty} \frac{\left((y-\varphi(x, \mathbf{a}))^{2}-\Delta\right) \partial_{x} \varphi(x, \mathbf{a})^{2}+\Delta \partial_{x x} \varphi(x, \mathbf{a})(y-\varphi(x, \mathbf{a}))}{\Delta^{2}} \\
& \quad \cdot \frac{e^{-\frac{(y-\varphi(x, \mathbf{a}))^{2}}{2 \Delta}} d y}{\sqrt{2 \pi \Delta}} \\
& =\int d P_{A}(\mathbf{a}) \int_{-\infty}^{+\infty} \frac{\left(\widetilde{y}^{2}-1\right) \partial_{x} \varphi(x, \mathbf{a})^{2}+\sqrt{\Delta} \partial_{x x} \varphi(x, \mathbf{a}) \widetilde{y}}{\Delta} \frac{e^{-\tilde{y}^{2}}}{\sqrt{2 \pi}} d \widetilde{y} \\
& = \tag{7.114}
\end{align*}
$$

Therefore,

$$
\begin{aligned}
& \mathbb{E}\left[\left.\left(\ell_{\tilde{Y}^{(q, \rho)}}^{\prime \prime}(x)+\ell_{\tilde{Y}(q, \rho)}^{\prime}(x)^{2}\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V}\right] \\
& \quad=\mathbb{E}\left[\left.\left(\int_{-\infty}^{+\infty}\left(\ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}\right) e^{\ell_{y}(x)} d y\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V}\right]=0 .
\end{aligned}
$$

The latter directly implies that

$$
\begin{align*}
& \mathbb{E}\left[\left.\left(\ell_{\tilde{Y}_{(q, \rho)}^{\prime}}^{\prime \prime}(x)+\ell_{\tilde{Y}_{(q, \rho)}^{\prime}}^{\prime}(x)^{2}\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V}(\ln \mathcal{Z}(q, \rho)+1)\right] \\
& =\mathbb{E}\left[\left.\left(\ell_{\tilde{Y}^{(q, \rho)}}^{\prime \prime}(x)+\ell_{\tilde{Y}^{(q, \rho)}}^{\prime}(x)^{2}\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V}\left(\ln \mathcal{Z}(q, \rho)+\frac{\ln (2 \pi \Delta)}{2}\right)\right] . \tag{7.115}
\end{align*}
$$

We now use the identity 7.113 for $\ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}$ to get the upper bound

$$
\begin{align*}
& \mid \ell_{\tilde{Y}(q, \rho)}^{\prime \prime}(x)+\ell_{\tilde{Y}}^{\prime}(q, \rho) \\
& \leq \frac{\left((2 \| \varphi)^{2} \mid\right.}{} \quad \begin{array}{l}
\left.\Delta|Z|)^{2}+\Delta\right)\left\|\partial_{x} \varphi\right\|_{\infty}^{2}+\Delta\left\|\partial_{x x} \varphi\right\|_{\infty}\left(2\|\varphi\|_{\infty}+\sqrt{\Delta}|Z|\right) \\
\Delta^{2}
\end{array} . \tag{7.116}
\end{align*}
$$

We trivially have $P_{\text {out }}(y \mid x) \leq 1 / \sqrt{2 \pi \Delta}$ so

$$
\ln \mathcal{Z}(q, \rho)=\ln \int \frac{d u}{\sqrt{2 \pi}} e^{-\frac{u^{2}}{2}} P_{\mathrm{out}}\left(\widetilde{Y}^{(q, \rho)} \mid \sqrt{\rho-q} u+\sqrt{q} V\right) \leq-\frac{\ln (2 \pi \Delta)}{2}
$$

Besides, by Jensen's inequality,

$$
\begin{aligned}
\ln \mathcal{Z}(q, \rho) & =\ln \int \frac{d u}{\sqrt{2 \pi}} e^{-\frac{u^{2}}{2}} d P_{A}(\mathbf{a}) \frac{1}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2 \Delta}\left(\widetilde{Y}^{(q, \rho)}-\varphi(x, \mathbf{a})\right)^{2}} \\
& \geq \int \frac{d u}{\sqrt{2 \pi}} e^{-\frac{u^{2}}{2}} d P_{A}(\mathbf{a})\left(-\frac{\ln (2 \pi \Delta)}{2}-\frac{\left(\widetilde{Y}^{(q, \rho)}-\varphi(x, \mathbf{a})\right)^{2}}{2 \Delta}\right) \\
& \geq-\frac{\ln (2 \pi \Delta)}{2}-\frac{\left(2\|\varphi\|_{\infty}+\sqrt{\Delta}|Z|\right)^{2}}{2 \Delta} .
\end{aligned}
$$

Hence,

$$
\begin{equation*}
\left|\ln \mathcal{Z}(q, \rho)+\frac{\ln (2 \pi \Delta)}{2}\right| \leq \frac{\left(2\|\varphi\|_{\infty}+\sqrt{\Delta}|Z|\right)^{2}}{2 \Delta} \tag{7.117}
\end{equation*}
$$

Combining (7.115), (7.116), 7.117) yields an upper bound of the second term on the right-hand side of (7.112),

$$
\begin{aligned}
\left|\frac{1}{2} \mathbb{E}\left[\left.\left(\ell_{\tilde{Y}(q, \rho)}^{\prime \prime}(x)+\ell_{\tilde{Y}(q, \rho)}^{\prime}(x)^{2}\right)\right|_{x=\sqrt{\rho-q} U+\sqrt{q} V}(\ln \mathcal{Z}(q, \rho)+1)\right]\right|^{\prime} \\
\leq C\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right),
\end{aligned}
$$

where $C(a, b, c):=b^{2}\left(64 a^{4}+6 a^{2}+13.5\right)+c\left(8 a^{3}+2 \sqrt{\frac{2}{\pi}}\right)$. This upper bound holds for all $q \in[0, \rho]$. In particular, it holds for the first term on the right-hand side of (7.112) where $q=\rho$. We now have an upper bound for each summand on the right-hand side of (7.112) and we combine them to get $\forall \rho \in[q,+\infty)$ :

$$
\left.\frac{\partial I_{P_{\text {out }}}}{\partial \rho}\right|_{q, \rho} \leq 2 C\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)+2\left(4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+1\right)\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}
$$

It concludes the proof of the Lipschitzness of $I_{P_{\text {out }}}(q, \cdot)$ on $[q,+\infty)$.

## 7.E Properties of the interpolating mutual information

Remember that $\ell_{y}(x):=\ln P_{\text {out }}(y \mid x)$ and $\ell_{y}^{\prime}(\cdot), \ell_{y}^{\prime \prime}(\cdot)$ are the first and second derivatives of $\ell_{y}(\cdot)$. We denote by $P_{\text {out }}^{\prime}(y \mid x)$ and $P_{\text {out }}^{\prime \prime}(y \mid x)$ the first and second derivatives of $x \mapsto P_{\text {out }}(y \mid x)$. Finally, we define the scalar overlap $Q:=\frac{1}{k_{n}} \sum_{i=1}^{n} X_{i}^{*} x_{i}$.

## 7.E. 1 Derivative of the interpolating mutual information

Proposition 7.9 (extended version). Suppose that $\Delta>0$, (H1), (H2), (H3) hold, and $\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$. The derivative of the interpolating mutual information (7.44) with respect to $t$ satisfies for all $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\begin{align*}
& i_{n, \epsilon}^{\prime}(t)=O\left(\frac{1}{\sqrt{n \rho_{n}}}\right)+O\left(\sqrt{\frac{\alpha_{n}}{\rho_{n}} \operatorname{Var} \frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}}\right)+\frac{\rho_{n}}{2 \alpha_{n}} r_{\epsilon}(t)\left(1-q_{\epsilon}(t)\right) \\
& \quad+\frac{1}{2} \mathbb{E}\left\langle\left(Q-q_{\epsilon}(t)\right)\left(\frac{1}{m_{n}} \sum_{\mu=1}^{m_{n}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-\frac{\rho_{n}}{\alpha_{n}} r_{\epsilon}(t)\right)\right\rangle_{n, t, \epsilon} \tag{7.118}
\end{align*}
$$

where

$$
\left|O\left(\frac{1}{\sqrt{n \rho_{n}}}\right)\right| \leq \frac{S^{2} C}{\sqrt{n \rho_{n}}} \quad \text { and } \quad\left|O\left(\sqrt{\frac{\alpha_{n}}{\rho_{n}} \mathbb{V a r} \frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}}\right)\right| \leq S^{2} \sqrt{D \frac{\alpha_{n}}{\rho_{n}} \operatorname{Var} \frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}}
$$

with $\left(\partial_{x} \varphi\right.$ and $\partial_{x x} \varphi$ are the first and second partial derivatives of $\varphi$ with respect to its first coordinate)
$C:=\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\left(64\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{4}+2\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+\frac{25}{2}\right)+\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\left(8\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{3}+2 \sqrt{\frac{2}{\pi}}\right)$,
$D:=\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{4}+\frac{1}{2}\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}$.
In addition, if both sequences $\left(\alpha_{n}\right)_{n}$ and $\left(\rho_{n} / \alpha_{n}\right)_{n}$ are bounded, i.e., there exist real positive numbers $M_{\alpha}, M_{\rho / \alpha}$ such that $\forall n \in \mathbb{N}^{*}: \alpha_{n} \leq M_{\alpha}, \rho_{n} / \alpha_{n} \leq M_{\rho / \alpha}$, then for all $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\begin{align*}
& i_{n, \epsilon}^{\prime}(t)=O\left(\frac{1}{\sqrt{n} \rho_{n}}\right)+\frac{\rho_{n}}{2 \alpha_{n}} r_{\epsilon}(t)\left(1-q_{\epsilon}(t)\right) \\
& \quad+\frac{1}{2} \mathbb{E}\left\langle\left(Q-q_{\epsilon}(t)\right)\left(\frac{1}{m_{n}} \sum_{\mu=1}^{m_{n}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-\frac{\rho_{n}}{\alpha_{n}} r_{\epsilon}(t)\right)\right\rangle_{n, t, \epsilon} \tag{7.119}
\end{align*}
$$

where

$$
\left|O\left(\frac{1}{\sqrt{n} \rho_{n}}\right)\right| \leq \frac{S^{2} C+S^{2} \sqrt{D\left(\widetilde{C}_{1}+M_{\rho / \alpha} \widetilde{C}_{2}+M_{\alpha} \widetilde{C}_{3}\right)}}{\sqrt{n} \rho_{n}}
$$

and $\widetilde{C}_{1}, \widetilde{C}_{2}, \widetilde{C}_{3}$ are the polynomials in $\left(S,\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x \varphi}}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x \varphi}}{\sqrt{\Delta}}\right\|_{\infty}\right)$ defined in Proposition 7.23 .

Proof. Remember that $\mathcal{Z}_{t, \epsilon}$ is the normalization to the joint posterior density of $\left(\mathbf{X}^{*}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)$. We define the average interpolating free entropy $f_{n, \epsilon}(t):=\mathbb{E} \ln \mathcal{Z}_{t, \epsilon} / m_{n}$. Note that $i_{n, \epsilon}(t):=I\left(\left(\mathbf{X}^{*}, \mathbf{U}\right) ;\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}\right) \mid \mathbf{W}, \mathbf{V}\right) / m_{n}$ satisfies

$$
\begin{aligned}
i_{n, \epsilon}(t) & =-\frac{\mathbb{E} \ln \mathcal{Z}_{t, \epsilon}}{m_{n}}+\frac{1}{m_{n}} \mathbb{E}\left[\ln \left(e^{-\frac{\|\tilde{\mathbb{Z}}\|^{2}}{2}} P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)\right)\right] \\
& =-f_{n, \epsilon}(t)-\frac{1}{2 \alpha_{n}}+\mathbb{E}\left[\ln P_{\text {out }}\left(Y_{1}^{(t, \epsilon)} \mid S_{1}^{(t, \epsilon)}\right)\right]
\end{aligned}
$$

Given $\mathbf{X}^{*}, S_{1}^{(t, \epsilon)} \sim \mathcal{N}\left(0, V^{(t)}\right)$ where $\rho^{(t)}:=\frac{1-t}{k_{n}}\left\|\mathbf{X}^{*}\right\|^{2}+t+2 s_{n}$. Thus,

$$
\mathbb{E} \ln P_{\text {out }}\left(Y_{1}^{(t, \epsilon)} \mid S_{1}^{(t, \epsilon)}\right)=\mathbb{E}\left[\mathbb{E}\left[\ln P_{\text {out }}\left(Y_{1}^{(t, \epsilon)} \mid S_{1}^{(t, \epsilon)}\right) \mid \mathbf{X}^{*}\right]\right]=\mathbb{E}\left[h\left(\rho^{(t)}\right)\right],
$$

where $h: \rho \in[0,+\infty) \mapsto \mathbb{E}_{V \sim \mathcal{N}(0,1)} \int \ell_{y}(\sqrt{\rho} V) e^{\ell_{y}(\sqrt{\rho} V)} d y$. All in all, we have

$$
\begin{equation*}
i_{n, \epsilon}(t)=\mathbb{E}\left[h\left(\rho^{(t)}\right)\right]-f_{n, \epsilon}(t)-\frac{1}{2 \alpha_{n}} . \tag{7.120}
\end{equation*}
$$

We directly obtain that the derivative of $i_{n, \epsilon}(\cdot)$ is

$$
\begin{equation*}
i_{n, \epsilon}^{\prime}(t)=-\mathbb{E}\left[h^{\prime}\left(\rho^{(t)}\right)\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)\right]-f_{n, \epsilon}^{\prime}(t), \tag{7.121}
\end{equation*}
$$

where $h^{\prime}$ and $f_{n, \epsilon}^{\prime}$ are the derivatives of $h$ and $f_{n, \epsilon}$, respectively. In Lemma 7.22 of Appendix 7.D. we compute $h^{\prime}$ and show that $\forall \rho \in[0,+\infty)$ :

$$
\left|h^{\prime}(\rho)\right| \leq C:=C\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)
$$

where $C(a, b, c):=b^{2}\left(64 a^{4}+2 a^{2}+12.5\right)+c\left(8 a^{3}+2 \sqrt{2 / \pi}\right)$. The first term on the right-hand side of (7.121) thus satisfies

$$
\begin{equation*}
\left|\mathbb{E}\left[h^{\prime}\left(\rho^{(t)}\right)\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)\right]\right| \leq C \sqrt{\operatorname{Var}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}\right)}=\frac{C}{k_{n}} \sqrt{n \mathbb{V} \operatorname{Var}\left(\left(X_{1}^{*}\right)^{2}\right)}=\frac{C S^{2}}{\sqrt{n \rho_{n}}} \tag{7.122}
\end{equation*}
$$

We now turn to the computation of $f_{n, \epsilon}^{\prime}$.
Derivative of the average interpolating free entropy Note that

$$
\begin{equation*}
f_{n, \epsilon}(t)=\frac{1}{m_{n}} \mathbb{E}\left[\int \frac{d \mathbf{y} d \widetilde{\mathbf{y}}}{\sqrt{2 \pi}^{n}} e^{-\mathcal{H} t, \epsilon\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{y}, \tilde{\mathbf{y}}, \mathbf{W}, \mathbf{V}\right)} \ln \int d P_{X}^{(n)}(\mathbf{x}) \mathcal{D} \mathbf{u} e^{-\mathcal{H} t, \epsilon(\mathbf{x}, \mathbf{u}, \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}, \mathbf{V})}\right] \tag{7.123}
\end{equation*}
$$

where the expectation is over $\mathbf{X}^{*}, \mathbf{W}, \mathbf{V}$ and $\mathbf{U}, \mathcal{D} \mathbf{u}:=\frac{d \mathbf{u}}{\sqrt{2 \pi^{m n}}} e^{-\frac{\|\mathbf{u}\|^{2}}{2}}$, and the Hamiltonian $\mathcal{H}_{t, \epsilon}$ is

$$
\mathcal{H}_{t, \epsilon}(\mathbf{x}, \mathbf{u} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}, \mathbf{V}):=-\sum_{\mu=1}^{m_{n}} \ln P_{\text {out }}\left(y_{\mu} \mid s_{\mu}^{(t, \epsilon)}\right)+\frac{1}{2} \sum_{i=1}^{n}\left(\widetilde{y}_{i}-\sqrt{R_{1}(t, \epsilon)} x_{i}\right)^{2}
$$

where

$$
s_{\mu}^{(t, \epsilon)}:=\sqrt{\frac{1-t}{k_{n}}}(\mathbf{W} \mathbf{x})_{\mu}+\sqrt{R_{2}(t, \epsilon)} V_{\mu}+\sqrt{t+2 s_{n}-R_{2}(t, \epsilon)} u_{\mu}
$$

We will need the partial derivative of $\mathcal{H}_{t, \epsilon}$ with respect to $t$, denoted $\mathcal{H}_{t, \epsilon}^{\prime}$,

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}^{\prime}(\mathbf{x}, \mathbf{u} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}, \mathbf{V}):= & -\sum_{\mu=1}^{m_{n}} \frac{\partial s_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{y_{\mu}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right) \\
& -\frac{r_{\epsilon}(t)}{2 \sqrt{R_{1}(t, \epsilon)}} \sum_{i=1}^{n} x_{i}\left(\widetilde{y}_{i}-\sqrt{R_{1}(t, \epsilon)} x_{i}\right) . \tag{7.124}
\end{align*}
$$

Differentiating the right-hand side of (7.123) under the expectation and integral signs, we obtain

$$
\begin{align*}
f_{n, \epsilon}^{\prime}(t)=-\frac{1}{m_{n}} \mathbb{E} & {\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right) \ln \mathcal{Z}_{t, \epsilon}\right] } \\
& -\frac{1}{m_{n}} \mathbb{E}\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{x}, \mathbf{u} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)\right\rangle_{n, t, \epsilon} \\
=-\frac{1}{m_{n}} \mathbb{E}[ & \left.\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right) \ln \mathcal{Z}_{t, \epsilon}\right] \\
& -\frac{1}{m_{n}} \mathbb{E}\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)\right] \tag{7.125}
\end{align*}
$$

The last equality follows from the Nishimory identity

$$
\mathbb{E}\left\langle\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{x}, \mathbf{u} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)\right\rangle_{n, t, \epsilon}=\mathbb{E}\left[\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)\right] .
$$

Evaluating (7.124) at $(\mathbf{x}, \mathbf{u} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}, \mathbf{V})=\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)$ yields

$$
\begin{align*}
\mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)= & -\sum_{\mu=1}^{m_{n}} \frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \\
& -\frac{r_{\epsilon}(t)}{2 \sqrt{R_{1}(t, \epsilon)}} \sum_{i=1}^{n} X_{i}^{*} \widetilde{Z}_{i} . \tag{7.126}
\end{align*}
$$

The expectation of (7.126) is zero. Indeed,

$$
\begin{aligned}
\mathbb{E} \mathcal{H}_{t, \epsilon}^{\prime}\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{Y}^{(t, \epsilon)}, \mathbf{Y}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right) & =-\sum_{\mu=1}^{m_{n}} \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right)\right] \\
& =-\sum_{\mu=1}^{m_{n}} \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \mathbb{E}\left[\ell_{Y_{\mu}^{\prime}(t, \epsilon)}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \mid \mathbf{X}^{*}, \mathbf{U}, \mathbf{V}, \mathbf{W}\right]\right] \\
& =-\sum_{\mu=1}^{m_{n}} \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \int \ell_{y}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) P_{\text {out }}\left(y \mid S_{\mu}^{(t, \epsilon)}\right) d y\right] \\
& =-\sum_{\mu=1}^{m_{n}} \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \int P_{\text {out }}^{\prime}\left(y \mid S_{\mu}^{(t, \epsilon)}\right) d y\right] \\
& =0,
\end{aligned}
$$

where the last equality is because $\forall x \in \mathbb{R}$ :

$$
\int P_{\text {out }}^{\prime}(y \mid x) d y=\int d P_{A}(\mathbf{a}) \partial_{x} \varphi(x, \mathbf{a}) \int \frac{y-\varphi(x, \mathbf{a})}{\Delta} \frac{e^{-\frac{(y-\varphi(x, \mathbf{a}))^{2}}{2 \Delta}}}{\sqrt{2 \pi \Delta}} d y=0 .
$$

The expectation of (7.126) being zero, (7.125) simplifies to

$$
\begin{align*}
f_{n, \epsilon}^{\prime}(t)= & \frac{1}{m_{n}} \sum_{\mu=1}^{m_{n}} \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\right] \\
& +\frac{1}{m_{n}} \frac{r_{\epsilon}(t)}{2 \sqrt{R_{1}(t, \epsilon)}} \sum_{i=1}^{n} \mathbb{E}\left[X_{i}^{*} \widetilde{Z}_{i} \ln \mathcal{Z}_{t, \epsilon}\right] . \tag{7.127}
\end{align*}
$$

Let us compute the first kind of expectation on the right-hand side of (7.127), $\forall \mu \in\left\{1, \ldots, m_{n}\right\}$ :

$$
\begin{align*}
& \mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\right]=-\frac{1}{2} \mathbb{E}\left[\frac{\left(\mathbf{W X}^{*}\right)_{\mu}}{\sqrt{k_{n}(1-t)}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\right] \\
& \quad+\frac{1}{2} \mathbb{E}\left[\left(\frac{q_{\epsilon}(t) V_{\mu}}{\sqrt{R_{2}(t, \epsilon)}}+\frac{\left(1-q_{\epsilon}(t)\right) U_{\mu}}{\sqrt{t+2 s_{n}-R_{2}(t, \epsilon)}}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\right] . \tag{7.128}
\end{align*}
$$

An integration by parts w.r.t. the independent standard Gaussian random variables $\left(W_{\mu i}\right)_{i=1}^{n}$ gives

$$
\begin{align*}
& \mathbb{E}\left[\frac{\left(\mathbf{W X}^{*}\right)_{\mu}}{\sqrt{k_{n}(1-t)}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\right] \\
& =\sum_{i=1}^{n} \mathbb{E}\left[\frac{W_{\mu i} X_{i}^{*}}{\sqrt{k_{n}(1-t)}} \int d \mathbf{y} d \widetilde{\mathbf{y}} \ell_{y_{\mu}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) e^{-\mathcal{H} t, \epsilon\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{y}, \tilde{\mathbf{y}}, \mathbf{W}, \mathbf{V}\right)}\right. \\
& \left.\cdot \ln \left(\int d P_{X}^{(n)}(\mathbf{x}) \mathcal{D} \mathbf{u} e^{-\mathcal{H}_{t, \epsilon}(\mathbf{x}, \mathbf{u} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{w}, \mathbf{V})}\right)\right] \\
& =\sum_{i=1}^{n} \mathbb{E}\left[\frac{\left(X_{i}^{*}\right)^{2}}{k_{n}}\left(\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime \prime}\left(S_{\mu}^{(t, \epsilon)}\right)+\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right)^{2}\right) \ln \mathcal{Z}_{t, \epsilon}\right] \\
& +\sum_{i=1}^{n} \mathbb{E}\left[\frac{X_{i}^{*} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right)}{k_{n}}\left\langle x_{i} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right\rangle_{n, t, \epsilon}\right] \\
& =\mathbb{E}\left[\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)} \ln \mathcal{Z}_{t, \epsilon}\right]+\mathbb{E}\left\langle Q \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right\rangle_{n, t, \epsilon}, \tag{7.129}
\end{align*}
$$

where, in the last equality, we use the identity $\ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}=\frac{P_{\text {out }}^{\prime \prime}(y \mid x)}{P_{\text {out }}(y \mid x)}$. Another Gaussian integration by parts, this time with respect to $V_{\mu} \sim \mathcal{N}(0,1)$, gives

$$
\begin{align*}
& \mathbb{E}\left[\frac{q_{\epsilon}(t) V_{\mu}}{\sqrt{R_{2}(t, \epsilon)}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\right] \\
& =\mathbb{E}\left[\frac{q_{\epsilon}(t) V_{\mu}}{\sqrt{R_{2}(t, \epsilon)}} \int d \mathbf{y} d \widetilde{\mathbf{y}} \ell_{y_{\mu}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) e^{-\mathcal{H}_{t, \epsilon}\left(\mathbf{X}^{*}, \mathbf{U} ; \mathbf{y}, \widetilde{\mathbf{y}}, \mathbf{W}, \mathbf{V}\right)}\right. \\
& \left.\cdot \ln \left(\int d P_{X}^{(n)}(\mathbf{x}) \mathcal{D} \mathbf{u} e^{-\mathcal{H}_{t, \epsilon}(\mathbf{x}, \mathbf{u} \mathbf{:}, \widetilde{\mathbf{y}}, \mathbf{W}, \mathbf{V})}\right)\right] \\
& =\mathbb{E}\left[q_{\epsilon}(t)\left(\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime \prime}\left(S_{\mu}^{(t, \epsilon)}\right)+\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right)^{2}\right) \ln \mathcal{Z}_{t, \epsilon}\right] \\
& +\mathbb{E}\left[q_{\epsilon}(t) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right)\left\langle\ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right\rangle_{n, t, \epsilon}\right] \\
& =\mathbb{E}\left[q_{\epsilon}(t) \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)} \ln \mathcal{Z}_{t, \epsilon}\right]+\mathbb{E}\left\langle q_{\epsilon}(t) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right\rangle_{n, t, \epsilon}, \tag{7.130}
\end{align*}
$$

Finally, a Gaussian integration by part w.r.t. $U_{\mu} \sim \mathcal{N}(0,1)$ gives

$$
\begin{align*}
& \mathbb{E}\left[\frac{\left(1-q_{\epsilon}(t)\right) U_{\mu}}{\sqrt{t+2 s_{n}-R_{2}(t, \epsilon)}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\right] \\
&=\mathbb{E}\left[\left(1-q_{\epsilon}(t)\right) \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)} \ln \mathcal{Z}_{t, \epsilon}\right] \tag{7.131}
\end{align*}
$$

After plugging (7.129), (7.130) and (7.131) back in (7.128), we get

$$
\begin{align*}
\mathbb{E}\left[\frac{\partial S_{\mu}^{(t, \epsilon)}}{\partial t} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ln \mathcal{Z}_{t, \epsilon}\right]= & -\frac{1}{2} \mathbb{E}\left[\frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right) \ln \mathcal{Z}_{t, \epsilon}\right] \\
& -\frac{1}{2} \mathbb{E}\left\langle\left(Q-q_{\epsilon}(t)\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)\right\rangle_{n, t, \epsilon} . \tag{7.132}
\end{align*}
$$

We now compute the second kind of expectation on the right-hand side of (7.127),

$$
\begin{align*}
\mathbb{E}\left[X_{i}^{*} \widetilde{Z}_{i} \ln \mathcal{Z}_{t, \epsilon}\right] & =\mathbb{E}\left[X_{i}^{*} \widetilde{Z}_{i} \ln \int d P_{X}^{(n)}(\mathbf{x}) \mathcal{D} \mathbf{u} P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid s_{\mu}^{(t, \epsilon)}\right) e^{-\sum_{i=1}^{n} \frac{\left(\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{*}-x_{i}\right)+\tilde{Z}_{i}\right)^{2}}{2}}\right] \\
& =-\mathbb{E}\left[X_{i}^{*}\left\langle\sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{*}-x_{i}\right)+\widetilde{Z}_{i}\right\rangle_{n, t, \epsilon}\right] \\
& =-\sqrt{R_{1}(t, \epsilon) \mathbb{E}\left\langle\left(\rho_{n}-X_{i}^{*} x_{i}\right)\right\rangle_{n, t, \epsilon}} \text {, } \tag{7.133}
\end{align*}
$$

where the second equality follows from a Gaussian integration by parts w.r.t. $\widetilde{Z}_{i} \sim \mathcal{N}(0,1)$. Plugging the two simplified expectations (7.132) and 7.133) back in (7.127) yields

$$
\begin{array}{r}
f_{n, \epsilon}^{\prime}(t)=-\frac{\rho_{n}}{2 \alpha_{n}} r_{\epsilon}(t)\left(1-q_{\epsilon}(t)\right)-\frac{1}{2} \mathbb{E}\left[\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right) \frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}\right] \\
-\frac{1}{2} \mathbb{E}\left\langle\left(Q-q_{\epsilon}(t)\right)\left(\frac{1}{m_{n}} \sum_{\mu=1}^{m_{n}} \ell_{Y_{\mu}^{\prime(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{\prime}(t, \epsilon)}^{\prime}\left(s_{\mu}^{(t, \epsilon)}\right)-\frac{\rho_{n}}{\alpha_{n}} r_{\epsilon}(t)\right)\right\rangle_{n, t, \epsilon} . \tag{7.134}
\end{array}
$$

In order to end the proof of the proposition, we have to to upper bound the quantity

$$
\begin{equation*}
A_{n}^{(t, \epsilon)}:=\mathbb{E}\left[\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right) \frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}\right] \tag{7.135}
\end{equation*}
$$

that appears on the right-hand side of (7.134).
Upper bouding the quantity 7.135 Remember that $\ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}=\frac{P_{\text {out }}^{\prime \prime \prime}(y \mid x)}{P_{\text {out }}(y \mid x)}$ and $P_{\text {out }}(y \mid x)=e^{\ell_{y}(x)}$. Therefore, $\forall x \in \mathbb{R}$ :

$$
\int_{-\infty}^{+\infty} P_{\text {out }}^{\prime \prime}(y \mid x) d y=\int_{-\infty}^{+\infty}\left(\ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}\right) e^{\ell_{y}(x)} d y=0
$$

where the second equality is due to the computation (7.114) in the proof of Lemma 7.22, Appendix 7.D. Thus, using the tower property of the conditionnal expectation, we have for all $\mu \in\{1, \ldots, m\}$ :

$$
\begin{aligned}
& \mathbb{E}\left[\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)\right] \\
& \quad=\mathbb{E}\left[\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right) \sum_{\mu=1}^{m_{n}} \mathbb{E}\left[\left.\frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)} \right\rvert\, \mathbf{X}^{*}, \mathbf{S}^{(t, \epsilon)}\right]\right] \\
& \quad=\mathbb{E}\left[\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right) \sum_{\mu=1}^{m_{n}} \int_{-\infty}^{+\infty} P_{\text {out }}^{\prime \prime}\left(y \mid S_{\mu}^{(t, \epsilon)}\right) d y\right]=0
\end{aligned}
$$

The latter identity implies that

$$
\begin{align*}
\left|A_{n}^{(t, \epsilon)}\right| & =\left|\mathbb{E}\left[\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)\left(\frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}-f_{n, \epsilon}(t)\right)\right]\right| \\
& \leq \mathbb{E}\left[\left(\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)^{2}\right]^{\frac{1}{2}} \sqrt{\operatorname{Var} \frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}}, \tag{7.136}
\end{align*}
$$

where the inequality is due to Cauchy-Schwarz inequality. Again by tower property of the conditional expectation, we have

$$
\begin{align*}
& \mathbb{E}\left[\left(\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)^{2}\right] \\
& \quad=\mathbb{E}\left[\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)^{2} \mathbb{E}\left[\left.\left(\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2} \right\rvert\, \mathbf{X}^{*}, \mathbf{S}^{(t, \epsilon)}\right]\right] \tag{7.137}
\end{align*}
$$

 i.i.d. and centered. Therefore,

$$
\begin{align*}
\mathbb{E}\left[\left.\left(\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2} \right\rvert\, \mathbf{X}^{*}, \mathbf{S}^{(t, \epsilon)}\right] & =\mathbb{E}\left[\left.\left(\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2} \right\rvert\, \mathbf{S}^{(t, \epsilon)}\right] \\
& =m_{n} \mathbb{E}\left[\left(\frac{P_{\text {out }}^{\prime \prime}\left(Y_{1}^{(t, \epsilon)} \mid S_{1}^{(t, \epsilon)}\right)}{\left.\left.P_{\text {out }}^{\left(Y_{1}^{(t, \epsilon)} \mid S_{1}^{(t, \epsilon)}\right)}\right)^{2} \mid \mathbf{S}^{(t, \epsilon)}\right]}\right.\right. \\
& =m_{n} \mathbb{E}\left[\int_{-\infty}^{+\infty} \frac{P_{\text {out }}^{\prime \prime}\left(y \mid S_{1}^{(t, \epsilon)}\right)^{2}}{P_{\text {out }}\left(y \mid S_{1}^{(t, \epsilon)}\right)} d y\right] . \tag{7.138}
\end{align*}
$$

We now use the formula (7.113) for $\ell_{y}^{\prime \prime}(x)+\ell_{y}^{\prime}(x)^{2}=P_{\text {out }}^{\prime \prime}(y \mid x) / P_{\text {out }}(y \mid x)$ (see the proof of Lemma 7.22 in Appendix $7 . \mathrm{D})$. By Jensen's equality, $\forall x \in \mathbb{R}$ :

$$
\begin{aligned}
\left(\frac{P_{\text {out }}^{\prime \prime}(y \mid x)}{P_{\text {out }}(y \mid x)}\right)^{2} & \leq \frac{\int\left(\frac{\left((y-\varphi(x, \mathbf{a}))^{2}-\Delta\right) \partial_{x} \varphi(x, \mathbf{a})^{2}+\Delta \partial_{x x} \varphi(x, \mathbf{a})(y-\varphi(x, \mathbf{a}))}{\Delta^{2}}\right)^{2} \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi \Delta}} e^{-\frac{(y-\varphi(x, \mathbf{a}))^{2}}{2 \Delta}}}{\int \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi \Delta}} e^{-\frac{(y-\varphi(x, \mathbf{a}))^{2}}{2 \Delta}}} \\
& =\frac{\int\left(\frac{\left((y-\varphi(x, \mathbf{a}))^{2}-\Delta\right) \partial_{x} \varphi(x, \mathbf{a})^{2}+\Delta \partial_{x x} \varphi(x, \mathbf{a})(y-\varphi(x, \mathbf{a}))}{\Delta^{2}}\right)^{2} \frac{d P_{A}(\mathbf{a})}{\sqrt{2 \pi \Delta}} e^{-\frac{(y-\varphi(x, \mathbf{a}))^{2}}{2 \Delta}}}{P_{\text {out }}(y \mid x)} .
\end{aligned}
$$

It follows that $\forall x \in \mathbb{R}$ :

$$
\begin{aligned}
\int_{-\infty}^{+\infty} \frac{P_{\text {out }}^{\prime \prime}(y \mid x)^{2}}{P_{\text {out }}(y \mid x)} d y & =\int d P_{A}(\mathbf{a}) \int_{-\infty}^{+\infty}\left(\frac{\left(\widetilde{y}^{2}-1\right) \partial_{x} \varphi(x, \mathbf{a})^{2}+\sqrt{\Delta} \partial_{x x} \varphi(x, \mathbf{a}) \widetilde{y}}{\Delta}\right)^{2} \mathcal{D} \widetilde{y} \\
& \leq 4\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{4}+2\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}
\end{aligned}
$$

where $\mathcal{D} \widetilde{y}:=\frac{d \tilde{y}}{\sqrt{2 \pi}} e^{-\frac{\tilde{y}^{2}}{2}}$. Let $D:=\left\|\frac{\partial_{x \varphi}}{\sqrt{\Delta}}\right\|_{\infty}^{4}+\frac{1}{2}\left\|\frac{\partial_{x \varphi}}{\sqrt{\Delta}}\right\|_{\infty}^{2}$. Combining this last upper bound with (7.138) and (7.137) yields

$$
\begin{aligned}
\mathbb{E}\left[\left(\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}{P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid S_{\mu}^{(t, \epsilon)}\right)}\right)^{2}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)^{2}\right] & \\
& \leq 4 D m_{n} \operatorname{Var}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}\right)=\frac{4 D \alpha_{n} S^{4}}{\rho_{n}} .
\end{aligned}
$$

Going back to 7.136), we have $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\begin{equation*}
\left|A_{n}^{(t, \epsilon)}\right| \leq 2 S^{2} \sqrt{D \frac{\alpha_{n}}{\rho_{n}} \mathbb{V a r} \frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}} . \tag{7.139}
\end{equation*}
$$

Putting everything together: proofs of (7.118) and (7.119) Combining (7.121) and (7.134) yields the following formula for the derivative of $i_{n, \epsilon}$ (remember the definition (7.135) of $A_{n}^{(t, \epsilon)}$ ):

$$
\begin{align*}
& i_{n, \epsilon}^{\prime}(t)=\frac{A_{n}^{(t, \epsilon)}}{2}-\mathbb{E}\left[h^{\prime}\left(\rho^{(t)}\right)\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)\right]+\frac{\rho_{n}}{2 \alpha_{n}} r_{\epsilon}(t)\left(1-q_{\epsilon}(t)\right) \\
& +\frac{1}{2} \mathbb{E}\left\langle( Q - q _ { \epsilon } ( t ) ) \left(\frac{1}{m_{n}} \sum_{\mu=1}^{m_{n}} \ell_{Y_{\mu}^{(t, \epsilon)}}^{\prime}\left(S_{\mu}^{(t, \epsilon)}\right) \ell_{Y_{\mu}^{(t, \epsilon)}}^{\left.\left.\prime\left(s_{\mu}^{(t, \epsilon)}\right)-\frac{\rho_{n}}{\alpha_{n}} r_{\epsilon}(t)\right)\right\rangle_{n, t, \epsilon}} .\right.\right. \tag{7.140}
\end{align*}
$$

Thanks to the upper bounds (7.122) and (7.139), we see that the first and second summands on the right-hand side of 7.140$)$ are $\left.O\left(\sqrt{\frac{\alpha_{n}}{\rho_{n}} \operatorname{Var}\left(\ln \mathcal{Z}_{t, \epsilon} / m_{n}\right.}\right)\right)$ and $O\left(1 / \sqrt{n \rho_{n}}\right)$, respectively. It ends the proof of 7.118$)$.

It remains to prove the identity 7.119 that holds under the additional assumption that $\forall n: \alpha_{n} \leq M_{\alpha}, \rho_{n} / \alpha_{n} \leq M_{\rho / \alpha}$. Combining (7.139) with the upper bound 7.154 ) on the variance of $\operatorname{Var}\left(\ln \mathcal{Z}_{t, \epsilon} / m_{n}\right)$ (see Proposition 7.23 in Appendix 7.F gives

$$
\left|\frac{A_{n}^{(t, \epsilon)}}{2}\right| \leq \frac{S^{2} \sqrt{D\left(\widetilde{C}_{1}+M_{\rho / \alpha} \widetilde{C}_{2}+M_{\alpha} \widetilde{C}_{3}\right)}}{\sqrt{n} \rho_{n}}
$$

The constants $\widetilde{C}_{1}, \widetilde{C}_{2}, \widetilde{C}_{3}$ are defined in Proposition 7.23 while $D$ has been defined earlier in the proof. Besides, $1 / \sqrt{n \rho_{n}} \leq 1 / \sqrt{n} \rho_{n}$ as $\rho_{n} \leq 1$ and we can loosen the upper bound (7.122) to

$$
\left|\mathbb{E}\left[h^{\prime}\left(\rho^{(t)}\right)\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1\right)\right]\right| \leq \frac{C S^{2}}{\sqrt{n} \rho_{n}} .
$$

Hence, the term $A_{n}^{(t, \epsilon)} / 2-\mathbb{E}\left[h^{\prime}\left(\rho^{(t)}\right)\left(\left\|\mathbf{X}^{*}\right\|^{2} / k_{n}-1\right)\right]$ on the right-hand side of 7.140) is $O\left(1 / \sqrt{n} \rho_{n}\right)$, proving (7.119).

## 7.E. 2 Proof of Lemma 7.8

Proof. At $t=0$, the functions $r_{\epsilon}$ and $q_{\epsilon}$ do not play any role since $R_{1}(0, \epsilon)=\epsilon_{1}$ and $R_{2}(0, \epsilon)=\epsilon_{2}$. In Appendix 7.A, we restrict $\epsilon:=\left(\epsilon_{1}, \epsilon_{2}\right)$ to be in $\mathcal{B}_{n}:=\left[s_{n}, 2 s_{n}\right]^{2}$. However, nothing prevents us to define observations ( $\left.\mathbf{Y}^{(0, \epsilon)}, \widetilde{\mathbf{Y}}^{(0, \epsilon)}\right)$ using 7.41) with $t=0$ and $\epsilon \in\left[0,2 s_{n}\right]^{2} \supseteq \mathcal{B}_{n}$. We thus extend the interpolating normalized mutual information at $t=0$ to every $\epsilon$ in $\left[0,2 s_{n}\right]^{2} \supseteq \mathcal{B}_{n}$,

$$
i_{n, \epsilon}(0):=\frac{I\left(\left(\mathbf{X}^{*}, \mathbf{U}\right) ;\left(\mathbf{Y}^{(0, \epsilon)}, \tilde{\mathbf{Y}}^{(0, \epsilon)}\right) \mid \mathbf{W}, \mathbf{V}\right)}{m_{n}}
$$

Note that the variation we want to control in this lemma satisfies

$$
\begin{equation*}
\left|i_{n, \epsilon}(0)-\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}\right| \leq\left|i_{n, \epsilon}(0)-i_{n, \epsilon=(0,0)}(0)\right|+\left|i_{n, \epsilon=(0,0)}(0)-\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}\right| \tag{7.141}
\end{equation*}
$$

We upper bound the two terms on the right-hand side of (7.141) separately.

1) By the I-MMSE relationship (see [51]), we have for all $\epsilon \in\left[0,2 s_{n}\right]^{2}$ :

$$
\begin{equation*}
\left|\frac{\partial i_{n, \epsilon}(0)}{\partial \epsilon_{1}}\right|=\frac{1}{2 \alpha_{n}} \mathbb{E}\left[\left(X_{1}^{*}-\left\langle x_{1}\right\rangle_{n, 0, \epsilon}\right)^{2}\right] \leq \frac{\mathbb{E}\left[\left(X_{1}^{*}\right)^{2}\right]}{2 \alpha_{n}}=\frac{\rho_{n}}{2 \alpha_{n}} \tag{7.142}
\end{equation*}
$$

To upper bound the absolute value of the partial derivative with respect to $\epsilon_{2}$, we use that $\forall \epsilon \in\left[0,2 s_{n}\right]^{2}$ :

$$
\frac{\partial i_{n, \epsilon}(0)}{\partial \epsilon_{2}}=-\frac{1}{2} \mathbb{E}\left[\ell_{Y_{1}^{(0, \epsilon)}}^{\prime}\left(S_{1}^{(0, \epsilon)}\right)\left\langle\ell_{Y_{1}^{(0, \epsilon)}}^{\prime}\left(s_{1}^{(0, \epsilon)}\right)\right\rangle_{n, 0, \epsilon}\right]
$$

This identity is obtained in a similar fashion to the computation of the derivative of $i_{n, \epsilon}(\cdot)$ in Appendix 7.E.1 (see 7.130) and 7.131) in particular). Under the hypothesis (H2), we obtain in the proof of Lemma 7.22 the upper bound (7.108) on $\left|\ell_{y}^{\prime}(x)\right|$. Thus, $\forall x \in \mathbb{R}:\left|\ell_{Y_{1}^{(0, \epsilon)}}^{\prime}(x)\right| \leq\left(2\|\varphi\|_{\infty}+\left|Z_{1}\right|\right)\left\|\partial_{x} \varphi\right\|_{\infty}$ and

$$
\begin{equation*}
\left|\frac{\partial i_{n, \epsilon}(0)}{\partial \epsilon_{2}}\right| \leq \frac{1}{2} \mathbb{E}\left[\left(2\|\varphi\|_{\infty}+\left|Z_{1}\right|\right)^{2}\left\|\partial_{x} \varphi\right\|_{\infty}^{2}\right] \leq\left(4\|\varphi\|_{\infty}^{2}+1\right)\left\|\partial_{x} \varphi\right\|_{\infty}^{2} \tag{7.143}
\end{equation*}
$$

By the mean value theorem, and the upper bounds (7.142), (7.143), we have

$$
\begin{align*}
\left|i_{n, \epsilon}(0)-i_{n, \epsilon=(0,0)}(0)\right| & \leq \frac{\rho_{n}}{2 \alpha_{n}}\left|\epsilon_{1}\right|+\left(4\|\varphi\|_{\infty}^{2}+1\right)\left\|\partial_{x} \varphi\right\|_{\infty}^{2}\left|\epsilon_{2}\right| \\
& \leq\left(\frac{\rho_{n}}{2 \alpha_{n}}+\left(4\|\varphi\|_{\infty}^{2}+1\right)\left\|\partial_{x} \varphi\right\|_{\infty}^{2}\right) 2 s_{n} \\
& \leq\left(M_{\rho / \alpha}+2\left(4\|\varphi\|_{\infty}^{2}+1\right)\left\|\partial_{x} \varphi\right\|_{\infty}^{2}\right) s_{n} \tag{7.144}
\end{align*}
$$

2) We now upper bound the second term on the right-hand side of (7.141). Define the following observations where $\mathbf{X}^{*} \stackrel{\text { i.i.d. }}{\sim} P_{X}^{(n)}, \mathbf{W}:=\left(W_{\mu i}\right) \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$, $\mathbf{U}:=\left(U_{\mu}\right)_{\mu=1}^{m_{n}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ and $\eta \in[0,+\infty):$

$$
\begin{equation*}
Y_{\mu}^{(\eta)} \sim P_{\text {out }}\left(\cdot \left\lvert\, \frac{\left(\mathbf{W X}^{*}\right)_{\mu}}{\sqrt{k_{n}}}+\sqrt{\eta} U_{\mu}\right.\right)+Z_{\mu}, 1 \leq \mu \leq m_{n} . \tag{7.145}
\end{equation*}
$$

The joint posterior density of $\left(\mathbf{X}^{*}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{(\eta)}, \mathbf{W}\right)$ is

$$
d P\left(\mathbf{x}, \mathbf{u} \mid \mathbf{Y}^{(\eta)}, \mathbf{W}\right):=\frac{1}{\mathcal{Z}_{\eta}} d P_{X}^{(n)}(\mathbf{x}) \prod_{\mu=1}^{m_{n}} \frac{d u_{\mu}}{\sqrt{2 \pi}} e^{-\frac{u_{\mu}^{2}}{2}} P_{\text {out }}\left(Y_{\mu}^{(\eta)} \left\lvert\, \frac{(\mathbf{W} \mathbf{x})_{\mu}}{\sqrt{k_{n}}}+\sqrt{\eta} u_{\mu}\right.\right)
$$

where $\mathcal{Z}_{\eta}$ is the normalization factor. Define the average free entropy

$$
f_{n}(\eta):=\mathbb{E} \ln \mathcal{Z}_{\rho} / m_{n} .
$$

The normalized mutual information $i_{n}(\eta):=\frac{1}{m_{n}} I\left(\left(\mathbf{X}^{*}, \mathbf{U}\right) ; \mathbf{Y}^{(\eta)} \mid \mathbf{W}\right)$ satisfies

$$
\begin{equation*}
i_{n}(\rho)=\mathbb{E}\left[h\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}+\eta\right)\right]-f_{n}(\rho)-\frac{1}{2 \alpha_{n}} . \tag{7.146}
\end{equation*}
$$

where $h: \rho \in[0,+\infty) \mapsto \mathbb{E}_{V \sim \mathcal{N}(0,1)} \int \ell_{y}(\sqrt{\rho} V) e^{\ell_{y}(\sqrt{\rho} V)} d y$. The identity 7.146) can be obtained exactly as the identity (7.120) in Appendix 7.E.1. Under the assumptions of the lemma, all the hypotheses of domination are reunited to make sure that $\eta \mapsto i_{n}(\eta)$ is continuous on $\left[0,2 s_{n}\right]$ and differentiable on $\left(0,2 s_{n}\right)$. Therefore, by the mean-value theorem, there exists $\eta^{*} \in\left(0,2 s_{n}\right)$ such that:

$$
\begin{equation*}
\left|i_{n, \epsilon=(0,0)}(0)-\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}\right|=\left|i_{n}\left(2 s_{n}\right)-i_{n}(0)\right|=\left|i_{n}^{\prime}\left(\eta^{*}\right)\right| 2 s_{n} . \tag{7.147}
\end{equation*}
$$

Again, in a similar fashion to the computation of the derivative of $i_{n, \epsilon}(\cdot)$ in Appendix 7.E.1. we can show that $\forall \eta \in[0,+\infty)$ :

$$
\begin{align*}
i_{n}^{\prime}(\rho) & =\mathbb{E}\left[h^{\prime}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}+\eta\right)\right]-f_{n}^{\prime}(\rho)  \tag{7.148}\\
f_{n}^{\prime}(\rho) & =\frac{1}{2} \mathbb{E}\left[\sum_{\mu=1}^{m_{n}} \frac{P_{\text {out }}^{\prime \prime}\left(Y_{\mu}^{(\rho)} \left\lvert\, \frac{(\mathbf{W X})^{*}}{\sqrt{k_{n}}}+\sqrt{\eta} U_{\mu}\right.\right)}{P_{\text {out }}\left(Y_{\mu}^{(\rho)} \left\lvert\, \frac{\left.(\mathbf{W X})^{*}\right)_{\mu}}{\sqrt{k_{n}}}+\sqrt{\eta} U_{\mu}\right.\right)} \frac{\ln \mathcal{Z}_{\rho}}{m_{n}}\right] . \tag{7.149}
\end{align*}
$$

In Lemma 7.22 of Appendix 7.D, we compute $h^{\prime}$ and show that $\forall \rho \in[0,+\infty)$ :

$$
\left|h^{\prime}(\rho)\right| \leq C:=C\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)
$$

where $C(a, b, c):=b^{2}\left(64 a^{4}+2 a^{2}+12.5\right)+c\left(8 a^{3}+2 \sqrt{2 / \pi}\right)$. The first term on the right-hand side of (7.148) thus satisfies

$$
\begin{equation*}
\left|\mathbb{E}\left[h^{\prime}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}+\eta\right)\right]\right| \leq C . \tag{7.150}
\end{equation*}
$$

The second term, $f_{n}^{\prime}(\rho)$, is similar to the quantity $A_{n}^{(t, \epsilon)}$ defined in 7.135). We upper bound $A_{n}^{(t, \epsilon)}$ in the proof of Proposition 7.9 in Appendix 7.E.1 We can follow the same steps than for upper bounding $\bar{A}_{n}^{(t, \epsilon)}$ and obtain

$$
\begin{equation*}
\left|f_{n}^{\prime}(\eta)\right| \leq \sqrt{D m_{n} \operatorname{Var} \frac{\ln \mathcal{Z}_{\eta}}{m_{n}}} \tag{7.151}
\end{equation*}
$$

Note that $\mathcal{Z}_{\eta=2 s_{n}}=\mathcal{Z}_{t=0, \epsilon=(0,0)}$. By Proposition 7.23 in Appendix 7.F, we have $\operatorname{Var}\left(\ln \mathcal{Z}_{n=2 s_{n}} / m_{n}\right) \leq \widetilde{C} / n \alpha_{n} \rho_{n}$, where $\widetilde{C}$ is a polynomial in $S,\|\varphi\|_{\infty},\left\|\partial_{x} \varphi\right\|_{\infty}$, $\left\|\partial_{x x} \varphi\right\|_{\infty}, M_{\alpha}, M_{\rho / \alpha}$ with positive coefficients. In fact, this upper bound holds for all $\eta \in\left[0,2 s_{n}\right]$, i.e., $\forall \eta \in\left[0,2 s_{n}\right]$ :

$$
\operatorname{Var}\left(\frac{\ln \mathcal{Z}_{\eta}}{m_{n}}\right) \leq \frac{\widetilde{C}}{n \alpha_{n} \rho_{n}}
$$

The proof of this uniform bound on $\operatorname{Var}\left(\ln \mathcal{Z}_{\eta} / m_{n}\right)$ is the same as the one of Proposition 7.23, only that it is simpler because there is no second channel similar to $\mathbf{Y}^{(t, \epsilon)}$. We now combine (7.147), 7.148, 7.150, (7.151) to finally obtain:

$$
\begin{equation*}
\left|i_{n, \epsilon=(0,0)}(0)-\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}\right| \leq\left(C+\sqrt{\frac{D \widetilde{C}}{\rho_{n}}}\right) 2 s_{n} \tag{7.152}
\end{equation*}
$$

3) To end the proof of the lemma, we just need to plug (7.144, (7.152) back in (7.141) and use that $\rho_{n} \in(0,1]$,

$$
\left|i_{n, \epsilon}(0)-\frac{I\left(\mathbf{X}^{*} ; \mathbf{Y} \mid \mathbf{W}\right)}{m_{n}}\right| \leq\left(M_{\rho / \alpha}+2\left(4\|\varphi\|_{\infty}^{2}+1\right)\left\|\partial_{x} \varphi\right\|_{\infty}^{2}+2 C+\sqrt{D \widetilde{C}}\right) \frac{s_{n}}{\sqrt{\rho_{n}}}
$$

## 7.F Concentration of the free entropy

In this appendix we show that the log-partition function per data point, or free entropy, of the interpolating model studied in Section 7.A.1 concentrates around its expectation.

Proposition 7.23 (Free entropy concentration). Suppose that $\Delta>0$, (H1), (H2), (H3) hold, and $\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$. Then, $\forall(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\begin{equation*}
\operatorname{Var}\left(\frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}\right) \leq \frac{1}{n \alpha_{n} \rho_{n}}\left(\widetilde{C}_{1}+\frac{\rho_{n}}{\alpha_{n}} \widetilde{C}_{2}+\alpha_{n} \widetilde{C}_{3}\right) \tag{7.153}
\end{equation*}
$$

where $\left(\partial_{x} \varphi\right.$ and $\partial_{x x} \varphi$ are the first and second partial derivatives of $\varphi$ with respect to its first coordinate):

$$
\begin{aligned}
& \widetilde{C}_{1}:=1.5+4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+8 S^{2}\left(4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+1\right)\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2} \\
&+\left(2\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+\sqrt{\frac{2}{\pi}}\right)^{2}\left(2\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+\left(16+4 S^{2}\right)\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right) ; \\
& \widetilde{C}_{2}:=1.5+12 S^{2} ; \\
& \widetilde{C}_{3}:=8 S^{2}\left(3\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right. \\
&\left.+12\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+2 \sqrt{\frac{2}{\pi}}\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right)^{2}
\end{aligned}
$$

In addition, if there exist real positive numbers $M_{\alpha}, M_{\rho / \alpha}$ such that $\forall n \in \mathbb{N}^{*}$ : $\alpha_{n} \leq M_{\alpha}, \rho_{n} / \alpha_{n} \leq M_{\rho / \alpha}$ then for all $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\begin{equation*}
\mathbb{V a r}\left(\frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}\right) \leq \frac{C}{n \alpha_{n} \rho_{n}}, \tag{7.154}
\end{equation*}
$$

where $C:=\widetilde{C}_{1}+M_{\rho / \alpha} \widetilde{C}_{2}+M_{\alpha} \widetilde{C}_{3}$.
To lighten notations, we define $k_{1}:=\sqrt{R_{2}(t, \epsilon)}, k_{2}:=\sqrt{t+2 s_{n}-R_{2}(t, \epsilon)}$. Let $\mathbf{X}^{*}:=\left(X_{i}^{*}\right)_{i=1}^{n} \stackrel{\text { i.i.d. }}{\sim} P_{X}^{(n)}, \mathbf{W}:=\left(W_{\mu i}\right) \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1), \mathbf{V}:=\left(V_{\mu}\right)_{\mu=1}^{m_{n}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ and $\mathbf{U}:=\left(U_{\mu}\right)_{\mu=1}^{m_{n}} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ be independent random variables. Remember that

$$
\begin{equation*}
S_{\mu}^{(t, \epsilon)}:=\sqrt{\frac{1-t}{k_{n}}}\left(\mathbf{W X}^{*}\right)_{\mu}+k_{1} V_{\mu}+k_{2} U_{\mu} \tag{7.155}
\end{equation*}
$$

and, at $t \in[0,1]$, we observe

$$
\begin{cases}Y_{\mu}^{(t, \epsilon)} & \sim \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)+\sqrt{\Delta} Z_{\mu},  \tag{7.156}\\ \widetilde{Y}_{i}^{(t, \epsilon)}=\sqrt{R_{1}(t, \epsilon)} X_{i}^{*}+\widetilde{Z}_{i}, & 1 \leq i \leq n\end{cases}
$$

where $\left(Z_{\mu}\right)_{\mu=1}^{m_{n}},\left(\widetilde{Z}_{i}\right)_{i=1}^{n} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ and $\left.\left(\mathbf{A}_{\mu}\right)\right)_{\mu=1}^{m_{n}} \stackrel{\text { i.i.d. }}{\sim} P_{A} . \mathcal{Z}_{t, \epsilon}$ is the normalization to the joint posterior density of $\left(\mathbf{X}^{*}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{(t, \epsilon)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)$, i.e.,

$$
\mathcal{Z}_{t, \epsilon}:=\int d P_{X}^{(n)}(\mathbf{x}) \mathcal{D} \mathbf{u} e^{-\frac{\left\|\sqrt{R_{1}(t, \epsilon)} \mathbf{x}-\tilde{\mathbf{Y}}^{(t, \epsilon)}\right\|^{2}}{2}} P_{\mathrm{out}}\left(Y_{\mu}^{(t, \epsilon)} \mid s_{\mu}^{(t, \epsilon)}\right),
$$

where $\mathcal{D} \mathbf{u}:=\frac{d \mathbf{u} \mathbf{u}^{-\frac{\|u\|^{2}}{\|^{2}}}}{\sqrt{2 \pi^{m n}}}$ and $s_{\mu}^{(t, \epsilon)}:=\sqrt{\frac{1-t}{k_{n}}}(\mathbf{W} \mathbf{x})_{\mu}+k_{1} V_{\mu}+k_{2} u_{\mu}$. We define

$$
\Gamma_{\mu}^{(t, \epsilon)}:=\frac{\varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)-\varphi\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}\right)}{\Delta}
$$

By definition, $P_{\text {out }}\left(Y_{\mu}^{(t, \epsilon)} \mid s_{\mu}^{(t, \epsilon)}\right)=\int d P_{A}\left(\mathbf{a}_{\mu}\right) \frac{1}{\sqrt{2 \pi \Delta}} e^{-\frac{1}{2}\left(\Gamma_{\mu}^{(t, \epsilon)}+Z_{\mu}\right)^{2}}$. Therefore, the interpolating free entropy satisfies

$$
\begin{equation*}
\frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}=\frac{1}{2} \ln (2 \pi \Delta)-\frac{1}{2 m_{n}} \sum_{\mu=1}^{m_{n}} Z_{\mu}^{2}-\frac{1}{2 m_{n}} \sum_{i=1}^{n} \widetilde{Z}_{i}^{2}+\frac{\ln \widehat{\mathcal{Z}}_{t, \epsilon}}{m_{n}} \tag{7.157}
\end{equation*}
$$

where

$$
\begin{equation*}
\widehat{\mathcal{Z}}_{t, \epsilon}:=\int d P_{X}^{(n)}(\mathbf{x}) \mathcal{D} \mathbf{u} d P_{A}\left(\mathbf{a}_{\mu}\right) e^{-\widehat{\mathcal{H}}_{t, \epsilon}(\mathbf{x}, \mathbf{u}, \mathbf{a})} \tag{7.158}
\end{equation*}
$$

with

$$
\begin{align*}
\widehat{\mathcal{H}}_{t, \epsilon}(\mathbf{x}, \mathbf{u}, \mathbf{a}):= & \frac{1}{2} \sum_{\mu=1}^{m_{n}}\left(\Gamma_{\mu}^{(t, \epsilon)}\right)^{2}+2 Z_{\mu} \Gamma_{\mu}^{(t, \epsilon)} \\
& +\frac{1}{2} \sum_{i=1}^{n} R_{1}(t, \epsilon)\left(X_{i}^{*}-x_{i}\right)^{2}+2 Z_{i}^{\prime} \sqrt{R_{1}(t, \epsilon)}\left(X_{i}^{*}-x_{i}\right) . \tag{7.159}
\end{align*}
$$

From (7.157), it directly directly that

$$
\begin{align*}
\operatorname{Var}\left(\frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}\right) & \leq 3 \operatorname{Var}\left(\frac{1}{2 m_{n}} \sum_{\mu=1}^{m_{n}} Z_{\mu}^{2}\right)+3 \operatorname{Var}\left(\frac{1}{2 m_{n}} \sum_{i=1}^{n} \widetilde{Z}_{i}^{2}\right)+3 \operatorname{Var}\left(\frac{\ln \widehat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}\right) \\
& =\frac{3}{2 \alpha_{n} n}+\frac{3}{2 \alpha_{n}^{2} n}+3 \operatorname{Var}\left(\frac{\ln \widehat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}\right) \tag{7.160}
\end{align*}
$$

In order to prove Proposition 7.23 , it remains to show that $\ln \widehat{\mathcal{Z}}_{t, \epsilon} / m_{n}$ concentrates. We first show the concentration w.r.t. all Gaussian variables $\mathbf{W}, \mathbf{V}, \mathbf{Z}, \mathbf{Z}^{\prime}, \mathbf{U}$, then the concentration w.r.t. A and finally the one w.r.t. $\mathbf{X}^{*}$. The order in which we prove the concentrations does matter.

We denote $\partial_{x} \varphi$ and $\partial_{x x} \varphi$ the first and second partial derivatives of $\varphi$ with respect to its first coordinate. Note that $\left|R_{1}\right| \leq 2 s_{n}+\frac{\alpha_{n}}{\rho_{n}} r_{\max }$. Thanks to the inequality (7.109) in Appendix 7.D we obtain

$$
r_{\max }:=2\left|\frac{\partial I_{P_{\text {out }}}}{\partial q}\right|_{q=1, \rho=1} \left\lvert\, \leq 2 C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)\right.
$$

where $C_{1}(a, b):=\left(4 a^{2}+1\right) b^{2}$. Then, the quantity

$$
K_{n}:=2\left(s_{n}+\frac{\alpha_{n}}{\rho_{n}} C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)\right)
$$

upper bounds $\left|R_{1}\right|$. Besides, $\left|R_{2}\right|$ is upper bounded by 2.

## Concentration with respect to the Gaussian random variables

Lemma 7.24. Let $\mathbb{E}_{\mathbf{Z}, \tilde{\mathbf{Z}}}$ be the expectation w.r.t. $(\mathbf{Z}, \widetilde{\mathbf{Z}})$ only. Under the assumptions of Theorem 7.1, we have for all $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\mathbb{E}\left[\left(\frac{\ln \hat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}-\frac{1}{m_{n}} \mathbb{E}_{\mathbf{Z}, \mathbf{Z}^{\prime}} \ln \hat{\mathcal{Z}}_{t, \epsilon}\right)^{2}\right] \leq \frac{C_{2}}{n \alpha_{n} \rho_{n}}+\frac{C_{3}}{n \alpha_{n}^{2}}
$$

where $C_{2}:=4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+8 S^{2} C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)$ and $C_{3}=4 S^{2}$.

Proof. In this proof, we see $g(\mathbf{Z}, \widetilde{\mathbf{Z}}):=\ln \hat{\mathcal{Z}}_{t, \epsilon} / m_{n}$ as a function of $\mathbf{Z}$ and $\widetilde{\mathbf{Z}}$, and we work conditionally on all other random variables. We have

$$
\|\nabla g\|^{2}=\left\|\nabla_{\mathbf{Z}} g\right\|^{2}+\left\|\nabla_{\widetilde{\mathbf{z}}} g\right\|^{2},
$$

where each partial derivative has the form $\partial g / \partial x=-m_{n}^{-1}\left\langle\partial \widehat{\mathcal{H}}_{t, \epsilon} g / \partial x\right\rangle_{t, \epsilon}$. We find that

$$
\begin{aligned}
& \left\|\nabla_{\mathbf{Z}} g\right\|^{2}=m_{n}^{-2} \sum_{\mu=1}^{m_{n}}\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{t, \epsilon}^{2} \leq 4 m_{n}^{-1}\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}, \\
& \left\|\nabla_{\widetilde{\mathbf{Z}}} g\right\|^{2}=m_{n}^{-2} R_{1}(t, \epsilon) \sum_{i=1}^{n}\left(X_{i}^{*}-\left\langle x_{i}\right\rangle_{t, \epsilon}\right)^{2} \leq 4 K_{n} S^{2} m_{n}^{-2} n .
\end{aligned}
$$

Thus, $\|\nabla g\|^{2} \leq 4 m_{n}^{-1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+\frac{K_{n} S^{2}}{\alpha_{n}}\right)$. By the Gaussian-Poincaré inequality (Proposition 2.7), we have

$$
\begin{aligned}
& \mathbb{E}_{\mathbf{Z}, \tilde{\mathbf{Z}}}\left[\left(\frac{\ln \hat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}-\frac{\mathbb{E}_{\mathbf{Z}, \tilde{\mathbf{Z}}} \ln \hat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}\right)^{2}\right] \leq \frac{4}{n \alpha_{n}}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+\frac{K_{n} S^{2}}{\alpha_{n}}\right) \\
&=\frac{4}{n \alpha_{n}}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+\frac{2 S^{2} s_{n}}{\alpha_{n}}+\frac{2 S^{2}}{\rho_{n}} C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)\right) \\
& \leq \frac{4}{n \alpha_{n} \rho_{n}}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+2 S^{2} C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)\right)+\frac{4 S^{2}}{n \alpha_{n}^{2}} .
\end{aligned}
$$

The last inequality follows from $\rho_{n} \leq 1$ and $2 s_{n} \leq 1$. Taking the full expectation on both sides of this last inequality gives the result.

Lemma 7.25. Let $\mathbb{E}_{G}$ be the expectation w.r.t. $(\mathbf{Z}, \widetilde{\mathbf{Z}}, \mathbf{V}, \mathbf{U}, \mathbf{W})$ only. Under the assumptions of Theorem 7.1, we have for all $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\mathbb{E}_{\mathbf{Z}, \tilde{\mathbf{Z}}} \ln \widehat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}-\frac{\mathbb{E}_{G} \ln \widehat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}\right)^{2}\right] \leq \frac{C_{4}}{n \alpha_{n} \rho_{n}}, \tag{7.161}
\end{equation*}
$$

where $C_{4}:=\left(4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+2 \sqrt{2 / \pi}\right)^{2}\left(4+S^{2}\right)\left\|\frac{\partial_{x \varphi}}{\sqrt{\Delta}}\right\|_{\infty}^{2}$.
Proof. In this proof we see $g(\mathbf{V}, \mathbf{U}, \mathbf{W})=\mathbb{E}_{\mathbf{Z}, \widetilde{\mathbf{Z}}} \ln \widehat{\mathcal{Z}}_{t, \epsilon} / m_{n}$ as a function of $\mathbf{V}, \mathbf{U}$, $\mathbf{W}$ and we work conditionally on $\mathbf{A}, \mathbf{X}^{*}$. Each partial derivative of $g$ has the form $\partial g / \partial x=-m_{n}^{-1} \mathbb{E}_{\mathbf{Z}, \tilde{\mathbf{Z}}}\left[\left\langle\partial \widehat{\mathcal{H}}_{t, \epsilon} / \partial x\right\rangle_{t, \epsilon}\right]$. We first compute the partial derivative w.r.t. $V_{\mu}$,

$$
\begin{aligned}
\left|\frac{\partial g}{\partial V_{\mu}}\right| & =m_{n}^{-1}\left|\mathbb{E}_{\mathbf{z}, \tilde{\mathbf{Z}}}\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+Z_{\mu}\right) \frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial V_{\mu}}\right\rangle_{t, \epsilon}\right| \\
& \leq m_{n}^{-1} \mathbb{E}_{\mathbf{z}, \tilde{\mathbf{Z}}}\left[\left(\left(2\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+\left|Z_{\mu}\right|\right) 2 \sqrt{2}\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right]\right. \\
& =m_{n}^{-1}\left(4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+2 \sqrt{\frac{2}{\pi}}\right) \sqrt{2}\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty} .
\end{aligned}
$$

The same inequality holds for $\left|\partial g / \partial U_{\mu}\right|$. To compute the partial derivative w.r.t. $W_{\mu i}$, we first remark that

$$
\begin{aligned}
\frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial W_{\mu i}}=\sqrt{\frac{1-t}{\Delta k_{n}}}\left\{X_{i}^{*} \partial_{x} \varphi( \right. & \left.\sqrt{\frac{1-t}{k_{n}}}\left(\mathbf{W} \mathbf{X}^{*}\right)_{\mu}+k_{1} V_{\mu}+k_{2} U_{\mu}, \mathbf{A}_{\mu}\right) \\
& \left.-x_{i} \partial_{x} \varphi\left(\sqrt{\frac{1-t}{k_{n}}}(\mathbf{W} \mathbf{x})_{\mu}+k_{1} V_{\mu}+k_{2} u_{\mu}, \mathbf{a}_{\mu}\right)\right\}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\left|\frac{\partial g}{\partial W_{\mu i}}\right| & =m_{n}^{-1}\left|\mathbb{E}_{\mathbf{z}, \tilde{\mathbf{Z}}}\left\langle\left(\Gamma_{\mu}^{(t, \epsilon)}+Z_{\mu}\right) \frac{\partial \Gamma_{\mu}^{(t, \epsilon)}}{\partial W_{\mu i}}\right\rangle_{t, \epsilon}\right| \\
& \leq \frac{1}{m_{n} \sqrt{k_{n}}} \mathbb{E}_{\mathbf{z}, \tilde{\mathbf{Z}}}\left[\left(2\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+\left|Z_{\mu}\right|\right) 2 S\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right] \\
& =\frac{1}{m_{n} \sqrt{k_{n}}}\left(4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+2 \sqrt{\frac{2}{\pi}}\right) S\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}
\end{aligned}
$$

Putting together these inequalities on the partial derivatives of $g$, we find that

$$
\begin{aligned}
& \|\nabla g\|^{2}=\sum_{\mu=1}^{m_{n}}\left|\frac{\partial g}{\partial V_{\mu}}\right|^{2}+\sum_{\mu=1}^{m_{n}}\left|\frac{\partial g}{\partial U_{\mu}}\right|^{2}+\sum_{\mu=1}^{m_{n}} \sum_{i=1}^{n}\left|\frac{\partial g}{\partial W_{\mu i}}\right|^{2} \\
& \quad \leq \frac{4}{m_{n}}\left(4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+2 \sqrt{\frac{2}{\pi}}\right)^{2}\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+\frac{1}{m_{n} \rho_{n}}\left(4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+2 \sqrt{\frac{2}{\pi}}\right)^{2} S^{2}\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2} \\
& \\
& \leq \frac{1}{m_{n} \rho_{n}}\left(4\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+2 \sqrt{\frac{2}{\pi}}\right)^{2}\left(4+S^{2}\right)\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}
\end{aligned}
$$

We use $\rho_{n} \leq 1$ to obtain the last inequality. To end the proof of the lemma, we apply the Gaussian-Poincaré inequality (Proposition 2.7) as we did in the proof of Lemma 7.24 .

Concentration with respect to the random stream Next, we use EfronStein inequality (Proposition 2.5) to show that $\mathbb{E}_{G} \ln \widehat{\mathcal{Z}}_{t, \epsilon} / m_{n}$ concentrates w.r.t. A.

Lemma 7.26. Let $\mathbb{E}_{\mathbf{A}}$ be the expectation w.r.t. $\mathbf{A}$ only. Under the assumptions of Theorem 7.1, we have for all $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\begin{equation*}
\mathbb{E}\left[\left(\frac{\mathbb{E}_{G} \ln \widehat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}-\frac{\mathbb{E}_{G, \mathbf{A}} \ln \widehat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}\right)^{2}\right] \leq \frac{C_{5}}{n \alpha_{n}}, \tag{7.162}
\end{equation*}
$$

where $C_{5}:=2\left(2\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+\sqrt{\frac{2}{\pi}}\right)^{2}\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}$.
Proof. We see $g(\mathbf{A})=\mathbb{E}_{G} \ln \widehat{\mathcal{Z}}_{t, \epsilon} / m_{n}$ as a function of $\mathbf{A}$ only. Let $\nu \in\left\{1, \ldots, m_{n}\right\}$. Let $\mathbf{A}:=\left\{\mathbf{A}_{\mu}\right\}_{\mu=1}^{m_{n}}$ and $\mathbf{A}^{(\nu)}:=\left\{\mathbf{A}_{\mu}^{(\nu)}\right\}_{\mu=1}^{m_{n}}$ where $\mathbf{A}_{1}, \mathbf{A}_{2}, \ldots, \mathbf{A}_{m_{n}}, \mathbf{A}_{\nu}^{(\nu)} \stackrel{\text { i.i.d. }}{\sim} P_{A}$
and $A_{\mu}^{(\nu)}:=A_{\mu}$ for $\mu \neq \nu$. We want to bound the difference $g(\mathbf{A})-g\left(\mathbf{A}^{(\nu)}\right)$. We denote by $\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}$ and $\Gamma_{\mu}^{(t, \epsilon)(\nu)}$ the quantities $\widehat{\mathcal{H}}_{t, \epsilon}$ and $\Gamma_{\mu}^{(t, \epsilon)}$ where $\mathbf{A}$ is replaced by $\mathbf{A}^{(\nu)}$. By Jensen's inequality, we have

$$
\begin{equation*}
\frac{1}{m_{n}} \mathbb{E}_{G}\left\langle\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}-\widehat{\mathcal{H}}_{t, \epsilon}\right\rangle_{t, \epsilon}^{(\nu)} \leq g(\mathbf{A})-g\left(\mathbf{A}^{(\nu)}\right) \leq \frac{1}{m_{n}} \mathbb{E}_{G}\left\langle\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}-\widehat{\mathcal{H}}_{t, \epsilon}\right\rangle_{t, \epsilon}, \tag{7.163}
\end{equation*}
$$

where the angular brackets $\langle-\rangle_{t, \epsilon}$ and $\langle-\rangle_{t, \epsilon}^{(\nu)}$ denote the expectation with respect to the probability distributions $\propto d P_{X}^{(n)}(\mathbf{x}) \mathcal{D} \mathbf{u} d P_{A}\left(\mathbf{a}_{\mu}\right) e^{-\widehat{\mathcal{H}}_{t, \epsilon}(\mathbf{x}, \mathbf{u}, \mathbf{a})}$ and $\propto d P_{X}^{(n)}(\mathbf{x}) \mathcal{D} \mathbf{u} d P_{A}\left(\mathbf{a}_{\mu}\right) e^{-\hat{\mathcal{H}}_{t, \epsilon}^{(\nu)}(\mathbf{x}, \mathbf{u}, \mathbf{a})}$, respectively. From the definition 7.159) of $\widehat{\mathcal{H}}_{t, \epsilon}$, we have

$$
\widehat{\mathcal{H}}_{t, \epsilon}^{(\nu)}-\widehat{\mathcal{H}}_{t, \epsilon}=\frac{1}{2}\left(\left(\Gamma_{\nu}^{(t, \epsilon)(\nu)}\right)^{2}-\left(\Gamma_{\nu}^{(t, \epsilon)}\right)^{2}+2 Z_{\nu}\left(\Gamma_{\nu}^{(t, \epsilon)(\nu)}-\Gamma_{\nu}^{(t, \epsilon)}\right)\right) .
$$

Note that

$$
\left|\left(\Gamma_{\nu}^{(t, \epsilon)(\nu)}\right)^{2}-\left(\Gamma_{\nu}^{(t, \epsilon)}\right)^{2}+2 Z_{\nu}\left(\Gamma_{\nu}^{(t, \epsilon)(\nu)}-\Gamma_{\nu}^{(t, \epsilon)}\right)\right| \leq 8\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+4\left|Z_{\nu}\right|\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}
$$

Therefore, $\forall \nu \in\left\{1, \ldots, m_{n}\right\}$ :

$$
\begin{equation*}
\left|g(\mathbf{A})-g\left(\mathbf{A}^{(\nu)}\right)\right| \leq \frac{2}{m_{n}}\left(2\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}+\sqrt{\frac{2}{\pi}}\right)\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty} \tag{7.164}
\end{equation*}
$$

To conclude the proof, we just need to apply Efron-Stein inequality (Proposition 2.5).

## Concentration with respect to the signal

Lemma 7.27. Under the assumptions of Theorem 7.1, for all $(t, \epsilon) \in[0,1] \times \mathcal{B}_{n}$ :

$$
\mathbb{E}\left[\left(\frac{\mathbb{E}\left[\ln \widehat{\mathcal{Z}}_{t, \epsilon} \mid \mathbf{X}^{*}\right]}{m_{n}}-\frac{\mathbb{E} \ln \widehat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}\right)^{2}\right] \leq \frac{C_{6}}{n \rho_{n}}+\frac{C_{7} \rho_{n}}{n \alpha_{n}^{2}},
$$

where $C_{7}:=8 S^{2}$ and

$$
\begin{aligned}
C_{6}:=8 S^{2}\left(3\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+\right. & \left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty} \\
& \left.+12\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+2 \sqrt{\frac{2}{\pi}}\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right)^{2} .
\end{aligned}
$$

Proof. We see $g\left(\mathbf{X}^{*}\right)=\mathbb{E}\left[\ln \widehat{\mathcal{Z}}_{t, \epsilon} \mid \mathbf{X}^{*}\right] / m_{n}$ as a function of $\mathbf{X}^{*}$. For all $j \in\{1, \ldots, n\}$ :

$$
\begin{align*}
\frac{\partial g}{\partial X_{j}^{*}}= & -\frac{1}{m_{n}} \mathbb{E}\left[\left.\left\langle\frac{\partial \widehat{\mathcal{H}}_{t, \epsilon}}{\partial X_{j}^{*}}\right\rangle_{n, t, \epsilon} \right\rvert\, \mathbf{X}^{*}\right] \\
=-\frac{1}{m_{n}} & \sqrt{\frac{1-t}{\Delta k_{n}}} \sum_{\mu=1}^{m_{n}} \mathbb{E}\left[W_{\mu j} \partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)\left(\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{n, t, \epsilon}+Z_{\mu}\right) \mid \mathbf{X}^{*}\right] \\
& +\frac{1}{m_{n}} \mathbb{E}\left[\left\langle R_{1}(t, \epsilon)\left(X_{j}^{*}-x_{j}\right)+\sqrt{R_{1}(t, \epsilon)} \widetilde{Z}_{j}\right\rangle_{n, t, \epsilon} \mid \mathbf{X}^{*}\right] \\
=-\frac{1}{m_{n}} & \sqrt{\frac{1-t}{\Delta k_{n}}} \sum_{\mu=1}^{m_{n}} \mathbb{E}\left[W_{\mu j} \partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{n, t, \epsilon} \mid \mathbf{X}^{*}\right] \\
& +\frac{R_{1}(t, \epsilon)}{m_{n}} \mathbb{E}\left[\left(X_{j}^{*}-\left\langle x_{j}\right\rangle_{n, t, \epsilon}\right) \mid \mathbf{X}^{*}\right] \tag{7.165}
\end{align*}
$$

To get the last equality we use that $\left.\mathbb{E} \sqrt{R_{1}(t, \epsilon)} \widetilde{Z}_{j} \mid \mathbf{X}^{*}\right]=0$ and

$$
\mathbb{E}\left[W_{\mu j} \partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right) Z_{\mu} \mid \mathbf{X}^{*}\right]=\mathbb{E}\left[W_{\mu j} \partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right) \mid \mathbf{X}^{*}\right] \mathbb{E}\left[Z_{\mu}\right]=0
$$

A Gaussian integration by parts with respect to $W_{\mu j}$ yields

$$
\begin{aligned}
& \mathbb{E}\left[W_{\mu j} \partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{n, t, \epsilon} \mid \mathbf{X}^{*}\right] \\
& =\sqrt{\frac{1-t}{k_{n} \Delta}} \mathbb{E}\left[X_{j}^{*}\left(\partial_{x} \varphi^{2}+\varphi \partial_{x x} \varphi\right)\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right) \mid \mathbf{X}^{*}\right] \\
& \quad-\sqrt{\frac{1-t}{\Delta k_{n}}} \mathbb{E}\left[X_{j}^{*} \partial_{x x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)\left\langle\varphi\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}\right)\right\rangle_{n, t, \epsilon} \mid \mathbf{X}^{*}\right] \\
& \quad-\sqrt{\frac{1-t}{\Delta k_{n}}} \mathbb{E}\left[\partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)\left\langle x_{j} \partial_{x} \varphi\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}\right)\right\rangle_{n, t, \epsilon} \mid \mathbf{X}^{*}\right] \\
& +\sqrt{\frac{1-t}{\Delta k_{n}}} \mathbb{E}\left[\partial _ { x } \varphi ( S _ { \mu } ^ { ( t , \epsilon ) } , \mathbf { A } _ { \mu } ) \left\langle\varphi\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}\right)\right.\right. \\
& \\
& \left.\left.\quad \cdot\left(X_{j}^{*} \partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)-x_{j} \partial_{x} \varphi\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}\right)\right)\left(\Gamma_{\mu}^{(t, \epsilon)}+Z_{\mu}\right)\right\rangle_{n, t, \epsilon} \mid \mathbf{X}^{*}\right] \\
& -\sqrt{\frac{1-t}{\Delta k_{n}}} \mathbb{E}\left[\partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)\left\langle\varphi\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}\right)\right\rangle_{n, t, \epsilon}\right. \\
& \left.\quad \cdot\left\langle\left(X_{j}^{*} \partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)-x_{j} \partial_{x} \varphi\left(s_{\mu}^{(t, \epsilon)}, \mathbf{a}_{\mu}\right)\right)\left(\Gamma_{\mu}^{(t, \epsilon)}+Z_{\mu}\right)\right\rangle_{n, t, \epsilon} \mid \mathbf{X}^{*}\right]
\end{aligned}
$$

It directly follows that $\left|\mathbb{E}\left[W_{\mu j} \partial_{x} \varphi\left(S_{\mu}^{(t, \epsilon)}, \mathbf{A}_{\mu}\right)\left\langle\Gamma_{\mu}^{(t, \epsilon)}\right\rangle_{n, t, \epsilon} \mid \mathbf{X}^{*}\right]\right| \leq \sqrt{\frac{\Delta}{k_{n}}} \widetilde{C}_{6}$ where

$$
\begin{aligned}
& \widetilde{C}_{6}:=2 S\left(\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}+\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}\left\|\frac{\partial_{x x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}+4\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right. \\
&\left.+2 \sqrt{\frac{2}{\pi}}\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty}\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}^{2}\right) .
\end{aligned}
$$

Making use of this upper bound, we obtain for all $j \in\{1, \ldots, n\}$ :

$$
\begin{align*}
\left|\frac{\partial g}{\partial X_{j}^{*}}\right| & \leq \frac{\widetilde{C}_{6}}{k_{n}}+\frac{2 S K_{n}}{m_{n}} \\
& =\frac{\widetilde{C}_{6}}{k_{n}}+\frac{2 S}{m_{n}}\left(2 s_{n}+2 \frac{\alpha_{n}}{\rho_{n}} C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)\right) \\
& =\frac{1}{n \rho_{n}}\left(\widetilde{C}_{6}+4 S C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)\right)+\frac{2 S}{n \alpha_{n}} . \tag{7.166}
\end{align*}
$$

For a fixed $j \in\{1, \ldots, n\}$, let $\mathbf{X}^{(j)}$ be a random vector such that $X_{i}^{(j)}=X_{i}^{*}$ for $i \neq j$ and $X_{j}^{(j)} \sim P_{X}^{(n)}$ independently of everything else. By the mean-value theorem, and thanks to (7.166), we have

$$
\begin{aligned}
\mathbb{E}_{\mathbf{X}^{*}} \mathbb{E}_{X_{j}^{(j)}} & {\left[\left(g\left(\mathbf{X}^{*}\right)-g\left(\mathbf{X}^{*(j)}\right)\right)^{2}\right] } \\
& \leq\left(\frac{1}{n \rho_{n}}\left(\widetilde{C}_{6}+4 S C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)\right)+\frac{2 S}{n \alpha_{n}}\right)^{2} \mathbb{E}\left[\left(X_{j}^{*}-X_{j}^{(j)}\right)^{2}\right] \\
& \leq \frac{4}{n^{2} \rho_{n}}\left(\widetilde{C}_{6}+4 S C_{1}\left(\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty}\right)\right)^{2}+\frac{16 S^{2} \rho_{n}}{n^{2} \alpha_{n}^{2}}
\end{aligned}
$$

where in the last equality we use that

$$
\mathbb{E}\left[\left(X_{j}^{*}-X_{j}^{(j)}\right)^{2}\right]=2 \rho_{n} \mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]-2 \rho_{n}^{2} \mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}\right]^{2} \leq 2 \rho_{n} \mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=2 \rho_{n}
$$

and the simple inequality $(a+b)^{2} \leq 2 a^{2}+2 b^{2}$. To end the proof we just need to apply Proposition 2.5.

Proof of Proposition 7.23 Combining Lemmas 7.24, 7.25, 7.26 and 7.27 yields

$$
\begin{equation*}
\operatorname{Var}\left(\frac{\ln \widehat{\mathcal{Z}}_{t, \epsilon}}{m_{n}}\right) \leq \frac{C_{2}+C_{4}}{n \alpha_{n} \rho_{n}}+\frac{C_{3}+C_{7} \rho_{n}}{n \alpha_{n}^{2}}+\frac{C_{5}}{n \alpha_{n}}+\frac{C_{6}}{n \rho_{n}} . \tag{7.167}
\end{equation*}
$$

Plugging (7.167) back in 7.160 finally gives

$$
\begin{aligned}
\operatorname{Var}\left(\frac{\ln \mathcal{Z}_{t, \epsilon}}{m_{n}}\right) & \leq \frac{C_{2}+C_{4}}{n \alpha_{n} \rho_{n}}+\frac{C_{3}+C_{7} \rho_{n}+1.5}{n \alpha_{n}^{2}}+\frac{C_{5}+1.5}{n \alpha_{n}}+\frac{C_{6}}{n \rho_{n}} \\
& \leq \frac{C_{2}+C_{4}+C_{5}+1.5}{n \alpha_{n} \rho_{n}}+\frac{C_{3}+C_{7}+1.5}{n \alpha_{n}^{2}}+\frac{C_{6}}{n \rho_{n}} \\
& =\frac{1}{n \alpha_{n} \rho_{n}}\left(C_{2}+C_{4}+C_{5}+1.5+\frac{\rho_{n}}{\alpha_{n}}\left(C_{3}+C_{7}+1.5\right)+\alpha_{n} C_{6}\right),
\end{aligned}
$$

where we use $\rho_{n} \leq 1$ to obtain the second inequality.

## 7.G Concentration of the overlap

In this appendix we prove Proposition 7.10. Define the average free entropy $f_{n, \epsilon}(t):=\mathbb{E}\left[\ln \mathcal{Z}_{t, \epsilon]} / m_{n} \mathbb{E} \ln \mathcal{Z}_{t, \epsilon}\right.$. In this section we think of it as a function of $R_{1}=R_{1}(t, \epsilon)$ and $R_{2}=R_{2}(t, \epsilon)$, i.e., $\left(R_{1}, R_{2}\right) \mapsto f_{n, \epsilon}(t)$. Similarly, we view the free entropy for a realization of the quenched disorder variables as a function

$$
\left(R_{1}, R_{2}\right) \mapsto F_{n, \epsilon}(t):=\frac{\ln \mathcal{Z}_{t, \epsilon}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{(t, \epsilon)}, \mathbf{W}, \mathbf{V}\right)}{m_{n}}
$$

In this appendix, to lighten the notations, we drop the indices of the angular brackets $\langle-\rangle_{n, t, \epsilon}$ and simply write $\langle-\rangle$. We denote with • the scalar product between two vectors. We define:

$$
\mathcal{L}:=\frac{1}{k_{n}}\left(\frac{\|\mathbf{x}\|^{2}}{2}-\mathbf{x}^{\top} \mathbf{X}^{*}-\frac{\mathbf{x}^{\top} \widetilde{\mathbf{Z}}}{2 \sqrt{R_{1}}}\right) .
$$

The fluctuations of the overlap $Q:=\frac{\mathbf{x}^{\top} \mathbf{X}^{*}}{k_{n}}$ and those of $\mathcal{L}$ are related through the inequality

$$
\begin{equation*}
\frac{1}{4} \mathbb{E}\left\langle(Q-\mathbb{E}\langle Q\rangle)^{2}\right\rangle \leq \mathbb{E}\left\langle(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right\rangle \tag{7.168}
\end{equation*}
$$

The proof of (7.168) is based on integrations by parts with respect to $\widetilde{Z}$ and a repeated use of the Nishimori identity (see Lemma 2.1). Proposition 7.10 is then a direct consequence of the following result.

Proposition 7.28 (Concentration of $\mathcal{L}$ on $\mathbb{E}\langle\mathcal{L}\rangle)$. Suppose that $\Delta>0$, (H1), (H2), (H3) hold, $\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$, and the family of functions $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}},\left(q_{\epsilon}\right)_{\in \in \mathcal{B}_{n}}$ are regular. Further assume that there exist real positive numbers $M_{\alpha}, M_{\rho / \alpha}, m_{\rho / \alpha}$ such that $\forall n \in \mathbb{N}^{*}$ :

$$
\alpha_{n} \leq M_{\alpha} \quad \text { and } \quad \frac{m_{\rho / \alpha}}{n}<\frac{\rho_{n}}{\alpha_{n}} \leq M_{\rho / \alpha} .
$$

Let $\left(s_{n}\right)_{n \in \mathbb{N}^{*}}$ be a sequence of real numbers in $(0,1 / 2]$. Define $\mathcal{B}_{n}:=\left[s_{n}, 2 s_{n}\right]^{2}$. Then, $\forall t \in[0,1]$ :

$$
\begin{equation*}
\int_{\mathcal{B}_{n}} d \epsilon \mathbb{E}\left\langle\left(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle_{n, t, \epsilon}\right)^{2}\right\rangle_{n, t, \epsilon} \leq \frac{C}{\rho_{n}^{2}\left(\frac{\rho_{n} n}{\alpha_{n} m_{\rho / \alpha}}\right)^{\frac{1}{3}}-\rho_{n}^{2}} \tag{7.169}
\end{equation*}
$$

where $C$ is a polynomial in $\left(S,\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x}}{\sqrt{\Delta}}\right\|_{\infty}, M_{\alpha}, M_{\rho / \alpha}, m_{\rho / \alpha}\right)$ with positive coefficients.

Due to the identity $\mathbb{E}\left\langle(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right\rangle=\mathbb{E}\left\langle(\mathcal{L}-\langle\mathcal{L}\rangle)^{2}\right\rangle+\mathbb{E}\left[(\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right]$, Proposition 7.10 directly follows from the next two lemmas.

Lemma 7.29 (Concentration of $\mathcal{L}$ on $\langle\mathcal{L}\rangle$ ). Under the assumptions of Proposition 7.28, $\forall t \in[0,1]$ :

$$
\int_{\mathcal{B}_{n}} d \epsilon \mathbb{E}\left\langle\left(\mathcal{L}-\langle\mathcal{L}\rangle_{n, t, \epsilon}\right)^{2}\right\rangle_{n, t, \epsilon} \leq \frac{1}{n \rho_{n}} .
$$

The second lemma states that $\mathcal{L}$ concentrates w.r.t. the realizations of the quenched disorder variables. It is a consequence of the concentration of the free entropy (see Proposition 7.23 in Appendix 7.F).

Lemma 7.30 (Concentration of $\langle\mathcal{L}\rangle$ on $\mathbb{E}\langle\mathcal{L}\rangle$ ). Under the assumptions of Proposition 7.10, $\forall t \in[0,1]$ :

$$
\begin{equation*}
\int_{\mathcal{B}_{n}} d \epsilon \mathbb{E}\left[\left(\langle\mathcal{L}\rangle_{n, t, \epsilon}-\mathbb{E}\langle\mathcal{L}\rangle_{n, t, \epsilon}\right)^{2}\right] \leq \frac{C}{\rho_{n}^{2}\left(\frac{\rho_{n} n}{\alpha_{n} m_{\rho / \alpha}}\right)^{\frac{1}{3}}-\rho_{n}^{2}}, \tag{7.170}
\end{equation*}
$$

where $C$ is a polynomial in $\left(S,\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x} \varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x}}{\sqrt{\Delta}}\right\|_{\infty}, M_{\alpha}, M_{\rho / \alpha}, m_{\rho / \alpha}\right)$ with positive coefficients.

We now turn to the proof of Lemmas 7.29 and 7.30. The main ingredient will be a set of formulas for the first two partial derivatives of the free entropy w.r.t. $R_{1}=R_{1}(t, \epsilon)$. For any given realization of the quenched disorder,

$$
\begin{align*}
\frac{d F_{n, \epsilon}(t)}{d R_{1}} & =-\frac{\rho_{n}}{\alpha_{n}}\langle\mathcal{L}\rangle-\frac{1}{2 m_{n}}\left(\left\|\mathbf{X}^{*}\right\|^{2}+\frac{\widetilde{\mathbf{Z}}^{\top} \mathbf{X}^{*}}{\sqrt{R_{1}}}\right)  \tag{7.171}\\
\frac{1}{m_{n}} \frac{d^{2} F_{n, \epsilon}(t)}{d R_{1}^{2}} & =\left(\frac{\rho_{n}}{\alpha_{n}}\right)^{2}\left(\left\langle\mathcal{L}^{2}\right\rangle-\langle\mathcal{L}\rangle^{2}\right)+\frac{1}{4 m_{n}^{2} R_{1}^{3 / 2}} \widetilde{\mathbf{Z}}^{\top}\left(\mathbf{X}^{*}-\langle\mathbf{x}\rangle\right) . \tag{7.172}
\end{align*}
$$

Taking an expectation on both sides of (7.171) yields

$$
\begin{equation*}
\frac{d f_{n, \epsilon}(t)}{d R_{1}}=-\frac{\rho_{n}}{\alpha_{n}}\left(\mathbb{E}\langle\mathcal{L}\rangle+\frac{1}{2}\right)=\frac{\rho_{n}}{2 \alpha_{n}}\left(\frac{\mathbb{E}\|\langle\mathbf{x}\rangle\|^{2}}{k_{n}}-1\right) . \tag{7.173}
\end{equation*}
$$

To obtain the second equality, we simplify $\mathbb{E}\langle\mathcal{L}\rangle$ thanks to a Gaussian integration by parts w.r.t. $\widetilde{\mathbf{Z}}$ and the Nishimori identity $\mathbb{E}\left\langle\mathbf{x}^{\top} \mathbf{X}^{*}\right\rangle=\mathbb{E}\|\langle\mathbf{x}\rangle\|^{2}$ (see Lemma 2.1 . Taking an expectation on both sides of $(7.172)$ and integrating by parts w.r.t. the standard Gaussian random vector $\widetilde{\mathbf{Z}}$ gives

$$
\begin{equation*}
\frac{1}{m_{n}} \frac{d^{2} f_{n, \epsilon}(t)}{d R_{1}^{2}}=\left(\frac{\rho_{n}}{\alpha_{n}}\right)^{2} \mathbb{E}\left[\left\langle\mathcal{L}^{2}\right\rangle-\langle\mathcal{L}\rangle^{2}\right]-\frac{1}{4 m_{n}^{2} R_{1}} \mathbb{E}\left[\left\langle\|\mathbf{x}\|^{2}\right\rangle-\|\langle\mathbf{x}\rangle\|^{2}\right] . \tag{7.174}
\end{equation*}
$$

Proof of Lemma 7.29. From (7.174 we directly obtain that

$$
\begin{align*}
\mathbb{E}\left\langle(\mathcal{L}-\langle\mathcal{L}\rangle)^{2}\right\rangle & =\left(\frac{\alpha_{n}}{\rho_{n}}\right)^{2} \frac{1}{m_{n}} \frac{d^{2} f_{n, \epsilon}(t)}{d R_{1}^{2}}+\left(\frac{\alpha_{n}}{\rho_{n}}\right)^{2} \frac{1}{4 m_{n}^{2} R_{1}} \mathbb{E}\left[\left\langle\|\mathbf{x}\|^{2}\right\rangle-\|\langle\mathbf{x}\rangle\|^{2}\right] \\
& \leq \frac{\alpha_{n}}{\rho_{n}^{2} n} \frac{d^{2} f_{n, \epsilon}(t)}{d R_{1}^{2}}+\frac{1}{4 \epsilon_{1} n \rho_{n}}, \tag{7.175}
\end{align*}
$$

where we used that $\mathbb{E}\left\langle\|\mathbf{x}\|^{2}\right\rangle=\mathbb{E}\left\|\mathbf{X}^{*}\right\|^{2}=n \rho_{n}$ by the Nishimori identity and $R_{1} \geq \epsilon_{1}$. Remember that $\mathcal{B}_{n}:=\left[s_{n}, 2 s_{n}\right]^{2}$. By assumption the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ are regular. Therefore, $R^{t}:\left(\epsilon_{1}, \epsilon_{2}\right) \mapsto\left(R_{1}(t, \epsilon), R_{2}(t, \epsilon)\right)$ is a
$C^{1}$-diffeomorphism whose Jacobian determinant $\left|J_{R^{t}}\right|$ satisfies $\forall \epsilon \in \mathcal{B}_{n}:\left|J_{R^{t}}(\epsilon)\right| \geq$ 1. Integrating both sides of 7.175) over $\epsilon \in \mathcal{B}_{n}$ yields

$$
\begin{align*}
\int_{\mathcal{B}_{n}} d \epsilon \mathbb{E}\left\langle(\mathcal{L}-\langle\mathcal{L}\rangle)^{2}\right\rangle \leq & \frac{\alpha_{n}}{\rho_{n}^{2} n} \int_{R^{t}\left(\mathcal{B}_{n}\right)} \frac{d R_{1} d R_{2}}{\left|J_{R^{t}}\left(\left(R^{t}\right)^{-1}\left(R_{1}, R_{2}\right)\right)\right|} \frac{d^{2} f_{n, \epsilon}(t)}{d R_{1}^{2}} \\
& +\frac{1}{4 n \rho_{n}} \int_{\mathcal{B}_{n}} \frac{d \epsilon_{1}}{\epsilon_{1}} d \epsilon_{2} \\
\leq & \frac{\alpha_{n}}{\rho_{n}^{2} n} \int_{R^{t}\left(\mathcal{B}_{n}\right)} d R_{1} d R_{2} \frac{d^{2} f_{n, \epsilon}(t)}{d R_{1}^{2}}+\frac{s_{n}}{4 n \rho_{n}} \ln 2 \tag{7.176}
\end{align*}
$$

Note that $R^{t}\left(\mathcal{B}_{n}\right) \subseteq\left[s_{n}, 2 s_{n}+\frac{\alpha_{n}}{\rho_{n}} r_{\text {max }}\right] \times\left[s_{n}, 2 s_{n}+1\right]$ (by definition of the interpolation functions). Thus,

$$
\begin{align*}
\int_{\mathcal{B}_{n}} d \epsilon \mathbb{E}\left\langle(\mathcal{L}-\langle\mathcal{L}\rangle)^{2}\right\rangle & \leq \frac{\alpha_{n}}{\rho_{n}^{2} n} \int_{s_{n}}^{2 s_{n}+1} d R_{2}\left[\frac{d f_{n, \epsilon}(t)}{d R_{1}}\right]_{R_{1}=s_{n}}^{2 s_{n}+\frac{\alpha_{n}}{\rho_{n}} r_{\max }}+\frac{s_{n}}{4 n \rho_{n}} \ln 2 \\
& \leq \frac{1+s_{n}}{2 \rho_{n} n}+\frac{s_{n}}{4 n \rho_{n}} \ln 2 \leq \frac{1}{n \rho_{n}} \tag{7.177}
\end{align*}
$$

To obtain the second inequality, we bound the partial derivative of the free entropy using (7.173) and $\left.\mathbb{E}\|\langle\mathbf{x}\rangle\|^{2}\right\rangle \leq \mathbb{E}\left\langle\|\mathbf{x}\|^{2}\right\rangle=n \rho_{n}$ (again by the Nishimori identity),

$$
\begin{equation*}
\left|\frac{d f_{n, \epsilon}(t)}{d R_{1}}\right|=-\frac{d f_{n, \epsilon}(t)}{d R_{1}}=\frac{\rho_{n}}{2 \alpha_{n}}\left(1-\frac{\mathbb{E}\|\langle\mathbf{x}\rangle\|^{2}}{k_{n}}\right) \leq \frac{\rho_{n}}{2 \alpha_{n}} \tag{7.178}
\end{equation*}
$$

The last inequality follows from $s_{n} \leq 1 / 2$ and $(\ln 2) / 2<1$.
Proof of Lemma 7.30. We define the two functions

$$
\begin{align*}
\widetilde{F}\left(R_{1}\right) & :=F_{n, \epsilon}(t)-\frac{\sqrt{R_{1}}}{m_{n}} 2 S \sum_{i=1}^{n}\left|\widetilde{Z}_{i}\right|  \tag{7.179}\\
\widetilde{f}\left(R_{1}\right) & :=\mathbb{E} \widetilde{F}\left(R_{1}\right)=f_{n, \epsilon}(t)-\frac{\sqrt{R_{1}}}{\alpha_{n}} 2 S \mathbb{E}\left|\widetilde{Z}_{1}\right| . \tag{7.180}
\end{align*}
$$

Thanks to (7.172), we see that the second derivative of $\widetilde{F}\left(R_{1}\right)$ is nonnegative so $\widetilde{F}\left(R_{1}\right)$ is convex. Without the extra term on the right-hand side of 7.179), $F_{n, \epsilon}(t)$ is not necessarily convex in $R_{1}$ although $f_{n, \epsilon}(t)$ is convex (it can be shown easily). Note that $\widetilde{f}\left(R_{1}\right)$ is also convex. Define $A:=\frac{1}{m_{n}} \sum_{i=1}^{n}\left|\widetilde{Z}_{i}\right|-\mathbb{E}\left|\widetilde{Z}_{i}\right|$. From (7.179) and 7.180, we directly obtain that

$$
\begin{equation*}
\widetilde{F}\left(R_{1}\right)-\widetilde{f}\left(R_{1}\right)=F_{n, \epsilon}(t)-f_{n, \epsilon}(t)-\sqrt{R_{1}} 2 S A \tag{7.181}
\end{equation*}
$$

Thanks to (7.171) and 7.173), the difference of the derivatives (w.r.t. $R_{1}$ ) satisfies

$$
\begin{equation*}
\widetilde{F}^{\prime}\left(R_{1}\right)-\widetilde{f}^{\prime}\left(R_{1}\right)=\frac{\rho_{n}}{\alpha_{n}}(\mathbb{E}\langle\mathcal{L}\rangle-\langle\mathcal{L}\rangle)-\frac{\rho_{n}}{2 \alpha_{n}}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1+\frac{\widetilde{\mathbf{Z}}^{\top} \mathbf{X}^{*}}{k_{n} \sqrt{R_{1}}}\right)-\frac{S A}{\sqrt{R_{1}}} . \tag{7.182}
\end{equation*}
$$

For all $\delta \in\left(0, s_{n}\right), C_{\delta}\left(R_{1}\right):=\widetilde{f^{\prime}}\left(R_{1}+\delta\right)-\widetilde{f^{\prime}}\left(R_{1}-\delta\right) \geq 0$ (this is well-defined because $\delta<s_{n} \leq R_{1}$ ). It follows from Lemma 2.8 (applied to the convex functions $G=\widetilde{F}, g=\widetilde{f})$ and the two identities (7.181), (7.182) that $\forall \delta \in\left(0, s_{n}\right)$ :

$$
\begin{align*}
& \frac{\rho_{n}}{\alpha_{n}}|\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle| \\
& \quad \leq \delta^{-1} \sum_{u \in\left\{R_{1}-\delta, R_{1}, R_{1}+\delta\right\}}\left|\left(F_{n, \epsilon}(t)-f_{n, \epsilon}(t)\right)_{R_{1}=u}\right|+2 S|A| \sqrt{u} \\
& \quad+C_{\delta}\left(R_{1}\right)+\frac{S|A|}{\sqrt{R_{1}}}+\frac{\rho_{n}}{2 \alpha_{n}}\left|\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1+\frac{\widetilde{\mathbf{Z}}^{\top} \mathbf{X}^{*}}{k_{n} \sqrt{R_{1}}}\right| \\
& \quad \leq \delta^{-1} \sum_{u \in\left\{R_{1}-\delta, R_{1}, R_{1}+\delta\right\}}\left|\left(F_{n, \epsilon}(t)-f_{n, \epsilon}(t)\right)_{R_{1}=u}\right| \\
& \quad+C_{\delta}\left(R_{1}\right)+S|A|\left(\frac{1}{\sqrt{R_{1}}}+\frac{6 \sqrt{R_{1}}}{\delta}\right)+\frac{\rho_{n}}{2 \alpha_{n}}\left|\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}-1+\frac{\widetilde{\mathbf{Z}}^{\top} \mathbf{X}^{*}}{k_{n} \sqrt{R_{1}}}\right| . \tag{7.183}
\end{align*}
$$

The last inequality is due to $\sqrt{R_{1}+\delta}+\sqrt{R_{1}-\delta} \leq 2 \sqrt{R_{1}}$. Taking the square and then the expectation on both sides of the inequality (7.183), and making use of $\left(\sum_{i=1}^{6} v_{i}\right)^{2} \leq 6 \sum_{i=1}^{6} v_{i}^{2}$ (by convexity), yields

$$
\begin{align*}
& \mathbb{E}\left[(\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right] \leq \frac{6}{\delta^{2}}\left(\frac{\alpha_{n}}{\rho_{n}}\right)^{2} \sum_{u \in\left\{R_{1}-\delta, R_{1}, R_{1}+\delta\right\}} \operatorname{Var}\left(\left.F_{n, \epsilon}(t)\right|_{R_{1}=u}\right)+6\left(\frac{\alpha_{n} C_{\delta}\left(R_{1}\right)}{\rho_{n}}\right)^{2} \\
& \quad+6\left(\frac{\alpha_{n}}{\rho_{n}}\right)^{2} S^{2} \mathbb{E}\left[A^{2}\right]\left(\frac{1}{R_{1}}+\frac{12}{\delta}+\frac{36 R_{1}}{\delta^{2}}\right)+\frac{3}{2} \operatorname{Var}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}+\frac{\widetilde{\mathbf{Z}}^{\top} \mathbf{X}^{*}}{k_{n} \sqrt{R_{1}}}\right) . \tag{7.184}
\end{align*}
$$

By Proposition 7.23, under our assumptions, the free entropy $F_{n, \epsilon}(t)=\ln \mathcal{Z}_{t, \epsilon} / m_{n}$ concentrates such that

$$
\begin{equation*}
\operatorname{Var}\left(F_{n, \epsilon}(t)\right) \leq \frac{C}{n \alpha_{n} \rho_{n}} \tag{7.185}
\end{equation*}
$$

where $C$ is a polynomial in $\left(S,\left\|\frac{\varphi}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x \varphi}}{\sqrt{\Delta}}\right\|_{\infty},\left\|\frac{\partial_{x x}}{\sqrt{\Delta}}\right\|_{\infty}\right)$ with positive coefficients. Remark that, by independence of the noise variables, we have:

$$
\begin{equation*}
\mathbb{E}\left[A^{2}\right] \leq \frac{1-2 / \pi}{n \alpha_{n}^{2}}<\frac{1}{n \alpha_{n}^{2}} \tag{7.186}
\end{equation*}
$$

Also, the last term on the right hand side of (7.184) satisfies

$$
\begin{aligned}
\operatorname{Var}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}+\frac{\mathbf{X}^{*} \cdot \widetilde{\mathbf{Z}}}{k_{n} \sqrt{R_{1}}}\right) & =\operatorname{Var}\left(\frac{\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}\right)+\operatorname{Var}\left(\frac{\mathbf{X}^{*} \cdot \widetilde{\mathbf{Z}}}{k_{n} \sqrt{R_{1}}}\right) \\
& =\frac{n}{k_{n}^{2}} \operatorname{Var}\left(\left(X_{1}^{*}\right)^{2}\right)+\frac{n}{k_{n}^{2} R_{1}} \operatorname{Var}\left(X_{1}^{*} \widetilde{Z}_{1}\right) \leq \frac{S^{4}}{n \rho_{n}}+\frac{1}{n \rho_{n} R_{1}} .
\end{aligned}
$$

Plugging (7.185), 7.186) and the latter inequality back in (7.184) yields

$$
\begin{align*}
\mathbb{E}\left[(\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right] \leq & \frac{18 C \alpha_{n}}{n \rho_{n}^{3} \delta^{2}}+6\left(\frac{\alpha_{n} C_{\delta}\left(R_{1}\right)}{\rho_{n}}\right)^{2}+\frac{6 S^{2}}{n \rho_{n}^{2}}\left(\frac{1}{R_{1}}+\frac{12}{\delta}+\frac{36 R_{1}}{\delta^{2}}\right) \\
& \quad+\frac{3 S^{4}}{2 n \rho_{n}}+\frac{3}{2 n \rho_{n} R_{1}} \\
\leq & \frac{18 C \alpha_{n}}{n \rho_{n}^{3} \delta^{2}}+6\left(\frac{\alpha_{n} C_{\delta}\left(R_{1}\right)}{\rho_{n}}\right)^{2}+\frac{294 S^{2}}{n \rho_{n}^{2}} \frac{R_{1}}{\delta^{2}} \\
& \quad+\frac{3}{2 n \rho_{n}}\left(S^{4}+\frac{1}{R_{1}}\right) \tag{7.187}
\end{align*}
$$

where the last inequality follows from $R_{1}^{-1} \leq \delta^{-1} \leq R_{1} / \delta^{2}$.
The next step is to integrate both sides of (7.187) over $\epsilon \in \mathcal{B}_{n}:=\left[s_{n}, 2 s_{n}\right]^{2}$. By assumption, the families of functions $\left(q_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ and $\left(r_{\epsilon}\right)_{\epsilon \in \mathcal{B}_{n}}$ are regular so

$$
R^{t}:\left(\epsilon_{1}, \epsilon_{2}\right) \mapsto\left(R_{1}(t, \epsilon), R_{2}(t, \epsilon)\right)
$$

is a $C^{1}$-diffeomorphism whose Jacobian determinant $\left|J_{R^{t}}\right|$ satisfies $\forall \epsilon \in \mathcal{B}_{n}$ : $\left|J_{R^{t}}(\epsilon)\right| \geq 1$. Besides, $R^{t}\left(\mathcal{B}_{n}\right) \subseteq\left[s_{n}, K_{n}\right] \times\left[s_{n}, 2 s_{n}+1\right]$ where $K_{n}:=2 s_{n}+\frac{\alpha_{n}}{\rho_{n}} r_{\text {max }}$. Therefore,

$$
\begin{align*}
\int_{\mathcal{B}_{n}} d \epsilon \frac{294 S^{2}}{n \rho_{n}^{2}} \frac{R_{1}(t, \epsilon)}{\delta^{2}} & \leq \frac{294 S^{2}}{n \rho_{n}^{2}} \int_{\mathcal{B}_{n}} d \epsilon \frac{K_{n}}{\delta^{2}} \\
& =\frac{294 S^{2}}{n \rho_{n}^{2}} \frac{K_{n} s_{n}^{2}}{\delta^{2}} \leq 294 S^{2}\left(M_{\rho / \alpha}+r_{\max }\right) \frac{\alpha_{n} s_{n}^{2}}{n \rho_{n}^{3} \delta^{2}} \tag{7.188}
\end{align*}
$$

where we use that $K_{n}=\left(2 s_{n} \rho_{n} / \alpha_{n}+r_{\max }\right)\left(\alpha_{n} / \rho_{n}\right) \leq\left(M_{\rho / \alpha}+r_{\max }\right)\left(\alpha_{n} / \rho_{n}\right)$ as $s_{n} \leq 1 / 2$ and $\rho_{n} / \alpha_{n} \leq M_{\rho / \alpha}$. We now upper bound the integral of $\left(\alpha_{n} C_{\delta}\left(R_{1}\right) / \rho_{n}\right)^{2}$. Remember that $C_{\delta}\left(R_{1}\right):=\widetilde{f^{\prime}}\left(R_{1}+\delta\right)-\widetilde{f^{\prime}}\left(R_{1}-\delta\right) \geq 0$. By the definition (7.180) of $\widetilde{f}$ and the upper bound (7.178), we have

$$
\begin{equation*}
\left|\widetilde{f}^{\prime}\left(R_{1}\right)\right| \leq \frac{\rho_{n}}{2 \alpha_{n}}+\frac{S}{\alpha_{n} \sqrt{R_{1}}} \mathbb{E}\left|\widetilde{Z}_{1}\right| \leq \frac{\rho_{n}}{2 \alpha_{n}}+\frac{S}{\alpha_{n} \sqrt{R_{1}}} . \tag{7.189}
\end{equation*}
$$

This inequality directly implies that $\left|C_{\delta}\left(R_{1}\right)\right| \leq \frac{1}{\alpha_{n}}\left(\rho_{n}+2 S / \sqrt{s_{n}-\delta}\right)$, hence

$$
\begin{aligned}
& \int_{\mathcal{B}_{n}} d \epsilon C_{\delta}\left(R_{1}(t, \epsilon)\right)^{2} \leq \frac{1}{\alpha_{n}}\left(\rho_{n}+\frac{2 S}{\sqrt{s_{n}-\delta}}\right) \int_{\mathcal{B}_{n}} d \epsilon C_{\delta}\left(R_{1}(t, \epsilon)\right) \\
&=\frac{1}{\alpha_{n}}\left(\rho_{n}+\frac{2 S}{\sqrt{s_{n}-\delta}}\right) \int_{R^{t}\left(\mathcal{B}_{n}\right)} \frac{d R_{1} d R_{2}}{\left|J_{R^{t}}\left(\left(R^{t}\right)^{-1}\left(R_{1}, R_{2}\right)\right)\right|} C_{\delta}\left(R_{1}\right) \\
& \leq \frac{1}{\alpha_{n}}\left(\rho_{n}+\frac{2 S}{\sqrt{s_{n}-\delta}}\right) \int_{s_{n}}^{2 s_{n}+1} d R_{2} \int_{s_{n}}^{K_{n}} d R_{1} C_{\delta}\left(R_{1}\right) \\
& \leq \frac{1}{\alpha_{n}}\left(\rho_{n}+\frac{2 S}{\sqrt{s_{n}-\delta}}\right) \int_{s_{n}}^{2 s_{n}+1} d R_{2}\left(\widetilde{f}\left(K_{n}+\delta\right)-\widetilde{f}\left(K_{n}-\delta\right)\right. \\
&\left.\quad+\widetilde{f}\left(s_{n}-\delta\right)-\widetilde{f}\left(s_{n}+\delta\right)\right) .
\end{aligned}
$$

By the mean value theorem and the upper bound (7.189), we have

$$
\left|\widetilde{f}\left(R_{1}-\delta\right)-\widetilde{f}\left(R_{1}+\delta\right)\right| \leq 2 \delta\left(\frac{\rho_{n}}{2 \alpha_{n}}+\frac{S}{\alpha_{n} \sqrt{R_{1}-\delta}}\right) \leq \frac{\delta}{\alpha_{n}}\left(\rho_{n}+\frac{2 S}{\sqrt{s_{n}-\delta}}\right)
$$

uniformly in $R_{2}$. Therefore,

$$
\begin{align*}
\int_{\mathcal{B}_{n}} d \epsilon\left(\frac{\alpha_{n} C_{\delta}\left(R_{1}\right)}{\rho_{n}}\right)^{2} & \leq \frac{\left(1+s_{n}\right) \delta}{\rho_{n}^{2}}\left(\rho_{n}+\frac{2 S}{\sqrt{s_{n}-\delta}}\right)^{2} \\
& \leq \frac{3 \delta}{2 \rho_{n}^{2}}\left(\frac{1+2 S}{\sqrt{s_{n}-\delta}}\right)^{2}=\frac{3(1+2 S)^{2} \delta}{2 \rho_{n}^{2}\left(s_{n}-\delta\right)} . \tag{7.190}
\end{align*}
$$

For all $\epsilon \in \mathcal{B}_{n}$, we have $R_{1}(t, \epsilon) \geq s_{n}$ so (remember that $\int_{\mathcal{B}_{n}} d \epsilon=s_{n}^{2}$ )

$$
\begin{align*}
\int_{\mathcal{B}_{n}} d \epsilon \frac{3}{2 n \rho_{n}}\left(S^{4}+\frac{1}{R_{1}(t, \epsilon)}\right) & \leq \frac{3}{2 n \rho_{n}} \int_{\mathcal{B}_{n}} d \epsilon\left(S^{4}+\frac{1}{s_{n}}\right) \\
& \leq \frac{3 s_{n}}{2 n \rho_{n}}\left(S^{4} s_{n}+1\right) \leq \frac{3 s_{n}}{2 n \rho_{n}}\left(\frac{S^{4}}{2}+1\right) . \tag{7.191}
\end{align*}
$$

Integrating both sides of (7.187) over $\epsilon \in \mathcal{B}_{n}$ and making use of 7.188, 7.190), (7.191) yields

$$
\begin{array}{rl}
\int_{\mathcal{B}_{n}} & d \epsilon \mathbb{E}\left[(\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right] \\
& \leq \frac{\alpha_{n} s_{n}^{2}}{n \rho_{n}^{3} \delta^{2}}\left(18 C+294 S^{2}\left(M_{\rho / \alpha}+r_{\max }\right)\right)+\frac{3 s_{n}}{2 n \rho_{n}}\left(\frac{S^{4}}{2}+1\right)+\frac{9(1+2 S)^{2}}{\rho_{n}^{2}\left(\frac{s_{n}}{\delta}-1\right)} .
\end{array}
$$

We can use $\rho_{n} / \alpha_{n} \leq M_{\rho / \alpha}, \rho_{n} \leq 1, \delta \leq s_{n}$, and $s_{n} \leq 1 / 2$ (in this order) to show that

$$
\begin{aligned}
\frac{s_{n}}{n \rho_{n}}=\frac{\rho_{n}^{2} \delta^{2}}{\alpha_{n} s_{n}} \cdot \frac{\alpha_{n} s_{n}^{2}}{n \rho_{n}^{3} \delta^{2}} \leq \frac{M_{\rho / \alpha} \rho_{n} \delta^{2}}{s_{n}} \cdot \frac{\alpha_{n} s_{n}^{2}}{n \rho_{n}^{3} \delta^{2}} & \leq \frac{M_{\rho / \alpha} \delta^{2}}{s_{n}} \cdot \frac{\alpha_{n} s_{n}^{2}}{n \rho_{n}^{3} \delta^{2}} \\
& \leq M_{\rho / \alpha} s_{n} \cdot \frac{\alpha_{n} s_{n}^{2}}{n \rho_{n}^{3} \delta^{2}} \leq \frac{M_{\rho / \alpha}}{2} \cdot \frac{\alpha_{n} s_{n}^{2}}{n \rho_{n}^{3} \delta^{2}}
\end{aligned}
$$

Making use of this upper bound, we finally obtain

$$
\begin{equation*}
\int_{\mathcal{B}_{n}} d \epsilon \mathbb{E}\left[(\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right] \leq C_{1} \frac{\alpha_{n} s_{n}^{2}}{n \rho_{n}^{3} \delta^{2}}+C_{2} \frac{1}{\rho_{n}^{2}\left(\frac{s_{n}}{\delta}-1\right)}, \tag{7.192}
\end{equation*}
$$

where $C_{1}:=18 C+294 S^{2} r_{\max }+\left(294 S^{2}+3 S^{4} / 8+3 / 4\right) M_{\rho / \alpha}$ and $C_{2}:=9(1+2 S)^{2}$.
The convergence of the ratio $\delta / s_{n}$ to zero is a necessary condition for the second term on the right-hand side of (7.192) to vanish. If $\delta / s_{n} \rightarrow 0$ then $\left(\rho_{n}^{2}\left(s_{n} / \delta-1\right)\right)^{-1}=\Theta\left(\delta / \rho_{n}^{2} s_{n}\right)$. In that situation, both terms on the right-hand side of (7.192) are equivalent, that is, $\alpha_{n} s_{n}^{2} / n \rho_{n}^{3} \delta^{2}=\Theta\left(\delta / \rho_{n}^{2} s_{n}\right)$, if we choose $\delta \propto\left(\alpha_{n} / n \rho_{n}\right)^{\frac{1}{3}} s_{n}$. Note that we can choose $\delta \propto\left(\alpha_{n} / n \rho_{n}\right)^{\frac{1}{3}} s_{n}$ and still make sure that $\delta \in\left(0, s_{n}\right)$
because there exists $m_{\rho / \alpha}$ such that $\frac{\rho_{n}}{\alpha_{n}}>\frac{m_{\rho / \alpha}}{n}$ for all $n>0$. Plugging the choice $\delta=\left(\frac{m_{\rho / \alpha} \alpha_{n}}{n \rho_{n}}\right)^{\frac{1}{3}} s_{n}$ back in (7.192) ends the proof of the lemma,

$$
\begin{aligned}
\int_{\mathcal{B}_{n}} d \in \mathbb{E}\left[(\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right] & \leq \frac{C_{1}}{m_{\rho / \alpha}} \frac{1}{\rho_{n}^{2}\left(\frac{\rho_{n n}}{\alpha_{n} m_{\rho / \alpha}}\right)^{\frac{1}{3}}}+C_{2} \frac{1}{\rho_{n}^{2}\left(\frac{\rho_{n} n}{\alpha_{n} m_{\rho / \alpha}}\right)^{\frac{1}{3}}-\rho_{n}^{2}} \\
& \leq\left(\frac{C_{1}}{m_{\rho / \alpha}}+C_{2}\right) \frac{1}{\rho_{n}^{2}\left(\frac{\rho_{n} n}{\alpha_{n} m_{\rho / \alpha}}\right)^{\frac{1}{3}}-\rho_{n}^{2}}
\end{aligned}
$$

## 7.H Proof of Proposition 7.12

Before proving the proposition, we recall a few definitions for reader's convenience. We suppose that (H1), (H2), (H3) hold and $\Delta=\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$. For all $n \in \mathbb{N}^{*}$, we define the interval $\mathcal{B}_{n}:=\left[s_{n}, 2 s_{n}\right]$ where $\left(s_{n}\right)_{n \in \mathbb{N}^{*}}$ is a sequence of real numbers in $(0,1 / 2]$. Let $X_{i}^{*} \stackrel{\text { i.i.d. }}{\sim} P_{X}^{(n)}, \mathbf{A}_{\mu} \stackrel{\text { i.i.d. }}{\sim} P_{A}$ and $W_{\mu i}, V_{\mu}, U_{\mu}, Z_{\mu}, \widetilde{Z}_{i} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$ for $i=1 \ldots n$ and $\mu=1 \ldots m_{n}$. For $t \in[0,1]$ and $R:=\left(R_{1}, R_{2}\right) \in[0,+\infty) \times\left[0, t+2 s_{n}\right]$, consider the observations

$$
\left\{\begin{aligned}
Y_{\mu}^{\left(t, R_{2}\right)} & =\varphi\left(S_{\mu}^{\left(t, R_{2}\right)}, \mathbf{A}_{\mu}\right)+Z_{\mu}, 1 \leq \mu \leq m_{n} \\
& \sim P_{\text {out }}\left(\cdot \mid S_{\mu}^{\left(t, R_{2}\right)}\right) \\
\widetilde{Y}_{i}^{\left(t, R_{1}\right)} & =\sqrt{R_{1}} X_{i}^{*}+\widetilde{Z}_{i} \quad, 1 \leq i \leq n
\end{aligned}\right.
$$

where

$$
S_{\mu}^{\left(t, R_{2}\right)}:=\sqrt{\frac{1-t}{k_{n}}}\left(\mathbf{W X}^{*}\right)_{\mu}+\sqrt{R_{2}} V_{\mu}+\sqrt{t+2 s_{n}-R_{2}} U_{\mu}
$$

The joint posterior density of $\left(\mathbf{X}^{*}, \mathbf{U}\right)$ given $\left(\mathbf{Y}^{\left(t, R_{2}\right)}, \widetilde{\mathbf{Y}}^{\left(t, R_{1}\right)}, \mathbf{W}, \mathbf{V}\right)$ is

$$
\begin{aligned}
&\left.d P\left(\mathbf{x}, \mathbf{u} \mid \mathbf{Y}^{\left(t, R_{2}\right)}, \widetilde{\mathbf{Y}}^{\left(t, R_{1}\right)}, \mathbf{W}, \mathbf{V}\right):=\frac{1}{\mathcal{Z}_{t, R}} \prod_{i=1}^{n} d P_{X}^{(n)}\left(x_{i}\right) e^{-\frac{1}{2}\left(\sqrt{R_{1}} x_{i}-\widetilde{Y}_{i}^{\left(t, R_{1}\right)}\right)}\right)^{2} \\
& \cdot \prod_{\mu=1}^{m_{n}} \frac{d u_{\mu}}{\sqrt{2 \pi}} e^{-\frac{u_{\mu}^{2}}{2}} P_{\mathrm{out}}\left(Y_{\mu}^{\left(t, R_{2}\right)} \mid s_{\mu}^{\left(t, R_{2}\right)}\right),
\end{aligned}
$$

where $\mathcal{Z}_{t, R}$ is a normalization factor and

$$
s_{\mu}^{\left(t, R_{2}\right)}:=\sqrt{\frac{1-t}{k_{n}}}(\mathbf{W} \mathbf{x})_{\mu}+\sqrt{R_{2}} V_{\mu}+\sqrt{t+2 s_{n}-R_{2}} u_{\mu}
$$

We denote by the angular brackets $\langle-\rangle_{n, t, R}$ the expectation with respect to this posterior distribution. We define

$$
F_{1}^{(n)}(t, R):=-\left.2 \frac{\alpha_{n}}{\rho_{n}} \frac{\partial I_{P_{\text {out }}}}{\partial q}\right|_{q=\mathbb{E}\langle Q\rangle_{n, t, R,}, \rho=1}, \quad F_{2}^{(n)}(t, R):=\mathbb{E}\langle Q\rangle_{n, t, R}
$$

where $Q:=\frac{1}{k_{n}} \sum_{i=1}^{n} X_{i}^{*} x_{i}$. We now repeat and prove Proposition 7.12 .

Proposition 7.12. Suppose that (H1), (H2), (H3) hold and $\Delta=\mathbb{E}_{X_{0} \sim P_{0}}\left[X_{0}^{2}\right]=1$. For every $\epsilon \in \mathcal{B}_{n}$, there exists a unique global solution $R(\cdot, \epsilon):[0,1] \rightarrow[0,+\infty)^{2}$ to the initial value problem

$$
\frac{d y}{d t}=\left(F_{1}^{(n)}(t, y), F_{2}^{(n)}(t, y)\right), \quad y(0)=\epsilon .
$$

$R(\cdot, \epsilon)$ is continuously differentiable and the image of its derivative $R^{\prime}(\cdot, \epsilon)$ satisfies

$$
R^{\prime}([0,1], \epsilon) \subseteq\left[0, \frac{\alpha_{n}}{\rho_{n}} r_{\max }\right] \times[0,1]
$$

where $r_{\text {max }}:=-\left.2\left({ }^{\left(\partial I_{\text {Pout }}\right.} / \partial q\right)\right|_{q=1, \rho=1} \geq 0$. Besides, for all $t \in[0,1], R(t, \cdot)$ is a $C^{1}$-diffeomorphism from $\mathcal{B}_{n}$ onto its image whose Jacobian determinant is greater than, or equal to, one, i.e., $\forall \epsilon \in \mathcal{B}_{n}$ :

$$
\operatorname{det} J_{R(t, \cdot)}(\epsilon) \geq 1
$$

where $J_{R(t,)}$ denotes the Jacobian matrix of $R(t, \cdot)$. Finally, the same statement holds true if, for a fixed $r \in\left[0, r_{\max }\right]$, we instead consider the initial value problem

$$
\frac{d y}{d t}=\left(\frac{\alpha_{n}}{\rho_{n}} r, F_{2}^{(n)}(t, y)\right), y(0)=\epsilon .
$$

Proof. We only give the proof for the ODE $d y / d t=\left(F_{1}^{(n)}(t, y), F_{2}^{(n)}(t, y)\right)$ since the one for the ODE $d y / d t=\left(\alpha_{n} r / \rho_{n}, F_{2}^{(n)}(t, y)\right)$ is simpler and follows the same arguments.

By Jensen's inequality and the Nishimori identity (see Lemma 2.1),

$$
\mathbb{E}\langle Q\rangle_{n, t, R}:=\frac{\mathbb{E}\left[\langle\mathbf{x}\rangle_{n, t, R}^{\top} \mathbf{X}^{*}\right]}{k_{n}}=\frac{\mathbb{E}\left\|\langle\mathbf{x}\rangle_{n, t, R}\right\|^{2}}{k_{n}} \leq \frac{\mathbb{E}\left\langle\|\mathbf{x}\|^{2}\right\rangle_{n, t, R}}{k_{n}}=\frac{\mathbb{E}\left\|\mathbf{X}^{*}\right\|^{2}}{k_{n}}=1,
$$

i.e., $\mathbb{E}\langle Q\rangle_{n, t, R} \in[0,1]$. By Lemma 7.22 , the function $q \mapsto I_{P_{\text {out }}}(q, 1)$ is continuously twice differentiable, concave and nonincreasing on $[0,1]$. Therefore, $q \mapsto-\left.2\left(\partial I_{\text {out }} / \partial q\right)\right|_{q, \rho=1}$ is nonnegative and nondecreasing on $[0,1]$, which implies that $-\left.2\left(\partial I_{\text {Pout }} / \partial q\right)\right|_{q, \rho=1} \in\left[0, r_{\text {max }}\right]$. We have thus shown that the domain of definition of the function $F:(t, R) \mapsto\left(F_{1}^{(n)}(t, R), F_{2}^{(n)}(t, R)\right)$ is

$$
\mathcal{D}_{n}:=\left\{\left(t, R_{1}, R_{2}\right) \in[0,1] \times[0,+\infty)^{2}: R_{2} \leq 2 s_{n}+t\right\},
$$

and $F$ takes its values in $\left[0, \alpha_{n} r_{\max } / \rho_{n}\right] \times[0,1]$.
To invoke the Picard-Lindelöf theorem [104. Theorem 1.1], we have to check that $F$ is continuous in $t$ and uniformly Lipschitz continuous in $R$ (meaning that the Lipschitz constant is independent of $t$ ). We can show that $F$ is continuous on $\mathcal{D}_{n}$ and that, for all $t \in[0,1], F(t, \cdot)$ is differentiable on $(0,+\infty) \times\left(0, t+2 s_{n}\right)$ thanks to the standard theorems of continuity and differentiation under the integral sign. The domination hypotheses are indeed verified because (H1), (H2) hold. To
check the uniform Lipschitzness, we show that the Jacobian matrix $J_{F(t,)}(R)$ of $F(t, \cdot)$ is uniformly bounded in $(t, R)$. For all $\left(R_{1}, R_{2}\right) \in(0,+\infty) \times\left(0, t+2 s_{n}\right)$ :

$$
J_{F(t,))}(R)=\left[\begin{array}{cc}
c(t, R) & c(t, R)  \tag{7.193}\\
1 & 1
\end{array}\right]\left[\begin{array}{cc}
\left.\frac{\partial F_{2}^{(n)}}{\partial R_{1}}\right|_{t, R} & 0 \\
0 & \left.\frac{\partial F_{2}^{(n)}}{\partial R_{2}}\right|_{t, R}
\end{array}\right]
$$

where $c(t, R):=-\left.2 \frac{\alpha_{n}}{\rho_{n}} \frac{\partial^{2} I_{P_{\text {out }}}}{\partial q^{2}}\right|_{q=F_{2}^{(n)}(t, R), \rho=1}$ and

$$
\begin{align*}
& \left.\frac{\partial F_{2}^{(n)}}{\partial R_{1}}\right|_{t, R}=\frac{1}{k_{n}} \sum_{i, j=1}^{n} \mathbb{E}\left[\left(\left\langle x_{i} x_{j}\right\rangle_{n, t, R}-\left\langle x_{i}\right\rangle_{n, t, R}\left\langle x_{j}\right\rangle_{n, t, R}\right)^{2}\right]  \tag{7.194}\\
& \left.\frac{\partial F_{2}^{(n)}}{\partial R_{2}}\right|_{t, R}=\frac{1}{k_{n}} \sum_{\mu=1}^{m_{n}} \mathbb{E}\left\|\left\langle\ell_{Y_{\mu}^{(t, R)}}^{\prime}\left(s_{\mu}^{(t, R)}\right) \mathbf{x}\right\rangle_{n, t, R}-\left\langle\ell_{Y_{\mu}^{(t, R)}}^{\prime}\left(s_{\mu}^{(t, R)}\right)\right\rangle_{n, t, R}\langle\mathbf{x}\rangle_{n, t, R}\right\|^{2} \tag{7.195}
\end{align*}
$$

The function $\ell_{y}^{\prime}(\cdot)$ is the derivative of $\ell_{y}: x \mapsto \ln P_{\text {out }}(y \mid x)$. Both $\partial F_{2}^{(n)} / \partial R_{1}$ and $\partial F_{2}^{(n)} / \partial R_{2}$ are clearly nonnegative. Using the assumption (H1), we easily obtain from (7.194) that

$$
\begin{equation*}
0 \leq\left.\frac{\partial F_{2}^{(n)}}{\partial R_{1}}\right|_{t, R} \leq \frac{4 S^{4} n}{\rho_{n}} \tag{7.196}
\end{equation*}
$$

In the proof of Lemma 7.22 , under the hypothesis (H2), we obtain the upper bound (7.108) on $\left|\ell_{y}^{\prime}(x)\right|$. It yields $\forall x \in \mathbb{R}:\left|\ell_{Y_{\mu}^{(t, R)}}^{\prime}(x)\right| \leq\left(2\|\varphi\|_{\infty}+\left|Z_{\mu}\right|\right)\left\|\partial_{x} \varphi\right\|_{\infty}$. Thus, we easily see from (7.195) that

$$
\begin{equation*}
0 \leq\left.\frac{\partial F_{2}^{(n)}}{\partial R_{2}}\right|_{t, R} \leq 8 S^{2}\left(4\|\varphi\|_{\infty}^{2}+1\right)\left\|\partial_{x} \varphi\right\|_{\infty}^{2} \frac{\alpha_{n} n}{\rho_{n}} \tag{7.197}
\end{equation*}
$$

Finally, by Lemma $7.22, q \mapsto-\left.\left(\partial^{2} I_{P_{\text {out }}} / \partial q^{2}\right)\right|_{q, \rho=1}$ is nonnegative continuous on the interval $[0,1]$, so $-\left.\left(\partial^{2} I_{\text {out }} / \partial q^{2}\right)\right|_{q, \rho=1} \in[0, C]$ for some constant $C$ independent of $q$ and $c(t, R) \in\left[0,{ }^{2 C \alpha_{n}} / \rho_{n}\right]$. Combining the latter with (7.193), (7.196) and 7.197) shows that $J_{F(t, \cdot)}(R)$ is uniformly bounded in

$$
(t, R) \in\left\{\left(t, R_{1}, R_{2}\right) \in[0,1] \times(0,+\infty)^{2}: R_{2}<2 s_{n}+t\right\} .
$$

By the mean-value theorem, this implies that $F$ is uniformly Lipschitz continuous in $R$.

By the Picard-Lindelöf theorem [104, Theorem 1.1], for all $\epsilon \in \mathcal{B}_{n}$ there exists a unique solution to the initial value problem $d y / d t=F(t, y), y(0)=\epsilon$ that we denote $R(\cdot, \epsilon):[0, \delta] \rightarrow[0,+\infty)^{2}$. Here $\delta \in[0,1]$ is such that $[0, \delta]$ is the maximal interval of existence of the solution. The function $F$ takes its values in $\left[0, \alpha_{n} r_{\max } / \rho_{n}\right] \times[0,1]$ and $\epsilon \in \mathcal{B}_{n}$ so $\forall t \in[0, \delta]: R(t, \epsilon) \in\left[s_{n}, 2 s_{n}+t \alpha_{n} r_{\max } / \rho_{n}\right] \times\left[s_{n}, 2 s_{n}+t\right]$. It means that $\delta=1$ (the solution never leaves the domain of definition of $F$ ).

Each initial condition $\epsilon \in \mathcal{B}_{n}$ is tied to a unique solution $R(\cdot, \epsilon)$. This implies that the function $\epsilon \mapsto R(t, \epsilon)$ is injective. Its Jacobian determinant is given by Liouville's formula 104, Chapter V, Corollary 3.1]:

$$
\begin{aligned}
\operatorname{det} J_{R(t, \cdot)}(\epsilon) & =\left.\exp \int_{0}^{t} d s\left(\frac{\partial F_{1}^{(n)}}{\partial R_{1}}+\frac{\partial F_{2}^{(n)}}{\partial R_{2}}\right)\right|_{s, R(s, \epsilon)} \\
& =\exp \int_{0}^{t} d s\left(\left.c(s, R(s, \epsilon)) \frac{\partial F_{2}^{(n)}}{\partial R_{1}}\right|_{s, R(s, \epsilon)}+\left.\frac{\partial F_{2}^{(n)}}{\partial R_{2}}\right|_{s, R(s, \epsilon)}\right) .
\end{aligned}
$$

This Jacobian determinant is greater than, or equal to, one since we have shown earlier in the proof that $c(t, R), \partial F_{1}^{(n)} / \partial R_{1}$ and $\partial F_{2}^{(n)} / \partial R_{2}$ are nonnegative. The fact that the Jacobian determinant is bounded away from 0 uniformly in $\epsilon$ implies, by the inverse function theorem, that the injective function $\epsilon \mapsto R(t, \epsilon)$ is a $C^{1}$-diffeomorphism from $\mathcal{B}_{n}$ onto its image.

## Conclusion and possible directions

## 8

In this thesis, we have given precise characterizations, in high-dimensional regimes, of the statistical limits of estimation tasks associated with two ubiquitous statistical models, namely, noisy tensor factorization and generalized linear models. Our approach is information theoretic; we prove exact formulas for the normalized mutual informations associated with these estimation problems and leverage the relationship between mutual information and Bayes optimal inference. Our proofs are based on techniques that were first developed to prove that the replica ansatz for the free energy of large spin systems is exact. In recent years these techniques have been successfully applied to compute high-dimensional limits of normalized mutual informations, the latter being akin to free energies. Remarkable properties specific to the Bayes optimal setting, in particular the Nishimori identity, distinguish these proofs from the ones found in stastical physics. The adaptive interpolation method makes good use of these properties in order to propose a simplified and unified way to prove replica symmetric formulas for the asymptotic normalized mutual information. Prior knowledge of the replica ansatz remains valuable as it guides us towards the correct interpolation scheme.

The adaptive interpolation method was proposed in [37], [87]. Each chapter of the present work extends the method to more general settings. In Chapter 3 we use an adaptive interpolation to demonstrate the exactness of the RS formula even in situation where the prior distribution of the estimated signal is not factorable (the factors $\mathbf{U}$ and $\mathbf{V}$ of the rank-one matrix $\mathbf{U V}^{\top}$ do not have i.i.d. components). The adaptive interpolation relies on the concentration of the order parameter called the overlap, a measure of the correlation between the estimated signal and a sample drawn from the posterior distribution given the observations. In Chapter 3 we apply the adaptive interpolation to situations where the overlap is not a scalar but a matrix of size $K \times K, K$ being the rank of the estimated tensor, and prove a concentration result for this matrix overlap. The latter is not a straightforward extension of existing proofs for the concentration of the scalar overlap. It requires new ideas and technical arguments to make sure that, at a fixed step $t \in[0,1]$ of the interpolation, the interpolation path seen as a function
of the initial condition to the ODE has nice properties (invertibility, regularity, the image is a subset of the positive semidefinite cone appropriately bounded). In Chapters 5 and 6 we illustrate and exploit the modularity of the adaptive interpolation to analyze more complex models. For example, in Chapter 5, we go once more beyond the traditional model where the "spike" that generates the estimated rank-one tensor has i.i.d. components. The model differs from the one of Chapter 3 since the spike is generated by a latent vector. The modularity of the adaptive interpolation allows us to deduce a RS formula for the normalized mutual information by simply reusing the interpolation scheme of the i.i.d. case while the model towards which we interpolate is none other than a one-layer GLM (whose associated normalized mutual information is computed by another adaptive interpolation $[29]$ ). Thanks to this RS formula, we are able to analyze how the MMSE depends on the structure of the spike, that is, the ratio of latent to ambient space dimensions. Finally, in Chapter 7, we study the one-layer GLM of [29] in a high-dimensional regime that is different from the kind of regimes studied in statistical physics and the other chapters of this thesis. The estimated signal is extremely sparse in the sense that the number of nonzero entries is sublinear in its size, and we show that the signal can be effectively recovered with a (large enough) sublinear number of measurements. We prove that the RS formula for the normalized mutual information is still valid in such a high-dimensional limit. We then simplify this sup-inf formula into a minimization problem over a finite set of values. The latter highlights the phenomenology specific to the studied regime; the MMSE takes a finite number of values and sharply transitions from one to the other at values of the sampling rate for which the solution to the minimization problem is not unique.

Let us conclude with some of the challenges that we are faced with when trying to further extend the adaptive interpolation method.

Concentration of the free entropy A step that is common to all the proofs of RS formulas in this thesis is demonstrating that the free entropy concentrates around its expectation at all time $t$ of the interpolation. In most chapters, the proof is straightforward thanks to the usual concentration inequalities given in Chapter 2. However, in Chapters 5 and 6, we study models where a generative prior is stacked over a rank-one tensor factorization problem and a one-layer GLM, respectively, and the proof of the concentration becomes more involved.

Note that in Chapter 6 we prove the RS formula associated with a two-layer GLM by interpolating towards two decoupled channels, whose one of them is a one-layer GLM (see the observations $\widetilde{Y}_{i}^{(t, \epsilon)}$ in (6.19)). In fact, taking for granted the concentration of the free entropy, we can prove by induction the RS formula associated with a $L$-layer GLM with an adaptive interpolation that has a form similar to the one in Chapter 6, except that the one-layer GLM is replaced by a ( $L-1$ )-layer one. Unfortunately, for $L \geq 3$, it is not clear how to prove the concentration of the free entropy by using the same strategy that in Appendix 6.D.

Better concentration of the overlap In all the proofs of this thesis, we bound the variance of the overlap $Q$ by the variance of an auxiliary quantity $\mathcal{L}$. We then prove that the variance $\mathbb{E}\left\langle(\mathcal{L}-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right\rangle$, where the angular brackets $\langle-\rangle$ denote an expectation with respect to the posterior distribution, vanishes if we average it over the initial condition of the interpolation functions (the average smoothens the phase transitions that might appear for some initial conditions). To do so, we show separately that both averages of $\mathbb{E}\left\langle(\mathcal{L}-\langle\mathcal{L}\rangle)^{2}\right\rangle$ and $\mathbb{E}\left\langle(\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right\rangle$ vanish. However, the upper bound on the average of $\mathbb{E}\left\langle(\mathcal{L}-\langle\mathcal{L}\rangle)^{2}\right\rangle$ is stronger than the one on the average of $\mathbb{E}\left\langle(\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right\rangle$; the concentration of $\langle\mathcal{L}\rangle$ around its expectation is the bottleneck to better concentration results for the overlap.

In this thesis we focus on deriving asymptotic formulas for the normalized mutual information, hence we rarely care about the speed of convergence to the limit, except in Chapter 7. In the latter the sparse signal $\mathbf{X}^{*}$ has a sublinear sparsity $k_{n}=\Theta\left(n^{1-\lambda}\right)$ with $\lambda \in(0,1)$, but we need $\lambda \in(0,1 / 9)$ to prove the RS formula. This constraint is due to the upper bound on the average of $\mathbb{E}\left\langle(\langle\mathcal{L}\rangle-\mathbb{E}\langle\mathcal{L}\rangle)^{2}\right\rangle$ in Lemma 7.30. If this upper bound was as strong as the upper bound on the average of $\mathbb{E}\left\langle(\mathcal{L}-\langle\mathcal{L}\rangle)^{2}\right\rangle$ in Lemma 7.29 then we could extend the validity of the RS formula to $\lambda \in(0,1 / 2)$, that is, stronger regimes of sublinear sparsity. In that regard, let us mention that [33] proves the existence of an all-or-nothing phenomenon in sparse linear regression in a regime where $\lambda \in(1 / 2,1)$. It would be interesting to answer whether or not the bottleneck constituted by Lemma 7.30 can be widened.

More general measurement matrices In this thesis we study generalized linear models where the entries of the measurement matrices $\left(\mathbf{W}^{(1)}, \mathbf{W}^{(2)}\right.$ in Chapter 6, $\mathbf{W}$ in Chapter 7) are independent standard Gaussian random variables. There are a lot of interest in proving RS formulas for a wider variety of matrices. For example, the weights of a feedforward neural networks are surely not independent once the weights have been trained, even if they are initialized with independent standard Gaussian random variables. Having formulas that are not restricted to matrices whose entries are independent standard Gaussians would allow us to track how the mutual informations between different layers of the neural network evolve during training (75). As an example, let us consider the linear model

$$
\mathbf{Y}:=\mathbf{W} \mathbf{X}+\sqrt{\Delta} \mathbf{Z},
$$

where $\mathbf{W} \in \mathbb{R}^{m \times n}$ is a random matrix, $\mathbf{X} \in \mathbb{R}^{n}$ a random vector whose entries are i.i.d. with respect to a distribution $P_{X}, \mathbf{Z} \in \mathbb{R}^{m}$ a standard Gaussian random vector, and $\Delta>0$. This model has applications in compressed sensing and code-division multiple access (CDMA) systems. The first replica analysis of this model is due to Tanaka [163] in the context of CDMA (the entries of $\mathbf{X}$ are $\pm 1$ ) and for a matrix $\mathbf{W}$ with entries $W_{i j} \stackrel{\text { i.i.d. }}{\sim} \mathcal{N}(0,1)$. There are now multiple proofs of the corresponding RS formula [44], [73]. Since Tanaka's work, RS formulas for more general matrices have been derived [139], [164] but proving their validity remains a challenge. Suppose that $\mathbf{W}:=\mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^{\top}$ where $\mathbf{U}$ is an orthogonal matrix drawn from the Haar measure on $O(n)$ (the set of $n \times n$ orthogonal matrices), $\mathbf{V}$
an orthogonal matrix drawn from the Haar measure on $O(m)$, and $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times n}$ a diagonal matrix with nonnegative entries. Further assume that the empirical distribution of the diagonal entries of $\boldsymbol{\Sigma}$ converges to a measure $\mu$ in the highdimensional limit $n \rightarrow+\infty, m / n \rightarrow \alpha$. The definition of $\mathbf{W}$ is nothing but a singular value decomposition and thus encompasses a lot of applications. The RS ansatz for the normalized mutual information associated with this model is 164

$$
\begin{aligned}
\lim _{n \rightarrow+\infty} \frac{I(\mathbf{X} ; \mathbf{Y} \mid \mathbf{W})}{n} \underset{q \in\left[0, \mathbb{E} X^{2}\right]}{!?} \inf _{r \geq 0}\left\{I(X ; \sqrt{r} X+Z)+\frac{1}{2} \int_{0}^{\frac{\mathbb{E} X^{2}-q}{\Delta}}\right. & \mathcal{R}_{\mu}(-z) d z \\
& \left.-\frac{r\left(\mathbb{E} X^{2}-q\right)}{2}\right\}
\end{aligned}
$$

where $X \sim P_{X} \perp Z \sim \mathcal{N}(0,1)$ and $\mathcal{R}_{\mu}$ is the $\mathcal{R}$-transform of the measure $\mu$ [165]. There are two hindrances when trying to set up an interpolation for this model. First and foremost, it is not clear towards which channels we should interpolate. In particular, it is not clear what kind of channel we need in order for the primitive of the $\mathcal{R}$-transform to appear. Second, the derivative with respect to $t$ of the normalized mutual information associated with the interpolating model is simplified thanks to a Gaussian integration by parts when $\mathbf{W}$ has independent Gaussian entries. It is of course not possible to take this approach if $\mathbf{W}:=\mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^{\top}$. However, it might be possible to exploit the invariance by rotation of the Haar measure on $O(n)$; see for example Pastur and Vasilchuk's computation of the moments of traces of Haar-distributed matrices [166].

## Bibliography

[1] I. T. Jolliffe, Principal Component Analysis, 2nd ed., ser. Springer Series in Statistics. New York, NY, USA: Springer-Verlag, 2002, ISBN: 978-0-387-95442-4. DOI: $10.1007 / \mathrm{b} 98835$.
[2] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products", Journal of Mathematics and Physics, vol. 6, no. 1-4, pp. 164-189, 1927. DOI: 10.1002 /sapm192761164
[3] L. R. Tucker, "Some mathematical notes on three-mode factor analysis", Psychometrika, vol. 31, no. 3, pp. 279-311, Sep. 1966. DOI: 10.1007 bf02289464.
[4] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of "eckart-young" decomposition", Psychometrika, vol. 35, no. 3, pp. 283-319, Sep. 1970. DOI: 10.1007/bf02310791.
[5] C. J. Appellof and E. R. Davidson, "Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents", Analytical Chemistry, vol. 53, no. 13, pp. 2053-2056, Nov. 1981. DOI: 10.1021 ac00236a025.
[6] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning", IEEE Transactions on Signal Processing, vol. 65, no. 13, pp. 35513582, Jul. 2017, ISSN: 1941-0476. DOI: 10.1109/tsp.2017.2690524.
[7] M. Mørup, "Applications of tensor (multiway array) factorizations and decompositions in data mining", WIREs Data Mining and Knowledge Discovery, vol. 1, no. 1, pp. 24-40, Jan. 2011. DOI: $10.1002 /$ widm.1.
[8] A. Anandkumar, D. Hsu, and S. M. Kakade, "A method of moments for mixture models and hidden Markov models", (Edinburgh, Scotland), S. Mannor, N. Srebro, and R. C. Williamson, Eds., ser. Proceedings of Machine Learning Research, vol. 23, JMLR Workshop and Conference Proceedings, 2012, pp. 33.1-33.34.
[9] A. Anandkumar, R. Ge, D. Hsu, S. M. Kakade, and M. Telgarsky, "Tensor decompositions for learning latent variable models", J. Mach. Learn. Res., vol. 15, no. 1, pp. 2773-2832, Jan. 2014, ISSN: 1532-4435.
[10] A. Anandkumar, R. Ge, D. Hsu, and S. M. Kakade, "A tensor approach to learning mixed membership community models", J. Mach. Learn. Res., vol. 15, no. 1, pp. 2239-2312, Jan. 2014, ISSN: 1532-4435.
$[11]$ L. Xiong, X. Chen, T.-K. Huang, J. Schneider, and J. G. Carbonell, "Temporal collaborative filtering with Bayesian probabilistic tensor factorization", in Proceedings of the 2010 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, Apr. 2010. Doi: 10.1137/1.9781611972801.19.
[12] O. Duchenne, F. Bach, I. Kweon, and J. Ponce, "A tensor-based algorithm for high-order graph matching", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 12, pp. 2383-2395, Dec. 2011, ISSN: 1939-3539. DOI: 10.1109/TPAMI.2011.110.
[13] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications", SIAM Review, vol. 51, no. 3, pp. 455-500, 2009. DOi: 10.1137/07070111X.
[14] S. Rabanser, O. Shchur, and S. Günnemann, Introduction to tensor decompositions and their applications in machine learning, 2017. arXiv: 1711.10781 [stat.ML].
[15] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank", Psychometrika, vol. 1, no. 3, pp. 211-218, Sep. 1936. Doi: 10.1007/bf02288367.
[16] V. de Silva and L.-H. Lim, "Tensor rank and the ill-posedness of the best low-rank approximation problem", SIAM J. Matrix Anal. Appl., vol. 30, no. 3, pp. 1084-1127, Jan. 2008. Doi: 10.1137/06066518x.
[17] C. J. Hillar and L.-H. Lim, "Most tensor problems are NP-hard", J. ACM, vol. 60, no. 6, Nov. 2013, ISSN: 0004-5411. Doi: $10.1145 / 2512329$
[18] Y. Deshpande, E. Abbe, and A. Montanari, "Asymptotic mutual information for the balanced binary stochastic block model", Information and Inference: A Journal of the IMA, vol. 6, no. 2, pp. 125-170, Dec. 2016, ISSN: 2049-8764. DOI: 10.1093/imaiai/iaw017.
[19] E. Richard and A. Montanari, "A statistical model for tensor PCA", in Advances in Neural Information Processing Systems 27 (NIPS 2014), Red Hook, NY, USA: Curran Associates, 2014, pp. 2897-2905.
[20] H. Derksen, "Kruskal's uniqueness inequality is sharp", Linear Algebra Appl., vol. 438, no. 2, pp. 708-712, Jan. 2013. Doi: 10.1016/j.laa.2011.05.041.
[21] F. Krzakala, J. Xu, and L. Zdeborová, "Mutual information in rank-one matrix estimation", in 2016 IEEE Information Theory Workshop (ITW), Sep. 2016, pp. 71-75. DOI: 10.1109/ITW.2016.7606798.
[22] J. R. Fienup, "Phase retrieval algorithms: a comparison", Appl. Opt., vol. 21, no. 15, pp. 2758-2769, Aug. 1982. DOI: 10.1364/AO.21.002758.
[23] R. P. Millane, "Phase retrieval in crystallography and optics", J. Opt. Soc. Am. A, vol. 7, no. 3, pp. 394-411, Mar. 1990. Doi: 10.1364/JOSAA.7.000394.
[24] U. S. Kamilov, V. K. Goyal, and S. Rangan, "Message-passing de-quantization with applications to compressed sensing", IEEE Transactions on Signal Processing, vol. 60, no. 12, pp. 6270-6281, Dec. 2012, ISSN: 1941-0476. DOI: 10.1109/TSP.2012.2217334.
[25] P. T. Boufounos and R. G. Baraniuk, "1-bit compressive sensing", in 2008 42nd Annual Conference on Information Sciences and Systems, Mar. 2008, pp. 16-21. DOI: 10.1109/CISS.2008.4558487.
[26] J. A. Nelder and R. W. M. Wedderburn, "Generalized linear models", Journal of the Royal Statistical Society. Series A (General), vol. 135, no. 3, pp. 370-384, 1972, ISSN: 0035-9238. DOI: $10.2307 / 2344614$
[27] A. Manoel, F. Krzakala, M. Mézard, and L. Zdeborová, "Multi-layer generalized linear estimation", in 2017 IEEE International Symposium on Information Theory (ISIT), 2017, pp. 2098-2102. DOI: 10.1109/ISIT.2017.8006899.
[28] A. K. Fletcher, S. Rangan, and P. Schniter, "Inference in deep networks in high dimensions", in 2018 IEEE International Symposium on Information Theory (ISIT), 2018, pp. 1884-1888. DOI: 10.1109/ISIT.2018.8437792.
[29] J. Barbier, F. Krzakala, N. Macris, L. Miolane, and L. Zdeborová, "Optimal errors and phase transitions in high-dimensional generalized linear models", Proceedings of the National Academy of Sciences, vol. 116, no. 12, pp. 54515460, 2019, ISSN: 0027-8424. DOI: 10.1073/pnas. 1802705116
[30] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements", Communications on Pure and Applied Mathematics, vol. 59, no. 8, pp. 1207-1223, 2006. DOI: 10.1002 cpa. 20124.
$[31]$ D. L. Donoho, "Compressed sensing", IEEE Transactions on Information Theory, vol. 52, no. 4, pp. 1289-1306, Apr. 2006, ISSN: 1557-9654. DOI: 10.1109/TIT.2006.871582.
[32] D. Gamarnik and I. Zadik, "High dimensional regression with binary coefficients. Estimating squared error and a phase transtition", in Proceedings of the 2017 Conference on Learning Theory, S. Kale and O. Shamir, Eds., ser. Proceedings of Machine Learning Research, vol. 65, Amsterdam, Netherlands: PMLR, 2017, pp. 948-953.
[33] G. Reeves, J. Xu, and I. Zadik, "The all-or-nothing phenomenon in sparse linear regression", in Proceedings of the Thirty-Second Conference on Learning Theory, A. Beygelzimer and D. Hsu, Eds., ser. Proceedings of Machine Learning Research, vol. 99, Phoenix, USA: PMLR, 2019, pp. 2652-2663.
[34] J. Barbier, M. Dia, N. Macris, F. Krzakala, T. Lesieur, and L. Zdeborová, "Mutual information for symmetric rank-one matrix estimation: a proof of the replica formula", in Advances in Neural Information Processing Systems 29 (NIPS 2016), Red Hook, NY, USA: Curran Associates, 2016, pp. 424-432.
[35] A. E. Alaoui and F. Krzakala, "Estimation in the spiked Wigner model: a short proof of the replica formula", in 2018 IEEE International Symposium on Information Theory (ISIT), 2018, pp. 1874-1878. DOI: 10.1109/ISIT 2018.8437810 .
[36] M. Lelarge and L. Miolane, "Fundamental limits of symmetric low-rank matrix estimation", Probability Theory and Related Fields, vol. 173, no. 3, pp. 859-929, Apr. 1, 2019, ISSN: 1432-2064. Dor: 10.1007/s00440-018-0845-区.
[37] J. Barbier and N. Macris, "The adaptive interpolation method for proving replica formulas. Applications to the Curie-Weiss and Wigner spike models", Journal of Physics A: Mathematical and Theoretical, vol. 52, no. 29, p. 294002, Jun. 2019. DOI: 10.1088/1751-8121/ab2735.
[38] C. Robert and G. Casella, Monte Carlo Statistical Methods, 2nd ed., ser. Springer Texts in Statistics. New York, NY, USA: Springer-Verlag, 2004, ISBN: 978-1-4757-4145-2. DOI: 10.1007/978-1-4757-4145-2.
[39] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines", The Journal of Chemical Physics, vol. 21, no. 6, pp. 1087-1092, 1953. Doi: 10.1063/1.1699114.
[40] C. Robert and G. Casella, "A short history of Markov chain Monte Carlo: subjective recollections from incomplete data", Statistical Science, vol. 26, no. 1, pp. 102-115, 2011, ISSN: 0883-4237. [Online]. Available: http://www. jstor.org/stable/23059158.
[41] S. F. Edwards and P. W. Anderson, "Theory of spin glasses", Journal of Physics F: Metal Physics, vol. 5, no. 5, pp. 965-974, May 1975. Doi: 10.1088/0305-4608/5/5/017.
[42] M. Mezard, G. Parisi, and M. Virasoro, Spin Glass Theory and Beyond. World Scientific, 1986. DOI: $10.1142 / 0271$.
[43] A. Giurgiu, N. Macris, and R. Urbanke, "Spatial coupling as a proof technique and three applications", IEEE Transactions on Information Theory, vol. 62, no. 10, pp. 5281-5295, Oct. 2016, ISSN: 1557-9654. DOI: 10.1109/TIT.2016.2539144.
[44] J. Barbier, M. Dia, N. Macris, and F. Krzakala, "The mutual information in random linear estimation", in 2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Sep. 2016, pp. 625632. DOI: 10.1109/ALLERTON.2016.7852290.
[45] T. Lesieur, L. Miolane, M. Lelarge, F. Krzakala, and L. Zdeborová, "Statistical and computational phase transitions in spiked tensor estimation", in 2017 IEEE International Symposium on Information Theory (ISIT), 2017, pp. 511-515. DOI: 10.1109/ISIT.2017.8006580.
[46] L. Miolane, Fundamental limits of low-rank matrix estimation: the nonsymmetric case, 2017. arXiv: 1702.00473 [math.PR].
[47] A. Coja-Oghlan, F. Krzakala, W. Perkins, and L. Zdeborova, "Informationtheoretic thresholds from the cavity method", in Proceedings of the 49 th Annual ACM SIGACT Symposium on Theory of Computing, ser. STOC 2017, Montreal, Canada: Association for Computing Machinery, 2017, pp. 146-157, ISBN: 9781450345286 . DOI: $10.1145 / 3055399.3055420$.
[48] D. L. Donoho, "High-dimensional data analysis: the curses and blessings of dimensionality", in AMS Conference on Mathematical Challenges of the 21st century, 2000.
[49] A. N. Gorban and I. Y. Tyukin, "Blessing of dimensionality: mathematical foundations of the statistical physics of data", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 376, no. 2118, p. 20170 237, 2018. DOI: 10.1098/rsta.2017.0237.
[50] C. E. Shannon, "A mathematical theory of communication", Bell System Technical Journal, vol. 27, no. 3, pp. 379-423, 1948. DOI: 10.1002/j.15387305.1948.tb01338.x.
[51] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum meansquare error in Gaussian channels", IEEE Transactions on Information Theory, vol. 51, no. 4, pp. 1261-1282, Apr. 2005, ISSN: 1557-9654. DOI: 10.1109/TIT.2005.844072
[52] R. B. Griffiths, "A proof that the free energy of a spin system is extensive", Journal of Mathematical Physics, vol. 5, no. 9, pp. 1215-1222, 1964. DOI: 10.1063/1.1704228.
[53] M. Talagrand, Mean Field Models for Spin Glasses: Volume I: Basic Examples, ser. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. Springer-Verlag Berlin Heidelberg, 2011, vol. 54, ISBN: 978-3-642-15202-3. DOI: 10.1007/978-3-642-15202-3.
[54] S. Edwards. "Citation Classic commentaries, Theory of spin-glasses". (1985), [Online]. Available: http:/ / garfield. library. upenn.edu / classics1985 / A1985APZ5400001.pdf (visited on 10/31/2020).
[55] D. Sherrington and S. Kirkpatrick, "Solvable model of a spin-glass", Physical Review Letters, vol. 35, pp. 1792-1796, 26 Dec. 1975. DOI: 10.1103 PhysRevLett.35.1792.
[56] G. Parisi, "Infinite number of order parameters for spin-glasses", Phys. Rev. Lett., vol. 43, pp. 1754-1756, 23 Dec. 1979. Doi: 10.1103/PhysRevLett. 43. 1754.
[57] ——, "A sequence of approximated solutions to the s-k model for spin glasses", Journal of Physics A: Mathematical and General, vol. 13, no. 4, pp. L115-L121, Apr. 1980. Doi: 10.1088/0305-4470/13/4/009.
[58] D. Panchenko, The Sherrington-Kirkpatrick Model, ser. Springer Monographs in Mathematics. Springer-Verlag New York, 2013, ISBN: 978-1-4614-6289-7. DOI: $10.1007 / 978-1-4614-6289-7$.
[59] M. Talagrand, "The parisi formula", Annals of Mathematics, vol. 163, no. 1, pp. 221-263, Jan. 2006. DOI: 10.4007/annals.2006.163.221.
[60] F. Guerra and F. L. Toninelli, "The thermodynamic limit in mean field spin glass models", Communications in Mathematical Physics, vol. 230, no. 1, pp. 71-79, Sep. 1, 2002, ISSN: 1432-0916. DOI: 10.1007/s00220-002-0699-y.
[61] F. Guerra, "Broken replica symmetry bounds in the mean field spin glass model", Communications in Mathematical Physics, vol. 233, no. 1, pp. 1-12, Feb. 1, 2003, ISSN: 1432-0916. Doi: 10.1007/s00220-002-0773-5.
[62] N. Sourlas, "Spin-glass models as error-correcting codes", Nature, vol. 339, no. 6227, pp. 693-695, Jun. 1989. Doi: 10.1038/339693a0.
[63] T. Murayama, Y. Kabashima, D. Saad, and R. Vicente, "Statistical physics of regular low-density parity-check error-correcting codes", Phys. Rev. E, vol. 62, pp. 1577-1591, 2 Aug. 2000. Doi: 10.1103/PhysRevE.62.1577.
[64] H. Nishimori, Statistical Physics of Spin Glasses and Information Processing: An Introduction. New York, NY, USA: Oxford University Press, Jul. 19, 2001, ISBN: 9780198509417. DOI: 10.1093/acprof:oso/9780198509417.001. 0001.
[65] S. Franz, M. Leone, A. Montanari, and F. Ricci-Tersenghi, "Dynamic phase transition for decoding algorithms", Phys. Rev. E, vol. 66, p. 046 120, 4 Oct. 2002. DOI: $10.1103 /$ PhysRevE.66.046120.
[66] S. B. Korada and N. Macris, "Exact solution of the gauge symmetric p-spin glass model on a complete graph", Journal of Statistical Physics, vol. 136, no. 2, pp. 205-230, Jul. 1, 2009, ISSN: 1572-9613. DOI: 10.1007/s10955-009-9781-6.
[67] T. Lesieur, F. Krzakala, and L. Zdeborová, "MMSE of probabilistic lowrank matrix estimation: universality with respect to the output channel", in 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton), Sep. 2015, pp. 680-687. DoI: 10.1109/ALLERTON. 2015.7447070 .
[68] P. Milgrom and I. Segal, "Envelope theorems for arbitrary choice sets", Econometrica, vol. 70, no. 2, pp. 583-601, 2002. DOI: $10.1111 / 1468-$ 0262.00296
[69] J. Barbier, N. Macris, and L. Miolane, The layered structure of tensor estimation and its mutual information, 2017. arXiv: 1709.10368 [cs.IT].
[70] C. Luneau, J. Barbier, and N. Macris, "Mutual information for low-rank even-order symmetric tensor estimation", Information and Inference: A Journal of the IMA, Sep. 2020, iaaa022, ISSN: 2049-8772. DOI: 10.1093 imaiai/iaaa022.
[71] C. Luneau, N. Macris, and J. Barbier, "High-dimensional rank-one nonsymmetric matrix decomposition: the spherical case", in 2020 IEEE International Symposium on Information Theory (ISIT), Jun. 2020, pp. 2646-2651. DOI: 10.1109/ISIT44484.2020.9174104.
[72] J. Barbier and G. Reeves, "Information-theoretic limits of a multiview low-rank symmetric spiked matrix model", in 2020 IEEE International Symposium on Information Theory (ISIT), Jun. 2020, pp. 2771-2776. DOI: 10.1109/ISIT44484.2020.9173970.
[73] G. Reeves and H. D. Pfister, "The replica-symmetric prediction for random linear estimation with Gaussian matrices is exact", IEEE Transactions on Information Theory, vol. 65, no. 4, pp. 2252-2283, Apr. 2019, ISSN: 1557-9654. DOI: 10.1109/TIT.2019.2891664.
[74] J. Barbier, N. Macris, A. Maillard, and F. Krzakala, "The mutual information in random linear estimation beyond i.i.d. matrices", in 2018 IEEE International Symposium on Information Theory (ISIT), 2018, pp. 13901394. DOI: 10.1109/ISIT.2018.8437522.
[75] M. Gabrié, A. Manoel, C. Luneau, J. Barbier, N. Macris, F. Krzakala, and L. Zdeborová, "Entropy and mutual information in models of deep neural networks", Journal of Statistical Mechanics: Theory and Experiment, vol. 2019, no. 12, p. 124014 , Dec. 2019. Doi: 10.1088/1742-5468/ab3430.
[76] L. Zdeborová and F. Krzakala, "Statistical physics of inference: thresholds and algorithms", Advances in Physics, vol. 65, no. 5, pp. 453-552, 2016. DOI: $10.1080 / 00018732.2016 .1211393$.
[77] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing", Proceedings of the National Academy of Sciences, vol. 106, no. 45, pp. 18914-18 919, 2009, ISSN: 0027-8424. DOI: 10.1073 pnas. 0909892106
[78] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988, ISBN: 1558604790.
[79] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing", IEEE Transactions on Information Theory, vol. 57, no. 2, pp. 764-785, Feb. 2011, ISSN: 1557-9654. DOI: 10.1109/TIT.2010.2094817.
[80] A. K. Fletcher and S. Rangan, "Iterative reconstruction of rank-one matrices in noise", Information and Inference: A Journal of the IMA, vol. 7, no. 3, pp. 531-562, Jan. 2018, ISSN: 2049-8764. DOI: 10.1093/imaiai/iax014.
[81] C. Rush and R. Venkataramanan, "Finite sample analysis of approximate message passing algorithms", IEEE Transactions on Information Theory, vol. 64, no. 11, pp. 7264-7286, Nov. 2018, ISSN: 1557-9654. DOI: 10.1109 TIT.2018.2816681.
[82] C. Cademartori and C. Rush, "Exponentially fast concentration of vector approximate message passing to its state evolution", in 2020 IEEE International Symposium on Information Theory (ISIT), Jun. 2020, pp. 2670-2675. DOI: 10.1109/ISIT44484.2020.9174146.
[83] T. Lesieur, F. Krzakala, and L. Zdeborová, "Phase transitions in sparse pca", in 2015 IEEE International Symposium on Information Theory (ISIT), Jun. 2015, pp. 1635-1639. DOI: 10.1109/ISIT.2015.7282733.
[84] J. Kadmon and S. Ganguli, "Statistical mechanics of low-rank tensor decomposition", Journal of Statistical Mechanics: Theory and Experiment, vol. 2019, no. 12, p. 124016 , Dec. 2019. Doi: 10.1088/1742-5468/ab3216.
[85] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing", in 2011 IEEE International Symposium on Information Theory (ISIT), 2011, pp. 2168-2172. DOi: 10.1109/ISIT.2011.6033942.
[86] M. Aizenman, R. Sims, and S. L. Starr, "Extended variational principle for the sherrington-kirkpatrick spin-glass model", Phys. Rev. B, vol. 68, p. 214403,21 Dec. 2003. Doi: $10.1103 /$ PhysRevB.68.214403.
[87] J. Barbier and N. Macris, "The adaptive interpolation method: a simple scheme to prove replica formulas in Bayesian inference", Probability Theory and Related Fields, vol. 174, no. 3, pp. 1133-1185, Aug. 1, 2019, ISSN: 1432-2064. DOI: $10.1007 / \mathrm{s} 00440-018-0879-0$.
[88] G. Genovese and D. Tantari, "Legendre duality of spherical and Gaussian spin glasses", Mathematical Physics, Analysis and Geometry, vol. 18, no. 1, p. 10, Apr. 15, 2015, ISSN: 1572-9656. DOI: 10.1007/s11040-015-9181-x.
[89] J. Barbier, C. Luneau, and N. Macris, "Mutual information for low-rank even-order symmetric tensor factorization", in 2019 IEEE Information Theory Workshop (ITW), Aug. 2019, pp. 1-5. DOI: 10.1109/ITW44776. 2019.8989408 .
[90] C. Luneau and N. Macris, "Tensor estimation with structured priors", IEEE Journal on Selected Areas in Information Theory, vol. 1, no. 3, pp. 705-722, Nov. 2020, ISSN: 2641-8770. DOI: 10.1109/JSAIT.2020.3040336.
[91] C. Luneau, J. Barbier, and N. Macris, "Information theoretic limits of learning a sparse rule", in Advances in Neural Information Processing Systems 33 (NeurIPS 2020), 2020.
[92] S. Chatterjee. "Stein's method and applications, lecture notes scribed by David Rosenberg." (2007), [Online]. Available: https://statweb.stanford. edu/ ~souravc/AllLectures.pdf (visited on 10/31/2020).
[93] S. Boucheron, G. Lugosi, and P. Massart, Concentration Inequalities: A Nonasymptotic Theory of Independence. London, U.K.: Oxford Univ. Press, 2013, ISBN: 9780199535255.
[94] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. PHAN, "Tensor decompositions for signal processing applications: from two-way to multiway component analysis", IEEE Signal Processing Magazine, vol. 32, no. 2, pp. 145-163, Mar. 2015, ISSN: 1558-0792. DOI: 10.1109/MSP.2013.2297439.
[95] D. Féral and S. Péché, "The largest eigenvalue of rank one deformation of large Wigner matrices", Communications in Mathematical Physics, vol. 272, no. 1, pp. 185-228, May 1, 2007, ISSN: 1432-0916. DOI: $10.1007 / \mathrm{s} 00220-$ 007-0209-3.
[96] F. Benaych-Georges and R. R. Nadakuditi, "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices", Advances in Mathematics, vol. 227, no. 1, pp. 494-521, 2011, ISSN: 0001-8708. DOI: 10.1016/j.aim.2011.02.007.
[97] T. Lesieur, F. Krzakala, and L. Zdeborová, "Constrained low-rank matrix estimation: phase transitions, approximate message passing and applications", Journal of Statistical Mechanics: Theory and Experiment, vol. 2017, no. 7, p. 073 403, Jul. 2017. Doi: 10.1088/1742-5468/aa7284.
[98] B. Aubin, B. Loureiro, A. Maillard, F. Krzakala, and L. Zdeborová, "The spiked matrix model with generative priors", in Advances in Neural Information Processing Systems 32 (NIPS 2019), Red Hook, NY, USA: Curran Associates, 2019, pp. 8366-8377.
[99] R. Vershynin, High-Dimensional Probability: An Introduction with Applications in Data Science, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018. Doi: $10.1017 /$ 9781108231596.
[100] J. M. Kosterlitz, D. J. Thouless, and R. C. Jones, "Spherical model of a spin-glass", Physical Review Letters, vol. 36, pp. 1217-1220, 20 May 1976. DOI: 10.1103/PhysRevLett.36.1217.
[101] J. Ginibre, "Statistical ensembles of complex, quaternion, and real matrices", Journal of Mathematical Physics, vol. 6, no. 3, pp. 440-449, 1965. Doi: 10.1063/1.1704292.
[102] T. H. Berlin and M. Kac, "The spherical model of a ferromagnet", Physical Review, vol. 86, pp. 821-835, 6 Jun. 1952. DOI: 10.1103/PhysRev.86.821.
[103] A. Barra, G. Genovese, F. Guerra, and D. Tantari, "About a solvable mean field model of a Gaussian spin glass", Journal of Physics A: Mathematical and Theoretical, vol. 47, no. 15, p. 155 002, Mar. 2014. Doi: $10.1088 / 1751-$ 8113/47/15/155002.
[104] P. Hartman, Ordinary Differential Equations, 2nd ed. Philadelphia, PA, USA: SIAM, 2002, ISBN: 978-0-89871-922-2. DOI: 10.1137/1.9780898719222.
[105] E. S. Meckes, The Random Matrix Theory of the Classical Compact Groups, ser. Cambridge Tracts in Mathematics. Cambridge University Press, 2018. DOI: $10.1017 / 9781108303453$.
[106] J.-C. Mourrat, Hamilton-Jacobi equations for mean-field disordered systems, 2018. arXiv: 1811.01432 [math.PR].
[107] J. Schur, "Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen", Journal für die reine und angewandte Mathematik, vol. 140, pp. 1-28, 1911.
[108] R. A. Horn and C. R. Johnson, Matrix Analysis, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2012. Doi: $10.1017 / 9781139020411$.
[109] J. Barbier, "Overlap matrix concentration in optimal Bayesian inference", Information and Inference: A Journal of the IMA, May 2020, iaaa008, ISSN: 2049-8772. DOI: 10.1093/imaiai/iaaa008.
[110] P. D. Moral and A. Niclas, "A Taylor expansion of the square root matrix function", Journal of Mathematical Analysis and Applications, vol. 465, no. 1, pp. 259-266, 2018, ISSN: 0022-247X. DOI: 10.1016/j.jmaa.2018.05.005.
[111] J. M. Varah, "A lower bound for the smallest singular value of a matrix", Linear Algebra and its Applications, vol. 11, no. 1, pp. 3-5, 1975, ISSN: 0024-3795. DOI: 10.1016/0024-3795(75)90112-3.
[112] S. Mallat, A Wavelet Tour of Signal Processing, 3rd ed. Boston: Academic Press, 2009, ISBN: 978-0-12-374370-1. DOI: $10.1016 /$ B978-0-12-374370-1.00006-9.
[113] A. Mousavi, A. B. Patel, and R. G. Baraniuk, "A deep learning approach to structured signal recovery", in 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton), Sep. 2015, pp. 13361343. DOI: 10.1109/ALLERTON.2015.7447163.
[114] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets", in Advances in Neural Information Processing Systems 27 (NIPS 2014), Red Hook, NY, USA: Curran Associates, 2014, pp. 2672-2680.
[115] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks", Science, vol. 313, no. 5786, pp. 504-507, 2006, ISSN: 0036-8075. DOI: $10.1126 /$ science. 1127647 ,
[116] A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed sensing using generative models", in Proceedings of the 34th International Conference on Machine Learning - Volume 70, ser. ICML'17, Sydney, NSW, Australia: JMLR.org, 2017, pp. 537-546.
[117] R. Heckel and P. Hand, "Deep decoder: concise image representations from untrained non-convolutional networks", in 7th International Conference on Learning Representations (ICLR), 2019.
[118] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep image prior", in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Jun. 2018, pp. 9446-9454. DOI: 10.1109/CVPR.2018.00984.
[119] D. V. Veen, A. Jalal, M. Soltanolkotabi, E. Price, S. Vishwanath, and A. G. Dimakis, Compressed sensing with deep image prior and learned regularization, 2019. arXiv: 1806.06438 [stat.ML].
[120] P. Hand and V. Voroninski, "Global guarantees for enforcing deep generative priors by empirical risk", IEEE Transactions on Information Theory, vol. 66, no. 1, pp. 401-418, Jan. 2020, ISSN: 1557-9654. DOI: 10.1109 /TIT. 2019. 2935447.
[121] R. Heckel, W. Huang, P. Hand, and V. Voroninski, "Rate-optimal denoising with deep neural networks", Information and Inference: A Journal of the IMA, Jun. 2020, iaaa011, ISSN: 2049-8772. DOI: 10.1093/imaiai/iaaa011.
[122] P. Hand, O. Leong, and V. Voroninski, "Phase retrieval under a generative prior", in Advances in Neural Information Processing Systems 31 (NIPS 2018), Red Hook, NY, USA: Curran Associates, 2018, pp. 9136-9146.
[123] D. G. Mixon and S. Villar, SUNLayer: Stable denoising with generative networks, 2018. arXiv: 1803.09319 [cs.LG].
[124] P. McCullagh and J. A. Nelder, Generalized Linear Models, 2nd, ser. Chapman \& Hall/CRC Monographs on Statistics and Applied Probability. London, UK: Chapman \& Hall, 1989, ISBN: 9780412317606.
[125] F. Guerra and F. L. Toninelli, "Quadratic replica coupling in the Sher-rington-Kirkpatrick mean field spin glass model", Journal of Mathematical Physics, vol. 43, no. 7, pp. 3704-3716, 2002. Doi: 10.1063/1.1483378.
[126] S. B. Hopkins, J. Shi, and D. Steurer, "Tensor principal component analysis via sum-of-square proofs", in Proceedings of the 28th Conference on Learning Theory, vol. 40, Jul. 2015, pp. 956-1006.
[127] A. Zhang and D. Xia, "Tensor SVD: statistical and computational limits", IEEE Transactions on Information Theory, vol. 64, no. 11, pp. 7311-7338, Nov. 2018, ISSN: 1557-9654. DOI: 10.1109/TIT.2018.2841377.
[128] A. S. Wein, A. El Alaoui, and C. Moore, "The Kikuchi hierarchy and tensor PCA", in 2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS), Nov. 2019, pp. 1446-1468. DOI: 10.1109/FOCS.2019.000-2.
[129] M. Brennan and G. Bresler, "Reducibility and statistical-computational gaps from secret leakage", in Proceedings of Machine Learning Research, vol. 125, PMLR, 2020, pp. 648-847.
[130] Y. Luo and A. R. Zhang, Tensor clustering with planted structures: statistical optimality and computational limits, 2020. arXiv: 2005.10743 [math.ST].
[131] L. Paninski, "Estimation of entropy and mutual information", Neural Computation, vol. 15, no. 6, pp. 1191-1253, Jun. 2003. DOI: 10.1162 089976603321780272 .
[132] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method", in 37th annual Allerton Conference on Communication, Control, and Computing, 1999, pp. 368-377.
[133] N. Tishby and N. Zaslavsky, "Deep learning and the information bottleneck principle", in 2015 IEEE Information Theory Workshop (ITW), Apr. 2015, pp. 1-5. DOI: $10.1109 /$ ITW.2015.7133169.
[134] R. Shwartz-Ziv and N. Tishby, Opening the black box of deep neural networks via information, 2017. arXiv: 1703.00810 [cs.LG].
[135] A. Alemi, I. Fischer, J. Dillon, and K. Murphy, "Deep variational information bottleneck", in 5th International Conference on Learning Representations, 2017.
[136] M. I. Belghazi, A. Baratin, S. Rajeshwar, S. Ozair, Y. Bengio, A. Courville, and D. Hjelm, "Mutual information neural estimation", in Proceedings of the 35th International Conference on Machine Learning, vol. 80, PMLR, 2018, pp. 531-540.
[137] A. Kolchinsky, B. D. Tracey, and D. H. Wolpert, "Nonlinear information bottleneck", Entropy, vol. 21, no. 12, p. 1181, 2019. Dor: 10.3390/e21121181.
[138] A. M. Saxe, Y. Bansal, J. Dapello, M. Advani, A. Kolchinsky, B. D. Tracey, and D. D. Cox, "On the information bottleneck theory of deep learning", Journal of Statistical Mechanics: Theory and Experiment, vol. 2019, no. 12, p. 124020, Dec. 2019. DOI: 10.1088/1742-5468/ab3985.
[139] Y. Kabashima, "Inference from correlated patterns: a unified theory for perceptron learning and linear vector channels", Journal of Physics: Conference Series, vol. 95, p. 012 001, Jan. 2008. DOI: 10.1088/1742-6596/95/1/012001.
[140] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", Nature, vol. 521, no. 7553 , pp. 436-444, May 1, 2015, ISSN: 1476-4687. DOI: 10.1038 nature14539.
[141] P. Bühlmann and S. Van De Geer, Statistics for High-Dimensional Data: Methods, Theory and Applications, ser. Springer Series in Statistics. SpringerVerlag Berlin Heidelberg, 2011, ISBN: 978-3-642-20192-9. DOI: 10.1007/978-3-642-20192-9.
[142] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?", IEEE Transactions on Information Theory, vol. 52, no. 12, pp. 5406-5425, Dec. 2006, ISSN: 1557-9654. DOI: 10.1109/TIT.2006.885507.
[143] A. Engel and C. Van den Broeck, Statistical Mechanics of Learning. Cambridge University Press, 2001. DOI: 10.1017/CBO9781139164542.
[144] M. Mezard and A. Montanari, Information, Physics, and Computation. New York, NY, USA: Oxford University Press, 2009, ISBN: 978-0-19-857083-7. DOI: 10.1093/acprof:oso/9780198570837.001.0001.
[145] F. Guerra, "An introduction to mean field spin glas theory: methods and results", in Mathematical Statistical Physics, ser. Les Houches, A. Bovier, F. Dunlop, A. van Enter, F. den Hollander, and J. Dalibard, Eds., vol. 83, Elsevier, 2006, pp. 243-271. DOI: 10.1016/S0924-8099(06)80042-9.
[146] A. Montanari, "Tight bounds for LDPC and LDGM codes under MAP decoding", IEEE Transactions on Information Theory, vol. 51, no. 9, pp. 32213246, Sep. 2005, ISSN: 1557-9654. DOI: 10.1109/TIT.2005.853320.
[147] N. Macris, "Griffith-Kelly-Sherman correlation inequalities: a useful tool in the theory of error correcting codes", IEEE Transactions on Information Theory, vol. 53, no. 2, pp. 664-683, Feb. 2007, ISSN: 1557-9654. DOI: 10. 1109/TIT.2006.889002.
[148] ——, "Sharp bounds on generalized EXIT functions", IEEE Transactions on Information Theory, vol. 53, no. 7, pp. 2365-2375, Jul. 2007, ISSN: 1557-9654. DOI: 10.1109/TIT.2007.899536
[149] S. Kudekar and N. Macris, "Sharp bounds for optimal decoding of lowdensity parity-check codes", IEEE Transactions on Information Theory, vol. 55, no. 10, pp. 4635-4650, Oct. 2009, ISSN: 1557-9654. DOI: 10.1109 TIT.2009.2027523.
[150] S. B. Korada and N. Macris, "Tight bounds on the capacity of binary input random CDMA systems", IEEE Transactions on Information Theory, vol. 56, no. 11, pp. 5590-5613, Nov. 2010, ISSN: 1557-9654. DOI: 10.1109/ TIT.2010.2070131.
[151] J.-C. Mourrat, Hamilton-Jacobi equations for finite-rank matrix inference, 2019. arXiv: 1904.05294 [math. PR].
[152] G. Reeves, Information-theoretic limits for the matrix tensor product, 2020. arXiv: 2005.11273 [cs.IT]
[153] J. Barbier, N. Macris, M. Dia, and F. Krzakala, "Mutual information and optimality of approximate message-passing in random linear estimation", IEEE Transactions on Information Theory, vol. 66, no. 7, pp. 4270-4303, Jul. 2020, ISSN: 1557-9654. DOI: 10.1109/TIT.2020.2990880.
[154] G. Reeves, J. Xu, and I. Zadik, "All-or-nothing phenomena: from singleletter to high dimensions", in 2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Dec. 2019, pp. 654-658. DOI: 10.1109/CAMSAP45676.2019.9022473.
[155] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed., ser. Springer Series in Statistics. New York, NY, USA: Springer-Verlag New York, 2009, ISBN: 978-0-387-84858-7. DOI: 10.1007/ 978-0-387-84858-7.
[156] I. Rish and G. Grabarnik, Sparse Modeling: Theory, Algorithms, and Applications. USA: CRC Press, Inc., 2014, ISBN: 1439828695.
[157] J. A. Costa and A. O. Hero, "Learning intrinsic dimension and intrinsic entropy of high-dimensional datasets", in 2004 12th European Signal Processing Conference, Sep. 2004, pp. 369-372.
[158] M. Hein and J.-Y. Audibert, "Intrinsic dimensionality estimation of submanifolds in $\mathbb{R}^{d \prime}$, in Proceedings of the 22nd International Conference on Machine Learning, ser. ICML '05, Bonn, Germany: Association for Computing Machinery, 2005, pp. 289-296, ISBN: 1595931805. DOI: 10.1145 1102351.1102388 .
[159] E. Gardner and B. Derrida, "Three unfinished works on the optimal storage capacity of networks", Journal of Physics A: Mathematical and General, vol. 22, no. 12, pp. 1983-1994, Jun. 1989. DOI: 10.1088/0305-4470/22/12 004.
[160] G. Györgyi, "First-order transition to perfect generalization in a neural network with binary synapses", Physical Review A, vol. 41, pp. 7097-7100, 12 Jun. 1990. DOI: 10.1103/PhysRevA.41.7097.
[161] H. S. Seung, H. Sompolinsky, and N. Tishby, "Statistical mechanics of learning from examples", Physical Review A, vol. 45, pp. 6056-6091, 8 Apr. 1992. DOI: 10.1103/PhysRevA.45.6056.
[162] M. Opper and D. Haussler, "Generalization performance of Bayes optimal classification algorithm for learning a perceptron", Phys. Rev. Lett., vol. 66, pp. 2677-2680, 20 May 1991. DoI: 10.1103/PhysRevLett.66.2677.
[163] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors", IEEE Transactions on Information Theory, vol. 48, no. 11, pp. 2888-2910, Nov. 2002, ISSN: 1557-9654. DOI: 10.1109 TIT.2002.804053.
[164] A. M. Tulino, G. Caire, S. Verdú, and S. Shamai, "Support recovery with sparsely sampled free random matrices", IEEE Transactions on Information Theory, vol. 59, no. 7, pp. 4243-4271, Jul. 2013, ISSN: 1557-9654. DOI: 10.1109/TIT.2013.2250578.
[165] A. Tulino and S. Verdú, Random Matrix Theory and Wireless Communications. now, 2004, ISBN: 9781933019505. DOI: 10.1561/0100000001.
[166] L. Pastur and V. Vasilchuk, "On the moments of traces of matrices of classical groups", Communications in Mathematical Physics, vol. 252, no. 13, pp. 149-166, Oct. 2004. DOI: $10.1007 /$ s00220-004-1231-3.

# CLÉMENT LUNEAU 

Doctoral Researcher in Computer Science<br>표 French citizenship<br>Renens, Switzerland<br>Google Scholar

## EXPERIENCE

## Doctoral Researcher

EPF Lausanne, Communication Theory Laboratory<br>Sept. 2016 - Ongoing<br>- Lausanne, Switzerland

- Research on statistical inference problems involving highdimensional data (big data), establishing what performances are achievable both theoretically and algorithmically. Focusing on models of neural networks and tensor factorization.
- Published in top international conferences in machine learning (NeurIPS) and information theory (ISIT).


## Applied Scientist Intern

## Amazon

Jun. 2019 - Sep. 2019

- Edinburgh, United Kingdom

Worked on the design of the machine learning model used by Amazon for real-time bidding in display advertising.

## Research Assistant

## ETH Zürich, Signal and Information Processing Laboratory

Apr. 2016 - Jun. 2016
Zürich, Switzerland

- Continuation of the research started for my master's thesis under the direction of Prof. Hans-Andrea Loeliger.
- Proposed and studied a new algorithm to perform tomographic reconstruction. Advances presented at EUSIPCO'16.


## Electrical Engineer Intern

HFI (Haute Fréquence Ingénierie)
清 Jun. 2014 - Aug. 2014

- Meylan, France
- Wireless communication company designing products for professional mobile radio, lone worker monitoring and real-time indoor/outdoor localization.
- Worked on the software modernization of the wireless beacon used for localization with Contiki (OS for networked, memoryconstrained systems).


## TEACHING EXPERIENCE

Teaching assistant at EPFL for the following courses:

- Learning Theory - Spring 2019 \& 2020
- Markov Chains and Algorithmic Applications - Autumn 2017 \& 2018
- Information, Computation, Communication - Spring 2018 \& Autumn 2019
- Logic Systems - Spring 2017


## SKILLS

| Python | TensorFlow | Keras |
| :--- | :--- | :--- |
| Spark | SQL | Git | LaTeX

## LANGUAGES

## English

French
Spanish
Fluent, TOEFL IBT Score: 103/120 Native speaker
Basic (currently learning)

## EDUCATION

PhD in Computer Science

## EPF Lausanne

Sep. 2016 - Mar. 2021
Thesis title: Statistical limits of high-dimensional inference problems. Advised by Dr. Nicolas Macris.

## MSc in Electrical Engineering \& Information Technology <br> ETH Zürich

Sep. 2014 - Mar. 2016
Specialization in Communications, double degree program with Supélec

```
MEng (Diplôme d'Ingénieur)
Supélec
    Sep. 2012 - Jul. 2016 P Gif-sur-Yvette, France
```

Top-tier engineering school in Electrical Engineering, Computer Science and Telecommunications

## BSc. in Fundamental \& Applied Mathematics <br> Université Paris-Sud 11

Sep. 2012 - Jun. 2013 Orsay, France
Awarded with high honors, double degree program with Supélec

## Classes Préparatoires Scientifiques Lycée du Parc <br> Sept. 2010 - Jun. 2012 Lyon, France

2-year elite and intensive program in advanced Math and Physics to prepare the national competitive entrance exams to leading French engineering schools

- Invited talk on "Tensor estimation with structured priors" at the conference Youth in High-dimensions: Machine Learning, High-dimensional Statistics and Inference for the New Generation, invited by The International Centre for Theoretical Physics (ICTP), June 2020.
- Spotlight presentation on "Information theoretic limits of learning a sparse rule" at the 34th Annual Conference on Neural Information Processing Systems (NeurIPS 2020), December 2020.


## PUBLICATIONS

## Google Scholar Profile

## - ${ }^{-1}$ Journal Articles

- Luneau, C., J. Barbier, and N. Macris (2020). "Mutual information for low-rank even-order symmetric tensor estimation". In: Information and Inference: A Journal of the IMA.
- Luneau, C. and N. Macris (2020). "Tensor Estimation with Structured Priors". In: IEEE Journal on Selected Areas in Information Theory.
- Gabrié, M., A. Manoel, C. Luneau, J. Barbier, N. Macris, F. Krzakala, and L. Zdeborová (2019). "Entropy and mutual information in models of deep neural networks". In: Journal of Statistical Mechanics: Theory and Experiment 2019.12, p. 124014.


## :O: Conference Proceedings

- Luneau, C., J. Barbier, and N. Macris (2020). "Information theoretic limits of learning a sparse rule". In: 34th Annual Conference on Neural Information Processing Systems (NeurIPS 2020).
- Luneau, C., N. Macris, and J. Barbier (2020). "High-dimensional rank-one nonsymmetric matrix decomposition: the spherical case". In: IEEE International Symposium on Information Theory (ISIT 2020), pp. 2646-2651.
- Barbier, J., C. Luneau, and N. Macris (2019). "Mutual Information for Low-Rank Even-Order Symmetric Tensor Factorization". In: IEEE Information Theory Workshop (ITW 2019), pp. 1-5.
- Gabrié, M., A. Manoel, C. Luneau, J. Barbier, N. Macris, F. Krzakala, and L. Zdeborová (2018). "Entropy and mutual information in models of deep neural networks". In: 32nd Annual Conference on Neural Information Processing Systems (NeurIPS 2018), pp. 1821-1831.
- Zalmai, N., C. Luneau, C. Stritt, and H. Loeliger (2016). "Tomographic reconstruction using a new voxel-domain prior and Gaussian message passing". In: 24th European Signal Processing Conference (EUSIPCO 2016), pp. 2295-2299.


[^0]:    ${ }^{1}$ Note that, for a matrix (order- 2 tensor) $\mathbf{M}$, this definition of rank is equivalent to the usual definition where the rank is the dimension of the vector space spanned by the columns of $\mathbf{M}$.

[^1]:    ${ }^{2}$ It means that $\lim _{n \rightarrow+\infty} \mathbb{P}(\mathbf{X}$ is not full column rank $)=0$. For the distribution of the rows that we usually consider, the latter also holds for a finite $n \geq K$. For example, if the rows are i.i.d. Gaussian vectors then $\mathbb{P}(\mathbf{X}$ is not full column rank $)=0$ as long as $n \geq K$.

[^2]:    ${ }^{3}$ The model does not prevent us to work with deterministic activation function as we can choose a deterministic distribution $P_{A}$, or an activation function $\varphi: \mathbb{R} \times \mathbb{R}^{k_{A}} \rightarrow \mathbb{R}$ that is constant with respect to its second argument.

[^3]:    ${ }^{4}$ Here our words are intentionally vague. By bearing a relation, we mean that $\mathbf{X}$ and $\mathbf{Y}$ are not statistically independent. If it is the case then we can simply discard $\mathbf{Y}$, e.g., if $\mathbf{X}$ are real variables then any estimator of $\mathbf{X}$ based on $\mathbf{Y}$ is no better than $\mathbb{E} \mathbf{X}$.
    ${ }^{5}$ Note the use of $:=$ stressing that the right-hand side defines the left-hand side. Bayes' rule is an identity only if the beliefs match the true distributions.

[^4]:    ${ }^{6}$ For the sake of giving another loss function, let us mention the $0-1$ loss $\frac{1}{n} \sum_{i=1}^{n} \delta_{x_{i}, \widehat{x}_{i}}$ ( $\delta$ is the Kronecker delta) that is more meaningful for categorical variables and discrete real variables.

[^5]:    ${ }^{7}$ As long as the corresponding weight matrix is not diagonal, see the next sentence.

[^6]:    ${ }^{8}$ This canonical Boltzmann distribution is usually derived from the micro-canonical distribution which postulates that all microscopic configurations of an isolated system occur with equal probability. Here the spin system is in equilibrium with a thermal bath of constant temperature $T$, hence the isolated system is the union of the spin system and the thermal bath.

[^7]:    ${ }^{9}$ The Curie-Weiss model follows from the Ising model by approximating the local magnetization density $\frac{1}{2 d} \sum_{j: j \sim i} x_{j}$ at the site of each $x_{i}$ by the global magnetization density $\frac{1}{n} \sum_{j=1}^{n} x_{j}$.

[^8]:    ${ }^{10}$ For the upper bound, note that $1=(1-x+x)^{n}=\sum_{i=1}^{n}\binom{n}{i} x^{i}(1-x)^{n-i} \geq\binom{ n}{k} x^{k}(1-x)^{n-k}$ for $x \in[0,1]$ and let $x:=k / n$ to get $1 \geq\binom{ n}{k} \exp \left(-n h_{b}(k / n)\right)$. Thanks to Emre Telatar for pointing out this simpler argument.
    ${ }^{11}$ Spin glasses owe their name to their low-temperature behavior. The spins of a ferromagnet (think of Example 1.4 with $J, h>0$ ) all align in the same direction at low temperature $(\beta \rightarrow+\infty)$, yielding an average magnetization per spin that is nonzero. Instead, in a spin glass, the spins freeze at low temperature in states that result in the absence of an average magnetization. Like a window glass lacks the spatial order of a crystal, a spin glass lacks the spin order of a ferromagnet.

[^9]:    ${ }^{12}$ It assumes that the states of the replicated system that are significant in the partition function (1.34) possess some symmetry under permutation of the replica indices, e.g., that the overlap $\sum_{i} x_{i}^{(r)} x_{i}^{\left(r^{\prime}\right)} / n$ between two replicas does not depend on the pair $r \neq r^{\prime}$, or that the magnetization $\sum_{i} x_{i}^{(r)} / n$ does not depend on $r$.

[^10]:    ${ }^{13}$ "The mathematics were optimistic to say the least" 54 , S.F. Edwards about the replica method that he introduced in 41 .
    ${ }^{14}$ The entropy $S:=-\mathbb{E}\left[\sum_{\mathbf{x}} P_{\beta}(\mathbf{x} ; \mathbf{J}) \ln P_{\beta}(\mathbf{x} ; \mathbf{J})\right]$ where $P_{\beta}(\mathbf{x} ; \mathbf{J}) \propto e^{-\beta \mathcal{H}(\mathbf{x} ; \mathbf{J})}$ is by definition nonnegative but the replica symmetric ansatz implies a negative entropy at low temperature.

[^11]:    ${ }^{15}$ This is an application of the envelope theorem 68, Corollary 4].
    ${ }^{16}$ To distinguish from algorithmically possible.

[^12]:    ${ }^{17}$ Think about Example 1.2 where a spin $X_{i}$ interacts with all the others through the observations $\left(Y_{i j}\right)_{j \neq i}$, or Model 1.3 where all the signal entries interact at once through each of the observations $Y_{i}$.

[^13]:    ${ }^{18}$ The MMSE is different from the MSE of a random guess based on the prior of $\mathbf{X}$.

[^14]:    ${ }^{19}$ The black-and-white scan of a digit is a large matrix where each entry is the intensity of a pixel. Given that everyone writes digits more or less the same, the scanned digit effectively lives in a small subspace of all the possible images.

[^15]:    ${ }^{1} \mathcal{Z}_{t, R}\left(\mathbf{Y}^{(t)}, \widetilde{\mathbf{Y}}^{\left(t, R_{v}\right)}, \overline{\mathbf{Y}}^{\left(t, R_{u}\right)}\right)$ is the normalization factor of the right-hand side of (3.15).

[^16]:    ${ }^{1}$ Our theorems are proven here for bounded and smooth activation functions but, as explained, the proofs can be extended to unbounded and piecewise differentiable ones. Numerical solutions involve non-trivial integrals that are much easier to handle for piecewise linear functions.

[^17]:    ${ }^{1}$ Each weight matrix $\mathbf{W}^{(\ell)}$ is initialized with i.i.d. Gaussian entries and its singular-value decomposition $\mathbf{W}^{(\ell)}=\mathbf{U}_{\ell} \boldsymbol{\Sigma}_{\ell} \mathbf{V}_{\ell}^{\top}$ is computed. Then, $\mathbf{U}_{\ell}$ is an orthogonal matrix drawn from the Haar measure on $O(n)$ (the set of $n \times n$ orthogonal matrices), $\mathbf{V}_{\ell}$ is an orthogonal matrix drawn from the Haar measure on $O(m)$, and $\boldsymbol{\Sigma}_{\ell} \in \mathbb{R}^{m \times n}$ is a diagonal matrix with nonnegative entries. The matrices $\mathbf{U}_{\ell}, \mathbf{V}_{\ell}$ are kept fixed during training and only the singular values $\boldsymbol{\Sigma}_{\ell}$ are updated by SGD.

[^18]:    ${ }^{2}$ The third hypothesis is part of the definition of the model but we repeat it here to insist on the distribution that the weights have to follow.

[^19]:    ${ }^{1}$ The derivative of $I\left(\mathbf{X}^{*} ; \mathbf{Y}, \widetilde{\mathbf{Y}}^{(\tau)} \mid \mathbf{W}\right) / m_{n}$ with respect to $\tau$ at $\tau=0$ is equal to half the MMSE of the original problem.

[^20]:    ${ }^{2}$ Note that there is no loss of generality here. We can always rescale $\varphi$ and $P_{0}$ jointly to satisfy $\mathbb{E} X_{0}^{2}=1$.

[^21]:    ${ }^{3}$ Remember that $r_{\epsilon}$ takes its values in $\left[0, \frac{\alpha_{n}}{\rho_{n}} r_{\text {max }}\right]$. Besides, under (H2), $\ell_{Y_{\mu}^{(t, e)}}^{\prime}$ is upper bounded by $\left(\left|Y_{\mu}^{(t, \epsilon)}\right|+\|\varphi\|_{\infty}\right) \Delta^{-1}\left\|\partial_{x} \varphi\right\|_{\infty}=\left(\sqrt{\Delta}\left|Z_{\mu}\right|+2\|\varphi\|_{\infty}\right) \Delta^{-1}\left\|\partial_{x} \varphi\right\|_{\infty}$ (see the inequality 7.108) in Appendix 7.D). The noise $Z_{\mu}$ is averaged over thanks to the expectation.

