Dynamical Low Rank approximation of PDEs with random parameters

PAR

Eleonora MUSHARBASH

(EPFL

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2017

Abstract

In this work, we focus on the development and analysis of numerical techniques for the propagation of uncertainty through a large-scale dynamical system with random parameters. In this context, the numerical simulation of the random dynamics typically requires a large computational cost at each time, leading to a total effort which is often computationally unaffordable. Model order reduction techniques offer a remedy to overcome this difficulty, by deriving models of lower dimension, thus solvable at a relatively low computational cost, which accurately replicate the relation between input random parameter and solution dynamic. This idea relies on the observation that in many cases the collection of all solutions at all times, corresponding to all possible outcomes of the input random processes, can be well approximated in a low dimensional (low-rank) subspace. The main practical difficulty is that such subspace is, in general, not easy to characterize a priori and might significantly change during the evolution of the system. To overcome this problem we investigate a Dynamical Low Rank (DLR) approach, in which the approximation subspace is not fixed a priori and evolves in time by following the trajectory of the solution. The DLR can be interpreted as a reduced basis method, where the approximate solution is expanded in separable form over a set of few deterministic basis functions (modes) at each time, with the peculiarity that both the deterministic modes and the stochastic coefficients are computed on the fly and are free to adapt in time so as best describe the structure of the random solution.

Our first goal is to generalize and reformulate in a variational setting the Dynamically Orthogonal (DO) method, proposed by Sapsis and Lermusiaux (2009-2012) for the approximation of fluid dynamic problems with random initial conditions. The DO method is reinterpreted as a Galerkin projection of the governing equations onto the tangent space along the approximate trajectory to the manifold $\mathcal{M}_S$, given by the collection of all functions which can be expressed as a sum of S linearly independent deterministic modes combined with S linearly independent stochastic modes. Depending on the parametrization of the tangent space, one obtains a set of nonlinear differential equations, suitable for numerical integration, for both the coefficients and the basis functions of the approximate solution. By formalizing the DLR variational principle for parabolic PDEs with random parameters we establish a precise link with similar techniques developed in quite different contexts such as the Multi-Configuration Time-Dependent Hartree method (MCTDH) in quantum dynamics and the Dynamical Low-Rank approximation in a finite dimensional setting. On the other hand, the DLR approach gives a unified formulation for the DO method and other dynamical low-rank techniques

such as the DyBO and the DDO method, recently proposed in the UQ context. By the use of curvature estimates for the approximation manifold $\mathcal{M}_S$, we derive a theoretical bound for the approximation error of the $S-$terms DO solution by the corresponding S-terms best approximation, i.e. the truncated $S-$terms Karhunen-Loève expansion at each time instant. The bound is applicable for full rank DLR approximate solutions on the largest time interval in which the best $S-$terms approximation is continuously differentiable in time.

Secondly, we focus on parabolic equations, especially incompressible Navier Stokes equations, with random Dirichlet boundary conditions and we propose a DLR technique which allows for the strong imposition of such boundary conditions. We show that the DLR variational principle can be set in the constrained manifold of all S rank random fields with a prescribed value on the boundary, expressed in low-rank format, with rank M smaller than S. We characterize the tangent space to the constrained manifold by means of the Dual Dynamically Orthogonal (Dual DO) formulation, in which the stochastic modes are kept orthonormal and the deterministic modes satisfy suitable boundary conditions, consistent with the original problem. The same formulation is also used to conveniently include the incompressibility constraint when dealing with incompressible Navier Stokes equations with random parameters. Hence the latter is reduced to a set of S coupled PDEs for the evolution of the deterministic modes (M of which with non-homogeneous boundary conditions) coupled with $S - M$ ODEs for the evolution of the stochastic modes. The Dual DO method has been tested on two fluid dynamics problems: the classical benchmark of a laminar flow around a cylinder with random inflow velocity, and a biomedical application for simulating blood flow in a realistic carotid artery reconstructed from MRI data, where the inflow boundary conditions are taken as random due to the uncertainty and large errors in Doppler measurements.

Finally, we extend the DLR approach for the approximation of wave equations with random parameters. We propose the Symplectic DO method, according to which the governing equation is rewritten in Hamiltonian form and the approximate solution is sought in the low dimensional manifold of all complex-valued random fields with fixed rank. Recast in the real setting, the approximate solution is expanded over a set of few dynamical symplectic deterministic modes and satisfies the symplectic projection of the (real) Hamiltonian system into the tangent space of the approximation manifold along the approximate trajectory. As a result, the approximate solution preserves the mean Hamiltonian energy and continuously adapts in time to the structure of the solution.

Key words: Dynamical Low Rank, Dynamically Orthogonal approximation, Reduced Basis method, Uncertainty Quantification, Navier Stokes equations, wave equations

Résumé

Dans ce travail, nous nous concentrons sur le développement et l'analyse de techniques numériques concernant la propagation de l'incertitude par un système dynamique à grande échelle comportant des paramètres aléatoires. Dans ce contexte, chacune des simulations numériques des dynamiques aléatoires exige généralement un coût computationnel élevé, générant un effort total considérable. Les modèles d'ordre réduit offrent une solution pour surmonter cette difficulté, en proposant des modèles de plus petites dimensions, résolubles à un coût computationnel relativement faible et qui répliquent avec précision la relation entre les paramètres aléatoires et la dynamique de la solution. Cette idée repose sur l'observation que, dans de nombreux cas, la collecte de toutes les solutions en tout temps, correspondant à tous les résultats possibles des processus aléatoires en entrée, peut être approximée efficacement dans un sous-espace de faible dimension (bas rang). La difficulté pratique provient de la difficulté à caractériser a priori ce type de sous-espaces, qui peuvent de plus changer considérablement avec l'évolution du système. Pour surmonter ce problème, nous étudions une approche Dynamical Low Rank (DLR), dans laquelle le sous-espace d'approximation n'est pas fixé a priori et évolue dans le temps en suivant la trajectoire de la solution. Le DLR peut être interprété comme une méthode de base réduite, où la solution approximée est développée de manière séparable sur un ensemble de quelques fonctions (modes) de base déterministes à chaque instant, avec la particularité que les modes déterministes et les coefficients stochastiques sont calculés à la volée et sont libres de s'adapter dans le temps afin de mieux décrire la structure de la solution aléatoire.

Notre premier objectif est de généraliser et de reformuler dans un contexte variationnel la méthode dite Dynamique Orthogonale (DO), proposée par Sapsis et Lermusiaux (2009-2012) pour l'approximation de problèmes liés à la dynamique des fluides avec des conditions initiales aléatoires. La méthode DO est réinterprétée comme une projection de Galerkin des équations gouvernantes du problème sur l'espace tangent le long de la trajectoire approximée à la surface $\mathcal{M}_S$, donnée par la collection de toutes les fonctions, qui peut être exprimée par la somme de S modes déterministes linéairement indépendants combinée à S modes stochastiques linéairement indépendants. En fonction du paramétrage de l'espace tangent, on obtient un ensemble d'équations différentielles non linéaires, adaptées à l'intégration numérique, tant pour les coefficients que pour les fonctions de base de la solution approximée. En formalisant le principe variationnel DLR pour les EDP paraboliques avec des paramètres aléatoires, nous établissons un lien précis avec des techniques similaires développées dans des contextes

différents, tels que la méthode Multi-Configuration Time Dependent Hartree (MCTDH) en dynamique quantique et l'approximation DLR en dimension finie. D'autre part, l'approche DLR donne une formulation unifiée pour la méthode DO et d'autres techniques DLR telles que la DyBO et la méthode DDO, récemment proposées dans le contexte UQ. En utilisant les estimations de courbure pour la surface approximée $\mathcal{M}_S$, nous dérivons une borne théorique pour l'erreur d'approximation de la solution de DO à S-termes par la meilleure approximation S-termes, c'est-à- dire l'expansion tronquée S-Karhunen-Loève à chaque instant. La borne est applicable pour les solutions approximées DLR de rang plein au niveau du plus grand intervalle de temps dans lequel la meilleure approximation S-termes est différentiable en continu dans le temps. Deuxièmement, nous nous concentrons sur les équations paraboliques, en particulier les équations Navier- Stokes incompressibles, avec des conditions aux limites de Dirichlet aléatoires et nous proposons une technique DLR qui permet d'imposer ces conditions aux limites. Nous montrons que le principe variationnel DLR peut être défini dans une surface contrainte de tous les champs aléatoires de rang S avec une valeur prescrite sur la frontière, exprimée en format de rang faible, avec le rang M plus petit que S. Nous caractérisons l'espace tangent de la surface contrainte au moyen de la formulation Dual Dynamic Orthogonal (Dual DO), dans laquelle les modes stochastiques sont maintenus orthonormés et les modes déterministes satisfont des conditions de limites appropriées, compatibles avec le problème d'origine. La même formulation est également utilisée pour inclure commodément la contrainte d'incompressibilité, lorsqu'il s'agit d'équations Navier-Stokes incompressibles avec des paramètres aléatoires. Par conséquent, ce dernier est réduit à un ensemble S de PDE couplées pour l'évolution des modes éterministes (dont M avec des conditions aux limites non homogènes) associé aux S-M ODEs pour l'évolution des modes stochastiques. La méthode Dual DO a été testée sur deux problèmes de dynamique des fluides : la référence classique d'un flux laminaire autour d'un cylindre avec une vitesse d'entrée aléatoire et une application biomédicale, pour simuler le flux sanguin dans une artère carotide réaliste reconstituée à partir de données IRM, où les conditions limites d'entrée sont définies comme aléatoires en raison de l'incertitude et des erreurs importantes liées aux mesures Doppler. Enfin, nous étendons l'approche DLR pour l'approximation des équations des ondes avec des paramètres aléatoires. Nous proposons la méthode DO Symplectique, selon laquelle l'équation gouvernant du problème est réécrite sous forme hamiltonienne et la solution approximée est recherchée dans la surface de faible dimension de tous les champs aléatoires à valeur complexe de rang fixe. Dans une configuration réelle, la solution approximée est développée sur un ensemble de quelques modes déterministes symplectifs dynamiques et satisfait la projection symplectique du système hamiltonien (réel) dans l'espace tangent de la surface approximée le long de la trajectoire approximée. En conséquence, la solution approximée préserve l'énergie hamiltonienne moyenne et s'adapte en permanence à la structure de la solution.

Mots clefs : Dynamical Low Rank, Dynamically Orthogonal approximation, méthode de base réduite, quantification de l'incertitude, équations Navier-Stokes, équations des ondes

Acknowledgements

Lausanne, 18 May 2017   E. M.

Contents

Contents

List of Figures

Introduction

0.1 Forward UQ Problem: motivations and scopes

The last decades have witnessed a growing demand for mathematical modelling and numerical simulations in engineering applications across science and technology. Typically, the problem of interest is described by differential equations, discretized and approximated by numerical techniques and finally reproduced by computer simulations.

Reliable mathematical models need to take into account the presence of variability and/or lack of precise characterization of the input data. As an example, one may think for instance to biomedical applications simulating blow flow by means of computational fluid dynamics (e.g. Navier Stokes equations). In this context, numerical simulations can be used as a virtual platform to support the diagnosis of many cardiovascular diseases by predicting/quantifying hemodynamics indicators such as cardiac functional parameters affecting the cardiovascular anomalies. The input data of the corresponding mathematical model include physical parameters, as for instance flow viscosity, but also initial and boundary conditions, obtained in terms of flow rates and stresses after Doppler ultrasound tests, and computational domains, typically reconstructed from magnetic resonance images (MRI), acquired in vivo. The exact quantification of most of those parameters is typically compromised by measurement errors, reduced amount of data, for invasive data collections, intrinsic variability of the parameter itself and oversimplification of the model.

Uncertainty quantification offers a possible solution to overcome the limitations of deterministic mathematical models. The approach that we consider in this thesis consists in recasting the underlying (deterministic) problem in a probabilistic framework: the sources of uncertainty are included into the model with the goal of quantifying the variability of predicted output quantities, and perform robust and reliable simulations. In this context one can distinguish two main areas:

- the forward uncertainty propagation which aims to assess the impact of uncertain inputs into the model outputs,

- inverse problem in which the aim is to reduce the uncertainty in some input, by comparing the outputs of the model with experimental measurements.

In this thesis, we analyze only the first problem. We consider dynamical systems which are described by time-dependent partial differential equations (PDEs), including parabolic diffusion equations, Navier Stokes equations and second order wave equations, in which the uncertainty may affect the coefficients of the differential operator (e.g. diffusion, advection, reaction terms in elliptic PDEs) as well as initial and boundary conditions and forcing terms. Understanding how uncertainty propagates through a dynamical system is an important engineering problem, and it is currently the focus of many significant research efforts. Cases of interest come from diffusion and transport phenomena in highly heterogeneous porous media; propagation of seismic waves from unknown/uncertain sources and/or in randomly heterogeneous media; internal flows such as blood flow in an artery with uncertain inflow boundary conditions, just to name a few.

We always assume that the uncertainty in the parameters is known and properly described in terms of its probability distribution. In other words, the parameters of the model which are subject to uncertainty, are modelled as random variables or random fields with well know probabilistic law. The aim is to quantify the induced variability on outputs quantities, at each time. The numerical challenges in this context are several, especially when the distribution of the input parameters is complex or high dimensional, and when the parameters-to-solution map is non linear. In these cases, efficient numerical techniques are needed to guarantee accurate approximations with affordable computational cost.

The Monte Carlo (MC) method [46, 21] is the most straightforward and popular technique to solve forward UQ problems. In this method, the statistics of the solution are approximated by generating a sample of M independent realizations of the random data, solving the corresponding M deterministic PDEs, and averaging over all the realizations. The drawback of this method is the slow rate of convergence, meaning that large sample sizes M are needed for accurate approximations of output statistics. For practical problems, this often leads to a prohibitive computational effort.

In the last decades, many efforts have been devoted to build, upon the classic Monte Carlo method, improved versions with better convergence rate. We name here the Quasi Monte Carlo [51, 98] and the Multilevel Monte Carlo versions [37].

Contrarily to sampling methods of MC type, spectral methods aim at reconstructing the functional dependence of the solution on the input parameters. This strategy, which relays on the assumption that the parameters-to-solution map is smooth, consists in expanding the random solution over a suitable basis, e.g. of orthogonal polynomials with respect to the probability density function of the input parameters. Hence, practical approximations can be computed by Galerkin projection [6, 49, 89, 11] or collocation on tensor or sparse grids of Gauss points [5, 135, 75], with a computational cost which is drastically lower than that of sampling techniques, provided that the assumption of regularity is satisfied. On the other hand, this procedure becomes numerically challenging when the stochastic space has high (infinite) dimension or the parameter-to-solution map features low regularity. Indeed, despite

2

the great improvements obtained with the use of sparse techniques, spectral methods are often exposed to the so called curse of dimensionality, i.e. an error versus cost performance that is negatively affected by the high dimension of the underlying stochastic space and degenerates in the limit of infinite dimension. Moreover, independently from the gPC procedure one chooses, the propagation of the uncertainty through dynamical systems remains a challenging task, especially for long time integrations, essentially due to the fact that the probability distribution of the solution evolves as a function of time and might significantly deviate from the probability distribution of the input data. Roughly speaking, at early times the solution typically "remembers" the input parameters and its density functions stays close to that of the random data, while for later times might start to move away from the distribution of the input and develop its own stochastic characteristics. This implies that an approximation with fixed (in time) polynomial basis might become more and more demanding, in the sense that more and more terms have to be included during the evolution to properly approximate the solution. Such difficulty arises even when dealing with simple linear systems, a classical example are acoustic or elastic waves with uncertain random speed, as exhaustively discussed in [128, 48].

A different approach, recently proposed for the quantification of the uncertainty, is instead based on model order reduction techniques. In this context, the approximation strategy relays on the observation that, for certain classes of problems, the solution manifold $\mathscr{U}$, i.e. the collection of all solutions at all times and for all possible values of the input parameters, can be well approximated by a linear subspace $\mathscr{U}_S$ of small dimension S. Roughly speaking, this means that the solution of the original PDE problem, which belongs to an infinite dimensional space, can be well approximated by a linear combination of S terms, with S small. Assuming to be able to parametrize the approximation manifold $\mathscr{U}_S$ in terms of S orthogonal functions, called reduced basis, then the approximation problem reduces to solve a S dimensional system, obtained by the Galerkin projection of the governing equation onto $\mathscr{U}_S$. In practice, this means that the potentially high dimensional (forward) UQ problem is reduced in the on-line stage to low-cost reduced-order simulations, which, as a consequence, are not (deeply) affected by the curse of dimensionality. The main practical difficulty in this approach is that the approximation subspace $\mathscr{U}_S$ is, in general, not easy to characterize a priori. A technique widely used in the applications is the Proper Orthogonal Decomposition (POD) [75, 20, 132, 23], according to which the reduced basis is extracted by a singular value decomposition of the correlation matrix of the collected snapshots, i.e. a certain number of precomputed solutions corresponding to several values of the input parameters. When dealing with time-dependent problems, the snapshots typically need to be collected at several time instants and the constructed reduced basis is meant to work for all times. This negatively affects both the off-line stage, i.e. the precomputation of the snapshots and the SVD decomposition, which here requires a considerable computational cost, and the on-line stage, since the reduced basis typically needs to be quite rich to effectively approximate the solution at all times. Greedy algorithms to construct the reduced basis by optimizing the computational cost in the off-line stage, for both time-dependent and steady-state equations,

have been recently proposed in literature and successfully applied to large classes of problems [52, 97, 54, 43, 26, 18, 10]. However, the approximation of random dynamics, for which the solution might significantly change over time, remains a challenging task. In this situations, despite the use of greedy techniques, the number of reduced basis still needs to be sufficiently large to be able to approximate the solution manifold at each time. This is in some sense the same issue affecting the gPC approach (long time integration), if one reinterprets the set of multivariate polynomial functions which are used in the expansion of the solution, as a stochastic (reduced) basis with deterministic coefficients, symmetrically to the RB/POD methods which use deterministic reduced bases with stochastic coefficients. More generally this means that long time-integration negatively influences any low-rank methods which make use of (a fixed number of) basis functions which are constant in time. Applications which need to face such numerical difficulty can be found for instance in seismic engineering, to describe the wave propagation with random source location, or in atmospheric flows and weather forecasting, using fluid dynamic equations with random initial conditions.

The most direct attempt to overcome the limitations related to expansions of the solution on a fixed basis, either deterministic or stochastic, consists in developing dynamical low-rank techniques, by using time evolving basis. Roughly speaking the corresponding methods aim at evolving the low-rank approximate solution in time, hence adapting the spatial and stochastic basis so as to best describe the structure of the solution at each time instant, in order to preserve both the low-rank format and the approximation accuracy. Such approach, initially proposed in finite dimensional setting by Lubich [66, 56] with the name of Dynamical Low Rank (DLR) approximation, and widely used in quantum mechanics for the approximation of deterministic time-dependent Schrödinger equations, has been independently introduced in UQ context by Sapsis and Lermusiaux [116, 117] (Dynamically Orthogonal (DO) method) and successfully applied to problems in ocean dynamics with random data. The DLR can be essentially seen as a reduced basis method, thus solvable at a relatively low computational cost, in which the solution is expanded as a linear combination of few deterministic modes with random coefficients. Its peculiarity is that both the spatial and the stochastic bases are computed on the fly and are free to evolve, thus adjusting at each time to the current structure of the random solution. The computational saving of the DLR method is however contrasted by more complex implementation requirements. Contrary to sampling or collocation methods, which take the underlying deterministic model as a black box and can directly use available solvers, the implementation of DLR techniques depends on the structure of the governing equation and the computational cost is made heavier by the continuous update in time of the reduced bases. On the other hand, this allows obtaining a suitable approximation of the whole set of solutions corresponding to all possible realizations of the input random parameters, at each time.

From a variational point of view the approximate solution is sought in the low dimensional manifold of all S rank random fields (functions which can be expressed as sums of S linearly independent deterministic modes combined with S linearly independent stochastic modes);

4

it is obtained by performing a Galerkin projection of the governing equations onto the (time-dependent) tangent space to the manifold along the solution trajectory. After an explicit parametrization of the manifold, one obtains a set of nonlinear differential equations, suitable for numerical integration, for both the coefficients and the basis functions of the approximate solution.

In this work, we thoroughly investigate Dynamically Low-Rank techniques for the propagation of uncertainty through a large-scale dynamical system with random parameters. Specific types of equations that will be considered include parabolic diffusion equations, incompressible Navier Stokes equations, and second order wave equations.

1 Thesis Overview

1.1 Problem setting

In this thesis, we focus on time-dependent partial differential equations (PDEs) with random input parameters as e.g. initial or boundary conditions, model coefficients and forcing terms. Specific types of equations that will be considered include parabolic diffusion equations, incompressible Navier Stokes equations and second order wave equations.

We introduce here the problem setting that will be addressed in Chapter 3 and Chapter 4. The discussion about the problem analyzed in Chapter 5 is postponed to Section 1.4.3.

Let D be an open and bounded physical domain in R^d, $1 \le d \le 3$, with Lipschitz continuous boundary ∂D, and let $(\Omega, \mathscr{A}, \mathscr{P})$ be a complete probability space, where Ω is the set of outcomes, $\mathscr{A}$ a σ-algebra and $P : \mathscr{A} \to [0,1]$ a probability measure. We consider the following general real valued problem:

$$
\begin{cases}
\dot{u}(x,t,\omega) = \mathscr{L}(u(x,t,\omega),t,\omega), & x \in D, \ t \in (0,T], \ \omega \in \Omega, \\
u(x,0,\omega) = u_0(x,\omega) & x \in D, \ \omega \in \Omega, \\
\mathscr{B}(u(x,t,\omega),\omega) = g(x,t) & x \in D, \ t \in (0,T], \ \omega \in \Omega,
\end{cases}
\tag{1.1}
$$

where t and x are respectively the time variable in $[0,T]$ and the spatial variable in D, $\mathscr{L}$ is a linear or non-linear differential operator, and $\mathscr{B}$ an operator defining the boundary conditions. Here $\omega \in \Omega$ represents a random elementary event which may affect the operator $\mathscr{L}$ (as e.g. a coefficient of a forcing term), the boundary conditions or the initial conditions. (The case of a random domain can be recast to form (1.1) after introducing a mapping onto a fixed reference configuration, but this will not be discussed in the thesis.)

We denote by $L^2(D)$ the Hilbert space of square integrable real valued functions defined in D:

$$
L^2(D) = \left\{ u : D \to \mathbb{R} \text{ s.t. } \int_D |u(x)|^2 \, dx < \infty \right\}
$$

Without further specifications, hereafter we assume that the solution $u(\cdot, t, \omega)$ to problem (1.1) is in a certain real Hilbert space $\mathcal{H} \subset L^2(D)$ for (almost) all $t \in [0, T]$ and $\omega \in \Omega$ and that $\mathcal{L}(u, t, \omega) \in \mathcal{H}'$ for all $u \in \mathcal{H}$ and almost everywhere in $[0, T]$ and Ω, with $\mathcal{H}'$ denoting the dual of $\mathcal{H}$.

1.2 Reduced Order Models: state of the art

The quantification of uncertainty for large scale problems is often a computationally challenging task, complicated by the typically high (infinite) dimension of the stochastic space and made even more prohibitive when dealing with time-dependent problems. In this Section, we describe approaches proposed in the literature, which accelerate the computation of the approximate solution by deriving low complexity (reduced) models of the governing equation. These are solvable at relatively low computational cost and accurately preserve the relation between input parameters and outputs. Typically the reduction is achieved via state projection onto a low dimensional spectral basis. In this spirit, we include here the generalized Polynomial Chaos approach, in which the solution is approximated by a truncated expansion over a set of stochastic global multivariate polynomial functions, as a stochastic reduced basis with deterministic coefficients, symmetrical to the "classical" reduced order methods which use deterministic reduced bases with stochastic coefficients.

1.2.1 Reduced Basis models

The Reduced Basis (RB) is a class of reduced order techniques that has been initially introduced to approximate parametrized problems and used to speed up the computational cost in real-time simulations and many-query contexts (e.g. optimization, control or parameter identification), see e.g. [58, 110] and references therein. More recently, RB models have also been applied to UQ problems for the computation of statistics of random solutions [44, 28, 29, 17, 27].

The central idea of the RB approach is to approximate the potentially infinite dimensional problem onto a linear subspace spanned by few, well chosen, deterministic basis functions. When applied to problem (1.1), the approximate solution is sought in the form:

$$u_S^{RB}(x, t, \omega) = \sum_{i=1}^{S} U_i(x) Y_i(t, \omega) \tag{1.2}$$

and satisfies for each $\omega \in \Omega$, the governing equation (1.1) projected on the subspace $\mathcal{U}_S^{RB}$, spanned by $U_1, ..., U_S$, at each time, i.e.

$$\langle \dot{u}_S^{RB}(x, t, \omega) - \mathcal{L}(u_S^{RB}, t, \omega), v \rangle = 0, \quad \forall v \in \mathcal{U}_S^{RB}, \ \forall (t, \omega) \in (0, T] \times \Omega \tag{1.3}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality paring between $\mathcal{H}$ and $\mathcal{H}'$, completed with proper imposition

of the initial and boundary condition.

Computationally speaking, the key idea of this approach is to split the computations into two parts: off-line stage and an on-line stage. The former is computationally the most expensive and consists in the selection of the deterministic modes $U_1, ..., U_S$. Depending on the strategy used to select the reduced basis, we distinguish the proper orthogonal decomposition (POD) [75, 20, 132, 23] and the greedy reduced basis (greedy-RB) methods [52, 97, 54, 43, 26, 18, 10]. The first is based in a "brute-force" sampling of the random parameters, which are used to generate the set of snapshots, i.e approximate solutions to the corresponding problems stored at several time instants (typically computed by some very accurate Finite Element discretization). The reduced basis is then extracted by performing a truncated Singular Value Decomposition (SVD) of the matrix collecting all snapshots. In the greedy-RB the sampling is instead calibrated by greedy and goal oriented algorithms, usually based on some specific a posteriori error estimators of residual type. Typically, the random parameter space is explored in an iterative way and at each iteration, a new solution is computed for the values of the input parameters which maximize, in some proper norm, the predicted error. Such solution is then added into the set of collected snapshots, and the procedure is repeated until satisfying a prescribed error tolerance. For time-dependent problems, a POD-greedy technique, which at each greedy step is invoking a POD compression in time, is usually preferred to a pure greedy approach in both the time and the parameter space, as the latter may generally "stall" before arriving at convergence [53]. The on-line stage consists in the assembly of the reduced order system (1.3), which is then solved with a computational cost that is independent of the dimension of the algebraic full order system, obtained by the discretization of the governing equation. In a UQ framework, the computation of statistics of the solution, and hence of the coefficients $Y_1, ..., Y_S$, still requires many solutions of the system (1.3) (generally obtained by Monte Carlo or Stochastic Collocation method) which is however much cheaper than the original problem (1.1).

Even if very effective in may situations, RB approximations may still require a high computational cost for certain classes of potentially compressible problems, for which the exact solution is "nearly low rank" at any fixed time. In these cases, the computational efficiency of the reduced model is compromised by the fact that the solution manifold at time t, i.e. the collection $\mathscr{U}(t) = \{u(t, \omega), \omega \in \Omega\}$ of all solutions at time t for all parameters $\omega \in \Omega$, although being nearly contained in a low dimensional subspace, may significantly change over time. This implies that the (fixed in time) reduced basis $(U_1, ..., U_S)$ has to be sufficiently rich to be able to approximate $\mathscr{U}(t)$ at all time, leading to a fairly large reduced model.

1.2.2 Generalized Polynomial Chaos expansion

An alternative approach that has been proposed in literature consists in first parametrizing the randomness in terms of a finite dimensional random vector $\mathbf{y}: \Omega \to \mathbb{R}^N$, possibly through some truncation step, and then approximating the functional dependence of the solution $u(t, x, \mathbf{y}(\omega))$ on the random vector $\mathbf{y}$ by deterministic approximation strategies. Widely used is the generalized Polynomial Chaos (gPC) expansion, in which the solution is expanded over a fixed stochastic basis of global multivariate polynomial functions which are orthogonal with respect to the density function of the input random parameters. This idea is motivated by the fact that the parameters-to-solution map is often very smooth for several types of random PDEs. More precisely, under the assumption that the random solution $u(t)$ belongs to $\mathscr{H} \otimes L^2(\Omega)$ for all $t \in [0, T]$, where $L^2(\Omega)$ denotes the space of all square integrable random variables, the gPC-approximate solution is sough in the form:

$$u_S^{gPC}(x, t, \omega) = \sum_{i=1}^{S} U_i(x, t) Y_i(\mathbf{y}(\omega)) \tag{1.4}$$

where $Y_1, ..., Y_S \in [L^2(\Omega)]^S$ is a set of polynomial basis functions in the variables $\mathbf{y}$ orthonormal with respect of the underlying (joint) probability measure of the random vector $\mathbf{y}$. The coefficients $U_1, ..., U_S$ can be computed e.g. by Stochastic Collocation on tensor or sparse grids of Gauss points [5, 135, 75] or Galerkin projection [6, 49, 89, 11]. The first consists in collocating the governing equation (1.1) into a set of points, such as tensor or sparse grids of Gauss points, computing the corresponding solutions and building a global polynomial approximation using such solutions. The second is a projection strategy which aims at computing the coefficients in the gPC expansion by Galerkin projection in the stochastic variables. This generally yields to a large coupled system of differential equations.

The Polynomial Chaos (PC) was first proposed in [131, 22] to discretize Gaussian densities with Hermite polynomials, lately generalized in [136] to any arbitrary random distribution, and successfully used e.g. [121, 135]. Approximation methods based on the gPC expansion work effectively for many classes of problems for which the solution features an analytical dependence with respect to $\mathbf{y}$, see [38] for (steady) elliptic problems and [100, 118] for linear parabolic equations with random coefficients. On the other hand, numerically challenging are problems for which the parameter-to-solution map has low regularity and the stochastic space has high (infinite) dimension. Indeed, despite the great improvements obtained with the use of sparse techniques, spectral methods are often exposed to the so called curse of dimensionality, i.e. an error versus cost performance that is negatively affected by the high dimension of the underlying stochastic space, and degenerates in the limit of infinite dimension. An additional issue concerning the effectiveness of the gPC approach is due to long time integrations. Indeed, as reported in literature [128], for certain classes of evolution equations, the dependence of the solution on the random parameters may significantly vary in time, and the approximation on a fixed polynomial basis might need an increasing number of terms in the expansion

(1.4), to keep an acceptable accuracy level for long time integration. Traveling waves with random speed are a classical example in which the gPC approximation with fixed number of terms, generally fails to provide acceptable solutions after short times. To fix ideas consider the following simple problem describing a random wave in the physical domain $[0, 2\pi]$, with periodic boundary conditions and a sinusoidal initial condition:

$$
\begin{cases}
\partial_{tt} u - y^2(\omega)\partial_{xx} u = 0, \\
u|_{t=0} = e^{ikx}, \quad k \in \mathbb{N}, \quad \partial_t u|_{t=0} = 0,
\end{cases}
\quad \text{in } [0, 2\pi]
\tag{1.5}
$$

where we assume y to be a uniformly distributed random variable in $[-1, 1]$, meaning that the wave travels with a velocity which may uniformly vary between -1 and 1. The analytical solution is given by $u(x, t, \omega) = \frac{1}{2}\left(e^{ik(x+y(\omega)t)} + e^{ik(x-y(\omega)t)}\right) = e^{ikx}\cos(ky(\omega)t)$. In order to effectively approximate the solution at time t on a spatial grid of size h and on a polynomial space of degree p, one should choose $h \approx k^{-1}$ and $p \approx kt/\pi$. Hence the total number of degrees of freedom is $N \approx p/h = k^2 t/\pi$ which shows that the simulation might become unfeasible for large t and k. This phenomenon has been well highlighted in [128]. Several adaptive techniques have been proposed in the literature to (partially) overcome this problem, e.g. time-dependent gPC [48, 59].

1.2.3 Proper Generalized Decomposition

The Proper Generalized Decomposition (PGD) [4, 33, 72, 104, 106] is a model reduction methodology for the approximation of multidimensional PDEs, based on the use of separated representations. Generally, the solution, which is defined in a M-dimensional domain, is approximated by iteratively building a sum of products of M one-dimensional functions, each one defined in a different space dimension. These functions are not known a priori, but constructed iteratively on-line.

The PGD approach was initially introduced in [72] and used to approximate time dependent deterministic PDEs by using a separated representation of the space and time coordinates, i.e. $u(\mathbf{x}, t) \approx u_S = \sum_{i=1}^{S} U_i(\mathbf{x})T_i(t)$, with $\mathbf{x} = (x_1, ..., x_M) \in D_M \subset \mathbb{R}^M$. This technique has been lately extended in [4] to the following more general separated representation:

$$
u(x_1, ... x_M, t) \approx u_S = \sum_{i=1}^{S} \left(\alpha_i \big[\prod_{k=1}^{M} U_{k,i}(x_k) \big] U_{M+1,i}(t) \right)
$$

where $U_{k,i}$ is the ith basis function which depends only on the kth coordinate. The same approach has been adopted in [32] for the approximation of parameter-dependent problems in the R-dimensional parameter space $\Omega_R \subset \mathbb{R}^R$. The solution procedure consists in assuming a space-time-parameter separated representation, i.e.

$$
u(x_1, ..., x_M, t, y_1, ..., y_R) \approx u_S = \sum_{i=1}^{S} \left(\alpha_i \big[\prod_{k=1}^{M} U_{k,i}(x_k) \big] U_{M+1,i}(t) \big[\prod_{j=1}^{R} Y_{j,i}(y_j) \big] \right),
$$

11

where $y_1, ..., y_R$ denote the variables in the parameter space Ω_R and $Y_{j,i}$ is the ith basis function which depends only on the jth parameter y_j. The numerical scheme consists of an iterative procedure, each Sth iteration aims at build the new set of bases functions $U_{S+1,1}, ... U_{S+1,M+1}$, $Y_{S+1,1}, ... Y_{S+1,R}$ and consists of the following 3 steps:

- projection of the solution onto a discrete basis: assumed that the bases $U_{S,1}, ... U_{S,M+1}$ and $Y_{S,1}, ... Y_{S,R}$ are known, the coefficients $\alpha_1, ..., \alpha_S$ are computed by performing a Galerkin projection of the governing equation into the space-time-parameter subspace spanned by $U_{S,1}, ..., U_{S,M+1}$, $Y_{S,1}, ... Y_{S,R}$;

- checking the convergence: this is done by computing the residual of the governing equation; if the residual if smaller then the required level of accuracy the process stops;

- enrichment of the approximation basis: this is achieved by considering the weak form of the governing equation in the space-time-parameter domain. As the problem of calculating the new basis $U_{S+1,1}, ... U_{S+1,M+1}$, $Y_{S+1,1}, ... Y_{S+1,R}$ is nonlinear, the use of an appropriate linearization scheme is needed. The simplest consists of using an alternated direction fixed point algorithm.

In UQ context, the PGD method was introduced in [102] and initially called Generalized Spectral Decomposition (GSD). In this context the approximate solution is sought in a separated form of the stochastic and deterministic variables and the problem of the enrichment of the approximation basis has been interpreted as a pseudo eigenproblem. This interpretation has led to the development of dedicated algorithms inspired from solution techniques for classical eigenproblems [103]. More recently, the PGD has been applied to the solution of high dimensional stochastic/parametric problems, with the introduction of suitable hierarchical tensor representations and associated algorithms [105].

1.2.4 Dynamical Low-Rank approximations

The Dynamically Orthogonal (DO) method is a reduced order technique which aims at overcoming the numerical challenges due to the course of dimensionality and long time integration. It has been first proposed in [116] for the approximation of problems governed by differential equations of type (1.1) and applied in the context of ocean dynamics [117, 127]. Equivalent approximations, with different formulations of the approximate solution, have been derived in [30, 31, 35], (Dynamically double orthogonal (DDO) and Bi-Orthogonal (BO) approximations). Further developments and adaptive strategies are currently under study [36, 14].

The DO approximate solution to problem (1.1) is written as a linear combination of S terms and has the following general form:

$$u_S(x, t, \omega) \quad = \bar{u}_S(x, t) + \sum_{i=1}^{S} U_i(x, t) Y_i(t, \omega) \tag{1.6}$$

where:

- $\bar{u}_S(x, t) \in \mathscr{H}$ is an approximation of the mean of the exact solution;

- $U_1, ..., U_1$ are $L^2(D)$-orthonormal deterministic functions in $\mathscr{H}$, at each t;

- $Y_1, ..., Y_S$ are square integrable random variables with zero mean and full rank covariance matrix.

By imposing the following constraints to the dynamics of the deterministic modes:

$$\langle \dot{U}_i(t), U_j(t) \rangle = 0 \qquad i, j = 1, ..., S \tag{1.7}$$

one can derive a set of nonlinear differential equations which is solved (on line) at each time, to determine the evolution of all the terms in (1.6). Thus, u_S provides an approximation to the whole set of solutions corresponding to all possible realizations of the input random parameters, at each time. In contrast with reduced order methods of type RB or gPC, both the deterministic and stochastic modes evolve in time in order to guarantee more flexibility to the approximation. The analysis and development of reduced order approximations of type DO is the central subject of this thesis. The mathematical details about the DO formulation will be briefly introduced in Section 1.4.1 and exhaustively discussed in the following Chapters.

Low-rank approximations with dynamical modes have been widely studied and used in quantum mechanics to approximate deterministic time-dependent Schrödinger equations. In this context, we recall the time-dependent Hartree method in which the wave function $\psi(x_1, \ldots, x_d)$ is approximated in separable form as $\psi \approx \alpha(t)\phi_1(x_1)\cdots\phi_d(x_d)$ and its generalization, known as multi-configuration time-dependent Hartree (MCTDH) method [92, 12, 65, 138] in which the solution is sought as a linear combination of terms in separable form.

In the finite dimensional setting, the same approach is known as Dynamical Low Rank approximation and used in [66, 56] for low rank approximation of evolution matrix equations. Similarly in [67, 39, 68] the same approach is used for the low rank approximation of evolution tensor equations in Tucker format, with a construction closely related to the one used for the MCTDH method in quantum physics. Extensions to other low-rank formats as Hierarchical Tucker (HT) or Tensor Train (TT) have also been investigated [84, 81]. We mention also the link with the algorithm proposed in [9] for updating singular value decompositions of time-varying matrices and the structured low-rank approximation considered in [71, 19] for time-independent matrices. Related formulations have been investigated in [69] for optimization problems with Low-rank tensor structure, arising for example in the reconstruction of high-dimensional data set.

1.3 Notation

In this section, we introduce the main notations which will be employed hereafter.

Let D be an open and bounded physical domain in $\mathbb{R}^d$, $1 \leq d \leq 3$ with Lipschitz continuous boundary ∂D, and let x stand for the spatial deterministic variable. We denote by $L^2(D)$ the Hilbert space of all square integrable (according the Lebesgue measure) real valued functions defined in D:

$$L^2(D) = \left\{u : D \to \mathbb{R} \text{ s.t. } \int_D |u(x)|^2 dx < \infty\right\}$$

endowed with the L^2 inner product defined as:

$$\langle u, v \rangle := \int_D u(x) v(x) dx \qquad \forall u, v \in L^2(D)$$

and associated L^2-norm: $\|u\|^2 = \sqrt{\langle u, u \rangle}$. Similarly, $H^1(D)$ denotes the space of square integrable, real valued functions with square integrable partial derivatives in D:

$$H^1(D) := \left\{u : D \to \mathbb{R} \,|\, \int_D |u(x)|^2 + |\nabla u(x)|^2 dx < \infty\right\}.$$

and $H_0^1(D)$ the subspace of $H^1(D)$- functions which vanish on the boundary ∂D.
A vector-valued random field will be denoted by small bold letters $\mathbf{u} := (u_1, .., u_N)^T$ and is conventionally a column vector. In particular, $H_{div}^1(D)$ denotes the following space:

$$H_{div}^1(D) := \left\{\mathbf{v} \in [H^1(D)]^d : \nabla \cdot \mathbf{v} = 0\right\},$$

In analogy with the real case, we denote by $L^2(D, \mathbb{C})$ $\left(H^1(D, \mathbb{C})\right)$ the Hilbert space of square integrable, complex valued functions (with square integrable partial derivatives) on D. We use the notation $\langle \cdot, \cdot \rangle_h$ to distinguish the Hermitian product of $L^2(D, \mathbb{C})$, defined as:

$$\langle \hat{u}, \hat{v} \rangle_h := \langle u^q, v^q \rangle + \langle u^p, v^p \rangle + i(\langle u^p, v^q \rangle - \langle u^q, v^p \rangle) \qquad \forall \hat{u} = u^q + i u^p, \hat{v} = v^q + i v^p \in L^2(D, \mathbb{C}),$$

$$(1.8)$$

from the real L^2 product.
Let $\mathbb{F}$ stand for $\mathbb{R}$ or $\mathbb{C}$, we introduce the following definitions, given for any $S \in \mathbb{N}$:

Definition 1.3.1. *We call Stiefel manifold, denoted by $St(S, H^1(D, \mathbb{F}))$, the collection of all $L^2(D)$-orthonormal frames of S functions in $H^1(D, \mathbb{F})$, i.e.:*

$$St(S, H^1(D, \mathbb{F})) = \left\{\mathbf{V} = (V_1, ..., V_S) : V_i \in H^1(D, \mathbb{F}) \text{ and } \langle V_i, V_j \rangle_* = \delta_{ij} \,\forall i, j = 1, ..., S\right\} \quad (1.9)$$

where $\langle \cdot, \cdot \rangle_*$ is the real L^2 product if $\mathbb{F} = \mathbb{R}$ and the hermitian product if $\mathbb{F} = \mathbb{C}$.

Observe that any element of $\mathbf{V} \in St(S, H^1(D, \mathbb{F}))$ is a S-dimensional basis in $H^1(D, \mathbb{F})$ which

is orthonormal with respect to the $L^2(D, \mathbb{F})$-product. For non orthonormal bases we use the notation

$$B(S, H^1(D, \mathbb{F})) = \{\mathbf{U} = (U_1, ..., U_S) : U_i \in H^1(D, \mathbb{F}) \text{ and } \mathbf{M} \in \mathbb{F}^{S \times S}, \quad \mathbf{M}_{ij} = < U_i, U_j >_*,$$
$$\text{with rank}(\mathbf{M}) = S\}$$

with the same convention on $\langle \cdot, \cdot \rangle_*$.

Definition 1.3.2. *We define $\mathscr{G}(S, H^1(D, \mathbb{F}))$ the Grassmann manifold of dimension S in $H^1(D, \mathbb{F})$, which consists of all the $S-$dimensional linear subspaces of $H^1(D, \mathbb{F})$.*

When $\mathbb{F}$ is omitted, we always assume $\mathbb{F} = \mathbb{R}$.
We recall that the Poisson operator in $[H^1(D)]^2$, here denoted by $\mathscr{J}_2$, is the linear map:

$$\mathscr{J}_2 : \quad [H^1(D)]^2 \quad \rightarrow [H^1(D)]^2$$
$$\mathbf{u} \qquad \mapsto \mathscr{J}_2(\mathbf{u}) = \mathscr{J}_2 \begin{bmatrix} u^1 \\ u^2 \end{bmatrix} := \begin{bmatrix} 0 & I_d \\ -I_d & 0 \end{bmatrix} \begin{bmatrix} u^1 \\ u^2 \end{bmatrix} = \begin{bmatrix} u^2 \\ -u^1 \end{bmatrix}$$

where I_d is the identity operator in $[H^1(D)]^2$.

Definition 1.3.3. *We denote $U(S, [H^1(D)]^2)$ the manifold of all L^2-orthonormal symplectic bases in $[H^1(D)]^2$, i.e.:*

$$U(S, [H^1(D)]^2) = \left\{ \mathbf{U} \in [H^1(D) \times H^1(D)]^{2S} \text{ such that} \quad \begin{aligned} \langle \mathbf{U}_j, \mathbf{U}_i \rangle &= \delta_{ij} \\ \langle \mathbf{U}_i, \mathscr{J}_2 \mathbf{U}_j \rangle &= (\mathbf{J}_{2S})_{ij} \end{aligned} \right\}, \qquad (1.10)$$

where $\mathbf{J}_{2S} \in \mathbb{R}^{2S \times 2S}$ is the Poisson matrix, i.e. $\mathbf{J}_{2S} = \begin{pmatrix} 0 & \mathbb{I}_S \\ -\mathbb{I}_S & 0 \end{pmatrix}$ and $\mathbb{I}_S$ is the identity matrix in $\mathbb{R}^{S \times S}$.

Let $(\Omega, \mathscr{A}, \mathscr{P})$ be a complete probability space, where Ω is the set of outcomes, $\mathscr{A}$ a σ-algebra and $P : \mathscr{A} \rightarrow [0, 1]$ a probability measure. Let $y : \Omega \rightarrow \mathbb{F}$ be an integrable random variable; we define the mean of y as:

$$\bar{y} = \mathbb{E}[y] = \int_\Omega y(\omega) d\mathscr{P}(\omega).$$

The symbol $L^2(\Omega, \mathbb{F})$ (respectively $L_0^2(\Omega, \mathbb{F})$) denotes the Hilbert space of $\mathbb{F}$-valued square integrable random variables (respectively with zero mean), that is:

$$L^2(\Omega, \mathbb{F}) := \left\{ y : \Omega \rightarrow \mathbb{F} : \mathbb{E}[y^2] = \int_\Omega (y(\omega))^2 d\mathscr{P}(\omega) < \infty \right\}$$

In this thesis we always assume that the stochastic space can be accurately parametrized in terms of a finite dimensional vector $\mathbf{y} : \Omega \rightarrow \mathbb{R}^N$.

We also recall that $L^2(D \times \Omega, \mathbb{F})$ denotes the space of all square integrable random fields:

$$L^2(D \times \Omega, \mathbb{F}) := \left\{ u : D \times \Omega \to \mathbb{F} \text{ s.t. } \mathbb{E}\big[\|u\|^2_{L^2(D,\mathbb{F})}\big] < \infty \right\},$$

and it is isometrically isomorphic to the tensor product space $L^2(D, \mathbb{F}) \otimes L^2(\Omega, \mathbb{F})$.
In analogy to Definition 1.3.1 and Definition 1.3.2, we define:

Definition 1.3.4. *We call Stiefel manifold, denoted by $St(S, L^2(\Omega))$, the collection of all $L^2(\Omega)$-orthonormal frames of S functions in $L^2(\Omega)$, i.e.:*

$$St(S, L^2(\Omega, \mathbb{F})) = \big\{ \mathbf{Y} = (Y_1, ..., Y_S) : Y_i \in L^2(\Omega, \mathbb{F}) \text{ and } \mathbb{E}[\mathbf{Y}_i \mathbf{Y}_j] = \delta_{ij} \ \forall i, j = 1, ..., S \big\} \qquad (1.11)$$

For non orthonormal bases we use the notation

$$B(S, L^2(\Omega, \mathbb{F})) = \big\{ \mathbf{Y} = (Y_1, ..., Y_S) : Y_i \in L^2(\Omega, \mathbb{F}) \text{ s.t. } \mathrm{rank}\big(\mathbb{E}[\mathbf{Y}\mathbf{Y}^T]\big) = S \big\}.$$

Definition 1.3.5. *We define $\mathcal{G}(S, L^2(\Omega, \mathbb{F}))$ the Grassmann manifold of dimension S in $L^2(\Omega)$ which consists of all the $S-$dimensional linear subspaces of $L^2(\Omega, \mathbb{F})$.*

1.4 Overview and main results

1.4.1 Variational Formulation for the DLR approximation

The first goal of this thesis has been the formalization of the variational setting of the DO method proposed [116, 117]. The aim has been to provide the basis for a suitable mathematical analysis of the DO method. By formalizing the link between the DO and the DLR method [66, 56], we reinterpreted the DO approximation as a Galerkin projection onto the tangent space to the approximation manifold (collection of all functions which can be expressed as a sum of S linearly independent deterministic modes combined with S linearly independent stochastic modes) along the approximate trajectory of the solution. This formulation has allowed us to set in a unified framework, different variants of the DO method that have been proposed in the literature as e.g. the DyBO method [30, 31]. The main achievement obtained in this direction is a quasi-optimal theoretical bound for the DO approximation of linear parabolic equations with random data. The result is inspired by the analogous one obtained in [39] for the MCTDH method (or in [66] for the DLR approximation), and adapted to the context of parabolic PDEs with random parameters.
We introduce here the DLR variational problem, which will be fully discussed in Chapter 3, and the Dual DO formulation which will be used in Chapter 4. For this aim, in this section, we assume that problem (1.1) is completed with Dirichlet homogeneous boundary conditions.

Consider the following manifold $\mathcal{M}_S$, for some fixed $S \in \mathbb{N}$.

Definition 1.4.1. *We call manifold of rank S the subset* $\mathcal{M}_S \subset \mathcal{H} \otimes L^2(\Omega)$ *defined as:*

$$\mathcal{M}_S = \Big\{ u_S \in \mathcal{H} \otimes L^2(\Omega) : u_S = \sum_{i=1}^{S} U_i Y_i \mid \quad \begin{array}{l} span(U_1, ..., U_S) \in \mathcal{G}(S, \mathcal{H}), \\ span(Y_1, ..., Y_S) \in \mathcal{G}(S, L^2(\Omega)) \end{array} \Big\} \tag{1.12}$$

and S rank random field any element $u_S \in \mathcal{M}_S$.

We define Dynamical Low Rank (DLR) approximation of rank S of problem (1.1) a function $u_S \in \mathcal{M}_S$ which satisfies the following variational principle at any time:

DLR Variational Principle. *At each* $t \in (0, T]$, *find* $u_S(t) \in \mathcal{M}_S$ *such that:* $u_S(0) = u_{0,S}$ *and*

$$\mathbb{E}[\langle \dot{u}_S(\cdot, t, \cdot) - \mathcal{L}(u_S(\cdot, t, \cdot), t, \omega), v \rangle] = 0, \quad \forall v \in \mathcal{T}_{u_S(t)} \mathcal{M}_S \tag{1.13}$$

where $u_{0,S}$ *is a suitable S rank approximation of* u_0 *by e.g. a truncated Karhunen-Loève expansion and* $\mathcal{T}_{u_S(t)} \mathcal{M}_S$ *is the tangent space to* $\mathcal{M}_S$ *at* $u_S(t)$.

The DLR variational problem is essentially a Galerkin projection of the governing equation (1.1) onto the tangent space to $\mathcal{M}_S$ at $u_S(t)$ at any time. In quantum mechanic this is known as Dirac-Frenkel time-dependent variational principle (see e.g. [80]) and leads to the MCTDH method [39, 67, 7] for the approximation of deterministic time-dependent Schrödinger equations.

What we have defined in (1.12) and (1.13) is an abstract manifold and an abstract variational principle. For computational reasons we need to:

- represent any elements of $\mathcal{M}_S$ in terms of deterministic (**U**) and stochastic (**Y**) modes, e.g. by using form (1.6);

- make sure that such representation is unique.

This is obtained by equipping $\mathcal{M}_S$ with a differential structure of quotient manifold. In particular, the choice of the equivalent relation (defining the quotient operation) should guarantee a unique parametrization of the tangent space to $\mathcal{M}_S$ at u_S in terms of separable variations in **U** and **Y**. The prerequisite of differential geometry and the differential construction of the approximation manifold $\mathcal{M}_S$ are discussed in Chapter 2.
Essentially, $\mathcal{M}_S$ admits many equivalent differential structures, which depend on the choice of the equivalent relation, and which lead to different parametrizations of the tangent space $\mathcal{T}_{u_S(t)} \mathcal{M}_S$. Based on the parametrization of $\mathcal{T}_{u_S(t)} \mathcal{M}_S$, one can derive from (1.13) different reduced systems, which however lead to the same approximate solutions.

In this thesis we use two alternative formulations:

- **DO decomposition** with orthonormal deterministic modes as proposed in [116, 117] in which any S rank random field is written as:

$$u_S(x,\omega) = \sum_{i=1}^{S} \tilde{Y}_i(\omega)\tilde{U}_i(x) = \tilde{\mathbf{U}}\tilde{\mathbf{Y}} \tag{1.14}$$

where:

- $\tilde{\mathbf{U}}$ is a row vector of $L^2(D)$−orthonormal deterministic functions in $\mathscr{H}$,
- $\tilde{\mathbf{Y}}$ is a column vector of S random variables with full rank covariance matrix $\mathbf{C} = \mathbb{E}[\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^T]$.

Let $\mathscr{O}_S$ be the subspace of all the orthogonal matrices of dimension S: $\mathscr{O}_S = \{O \in \mathbb{R}^{S \times S} : \mathbf{O}^T\mathbf{O} = \mathbf{OO}^T = \mathbb{I}\}$. By using the diffeomorphism between $\mathscr{M}_S$ and $\left(St(S, H^1(D))/\mathscr{O}_S\right) \times B(S, L^2(\Omega))$, one recovers the dynamical constraints (1.7) which allow to uniquely parametrize the tangent space to $\mathscr{M}_S$ at each point.

- **Dual DO decomposition** with uncorrelated stochastic modes, in which any S rank random field is written as:

$$u_S(x,\omega) = \sum_{i=1}^{S} Y_i(\omega)U_i(x) = \mathbf{UY} \tag{1.15}$$

where:

- $\mathbf{U}$ is a row vector of S linearly independent deterministic functions. Namely, $\mathbf{M} \in \mathbb{R}^{S \times S}$, defined as $\mathbf{M}_{ij} = \langle U_j, U_i \rangle$, is a full rank matrix.
- $\mathbf{Y}$ is a column vector of S $L^2(\Omega)$−orthonormal random variables.

By using the diffeomorphism between $\mathscr{M}_S$ and $B(S, \mathscr{H}) \times \left(St(S, L^2(\Omega))/\mathscr{O}_S\right)$, one gets the following dynamical constraints:

$$\mathbb{E}[\dot{Y}_i(t)Y_j(t)] = 0 \qquad \forall i, j = 1, ..., S \tag{1.16}$$

which allow to uniquely parametrize the tangent space to $\mathscr{M}_S$ at each point.

For convenience here, the DO or the Dual DO decompositions are defined without isolating the mean (i.e. $u_S = \mathbf{UY}$ instead of $u_S = \bar{u}_S + \mathbf{UY}$). However, for problems with homogeneous Dirichlet boundary conditions, both the formulations, by isolating the mean or without isolating the mean, can be adopted and lead to similar mathematical constructions. In practice, the formulation by isolating the mean implies that the stochastic mode associated to $\bar{u}_S$ is fixed "a priori" equal to 1. Thus the term $\bar{u}_S$ is extracted from the approximation manifold $\mathscr{M}_S$ and is determined by performing a Galerkin projection into the subspace spanned by $Y_0 = 1$, i.e. the subspace of all deterministic functions in $\mathscr{H}$, which simply corresponds to averaging the governing equation. For the term $\mathbf{UY}$ everything applies as discussed before, with the only

change that the stochastic modes are now constrained to have zero mean. This formulation may be worth in UQ context as we are typically interested in computing centered statistics of the solution.

According to the parametrization of the tangent space, one derives from (1.13) a set of non-linear differential equations for the deterministic and stochastic bases $(\mathbf{U}, \mathbf{Y})$, suitable for numerical integration. For instance, for the dual DO formulation we have:

Proposition 1.4.1. *Let* $(\mathbf{U}(t), \mathbf{Y}(t)) \in B(S, H^1(D)) \times St(S, L^2(\Omega))$ *be a solution of the following system:*

$$
\begin{cases}
\dot{U}_i(x, t) = \mathbb{E}\left[\mathscr{L}(u_S(x, t, \cdot)) Y_i(t, \cdot)\right] & (x, t) \in \partial D \times (0, T], \ i = 1, \cdots, S \\
U_i(x, t) = 0 & (x, t) \in \partial D \times (0, T], \ i = 1, ..., S \quad (1.17) \\
\sum_{i=1}^{S} \mathbf{M}_{ji}(t) \dot{Y}_i(t, \omega) = \Pi_{\mathscr{Y}}^{\perp} \langle \mathscr{L}(u_S(\cdot, t, \omega)), U_i(\cdot, t) \rangle & (t, \omega) \in (0, T] \times \Omega, \ j = 1, \cdots, S
\end{cases}
$$

with $U_1(x, 0) = U_{0i}$, $Y_i(t, \omega) = Y_{i0}(\omega)$ *and* $u_{0S} = \sum_{i=1}^{S} U_{i0} Y_{i0}$ *the truncated Karhunen-Loève expansion of* u_0, *then* $u_S(t) = \sum_{i=1}^{S} U_i(t) Y_i(t) \in \mathcal{M}_S$ *satisfies the DLR variational principle (1.13) at any* $t \in (0, T]$.

Here $\Pi_{\mathscr{Y}}$ is the orthogonal projection operator from the space $L^2(\Omega)$ to the S dimensional subspace $\mathscr{Y} = \text{span}\{Y_1, \cdots, Y_S\}$ and $\Pi_{\mathscr{Y}}^{\perp} \nu = (\mathbb{I} - \Pi_{\mathscr{Y}})(\nu)$ denotes the projection onto the orthogonal complement of $\mathscr{Y}$. We call (1.17) Dual DO reduced system.

In Chapter 3 we propose an error analysis for DLR approximation, by following the analogous one obtained in [39] for the MCTDH method. In particular we show that the DLR approximation error for a linear parabolic equation with random input data can be bounded in terms of the best rank S approximation of the solution (truncated Karhunen-Loève expansion), at each time instant. More precisely, under mild extra requirements on the data of the problem, the following holds:

Theorem. *Suppose that the best S-rank approximation $z_S(t)$ of the exact solution $u(t)$ is continuously differentiable in time for $0 \le t \le \bar{t}$ and the smallest singular value of $z_S(t)$ is uniformly bounded from below, with lower bound $\sigma(z_S(t)) \ge \rho > 0, \forall t \in [0, \bar{t}]$. Then there exists $0 < \hat{t} \le \bar{t}$ such that the approximation error of the DO approximate solution u_S, with initial value $u_S(0) = z_S(0)$, is bounded by*

$$
\|u_S(t) - z_S(t)\|_0^2 + a_{min} \int_0^t |u_S(\tau) - z_S(\tau)|_1^2 d\tau \le 2\alpha e^{2\beta(t)} \int_0^t \|z_S(\tau) - u(\tau)\|_1^2 d\tau, \quad (1.18)
$$

for all $0 < t \leq \hat{t}$, with

$$\beta(t) = 4\rho^{-1} \int_0^t \left(4\|\mathscr{L}(z_S(\tau))\|_0 + \|\mathscr{L}(u(\tau))\|_0 + \|\mathscr{L}(u_S(\tau))\|_0 + \|\dot{z}_S(\tau)\|_0^2 \right) d\tau,$$
$$\alpha = \max\left\{ \frac{a_{max}^2}{2a_{min}}, 4\rho^{-1} \right\},$$

where a_{min}, a_{max} are the coercivity and continuity constants of the elliptic operator $\mathscr{L}$ and $\|.\|_1, |.|_1$ denote respectively the norm and semi-norm in $H^1(D) \otimes L^2(\Omega)$.

The central ingredient for the proof is the use of curvature estimates for approximation manifold $\mathscr{M}_S$, which have been derived in [39]. The error bound is applicable for full rank DLR approximate solutions on the largest time interval in which the best $S-$terms approximation is continuously differentiable in time. The request on time differentiability is actually unavoidable and corresponds to asking that certain eigenvalues of the Karhunen-Loève decomposition do not cross in time.

The possibility of extending the error analysis to approximate solutions with deficient rank remains an open issue. The problem of the rank-deficiency is also the main obstacle in the analysis of the existence and uniqueness of the approximate solution. The only results available in literature concern the approximation of deterministic Schrödinger equations [67, 126] by the MCTDH method. Global in time results have been derive in [8, 7] by explicitly exploiting the fact that the conservation of energy guarantees the full rank condition for the approximate solution of Schrödinger equations, property that does not apply to our setting. In this work, we have not dealt with the analysis of well-posedness of the reduced model and we have simply assumed that a DLR approximate solution exists.

Similarly the stability of the DO reduced system is an open problem due to the fact that the correlation matrix of the deterministic modes $\mathbf{M}_{ij} = < U_i, U_j >$ in 1.17 (similarly the covariance matrix $\mathbf{C} = \mathbb{E}[\mathbf{YY}^T]$ by adopting the DO formulation), that has to be inverted, may become singular or nearly singular at some time instant t. This problem appears immediately if one considers, for instance, a parabolic equation with random coefficients and deterministic initial condition as the correlation matrix at time $t = 0$ will be identically zero. Despite the lack of theory on how to formulate the problem in the case of rank deficiency, we propose in Chapter 3 a numerical strategy to overcome this problem, that leads to satisfactory results. Other approaches have been proposed in [82, 36].

1.4.2 Dual DO approximation for PDEs with random Dirichlet boundary conditions

In Chapter 4 we extend the DLR variation approximation to problems with random boundary conditions of Dirichlet type. The Dual DO formulation recalled in the previous section turns out to be well suited in this context as it allows easily for a strong imposition of the non-homogeneous boundary conditions. The numerical results that we have obtained show that strong imposition of Dirichlet boundary conditions improves considerably the performance of the DLR method with respect to the simple projection of the boundary conditions, as proposed

in [116, 117].

In the analysis of random boundary conditions, we explicitly assume that $\mathscr{L}$ in (1.1) is a second order elliptic operator of the form

$$\mathscr{L}(u) = -div(A(x,\omega)\nabla u) - b(x,\omega) \cdot \nabla u + c(x,\omega)u - f(x,t,\omega)$$

where $A_{ij}(x,\omega), b_i(x,\omega), c(x,\omega), \ i,j = 1,...,d,$ are bounded random variables in the open bounded Lipschitz domain $D \subset \mathbb{R}^d$ and under the assumptions that $A(x,\omega)$ is uniformly coercive almost surely and $f \in L^2((0,T), L^2(D \times \Omega))$. The problem is set in $H^1(D) \otimes L^2(\Omega)$ and completed with random Dirichlet boundary conditions $u_{|\partial D} = g$. This setting has been generalized to vector-valued PDEs.

Problems of this type can be encountered for instance in fluid dynamics, in which small variations on inflow boundary conditions can have a strong impact on the dynamics of the flow. Applications can be found both in engineering and biomedical problems.

In dealing with random boundary data, the issue consists in establishing which boundary conditions should be satisfied by the low rank approximate solution and if and how the randomness coming from the boundary should be compressed. The difficulty is related to the fact that we are not able to say "a priori" which parameters have the strongest impact on the dynamics and at which time the dynamics of the solution is influenced by the uncertain parameters in the boundary data.

We propose in Chapter 4 a modified version of the DLR method which enforces the approximate solution to satisfy the same boundary conditions as the exact solution, or a well controlled approximation of them. This is obtained by setting the DLR variational principle in the constrained manifold of all S rank random fields which satisfy a prescribed value on the boundary, expressed in low rank format. After showing that this set is actually a manifold we characterize its tangent space at each point. The starting assumption in our model is that g, the datum on the boundary, is "almost low rank", which is not very restrictive in our context: since we are looking for an approximate solution u_S of rank S such that $u_S \approx u$, it is reasonable to ask that the boundary value $u_{|\partial D} = g$ is properly approximated in separable form by $g_M = \sum_{i=1}^M Z_i(\omega)v_i(t,x)$ with $M \leq S$.

The constrained manifold is constructed according to the following definition:

Definition 1.4.2. *A S rank random field under constraint g_M is a S rank function which is written as:*

$$u_S^{g_M}(x,\omega) = \sum_{i=1}^S U_i(x)Y_i(\omega) = \mathbf{U}\mathbf{Y} \tag{1.19}$$

and satisfies:

- $u_{S|\partial D} = g_M$ *a. s.,*

- $U_1,...,U_S$ *linearly independent deterministic functions.*

21

- $Y_1, ..., Y_S$ $L^2(\Omega)$-*orthonormal random variables.*

We denote by $\mathscr{M}_S^{g_M}$ the set of all S rank random fields under constraint g_M.

In other words, $\mathscr{M}_S^{g_M}$ is the collection of all S rank random fields which satisfy the condition $u_S = g_M$ on the boundary.

We show that $\mathscr{M}_S^{g_M}$ is actually a differential manifold and that any element $u_S^{g_M} \in \mathscr{M}_S^{g_M}$ can be written as:

$$u_S^{g_M}(x, \omega) = \sum_{i=1}^{R} U_i(x) Y_i(\omega) + \sum_{i=1}^{M} V_i(x) Z_i(\omega) \tag{1.20}$$

where:

- $R + M = S$,

- $u_S^{g_M}(x, \omega) = g_M(x, \omega) = \sum_{i=1}^{M} v_i(x) Z_i(\omega)$ for $x \in \partial D$ a.s.,

- all the random variables are mutually $L^2(\Omega)$-orthonormal:

  – $\mathbb{E}[Z_i Z_j] = \delta_{ij}$ for all $i, j = 1, ..., M$;

  – $\mathbb{E}[Y_i Y_j] = \delta_{ij}$ for all $i, j = 1, ..., R$;

  – $\mathbb{E}[Z_i Y_j] = 0$ for all $i = 1, ..., M$ and for all $j = 1, ..., R$.

- $U_1, ..., U_R$ are linearly independent.

Once $\mathscr{M}_S^{g_M}$ is equipped with a manifold structure, we are allowed to write the DLR variational principle on $\mathscr{M}_S^{g_M}$.

In this setting, the Dual DO formulation allows to derive the proper boundary conditions for each deterministic mode and leads to "strong" imposition of random Dirichlet boundary conditions.

The Dual DO reduced system results in a set of S coupled PDEs for the evolution of the deterministic modes (M of which with non homogeneous boundary conditions) coupled with $S - M$ ODEs for the evolution of the stochastic modes.

1.4.3 DLR approximation for Navier Stokes equations

We have focused on the application of DLR techniques to the incompressible Navier Stokes equations with random parameters. In this context, we propose again the Dual-DO formulation, which is very convenient to include the incompressibility constraint. The Dual DO reduced system results in S deterministic problems of Navier Stokes type, coupled to a system

of (at most) S stochastic ODEs. The DLR formulation for Navier Stokes equations and numerical tests are discussed in Chapter 4 of this thesis.

The problem under study is governed by incompressible Navier Stokes equations with random parameters:

$$
\begin{cases}
\dot{\mathbf{u}}(\mathbf{x}, t, \omega) - \nu(\mathbf{x}, t, \omega)\Delta\mathbf{u}(\mathbf{x}, t, \omega) + \mathbf{u}(\mathbf{x}, t, \omega) \cdot \nabla\mathbf{u}(\mathbf{x}, t, \omega) + \nabla p(\mathbf{x}, t, \omega) = \mathbf{f}(\mathbf{x}, t, \omega) & (\mathbf{x}, t) \in D \times (0, T] \\
\nabla \cdot \mathbf{u}(\mathbf{x}, t, \omega) = 0 & \\
\mathbf{u}(\mathbf{x}, 0, \omega) = \mathbf{u}_0(\mathbf{x}, \omega) & \mathbf{x} \in D \\
\mathbf{u}(\mathbf{x}, t, \omega) = \mathbf{g}(\mathbf{x}, t, \omega) & \mathbf{x} \in \Gamma_D, t \in (0, T] \\
\nu\partial_{\mathbf{n}}\mathbf{u}(\mathbf{x}, t, \omega) - p(\mathbf{x}, t, \omega) \cdot \mathbf{n} = \mathbf{h}(\mathbf{x}, t, \omega) & \mathbf{x} \in \Gamma_N, t \in (0, T]
\end{cases}
$$

$$(1.21)$$

where $\mathbf{u}$ is the velocity (column) vector field, p is the scalar pressure and ν is the kinematic viscosity that may eventually be modeled as a random variable or random field. Γ_D and Γ_N are disjointed parts of the boundary ∂D, such that $\overline{\Gamma}_D \cup \overline{\Gamma}_N = \partial D$, on which we impose Dirichlet and Neumann boundary conditions respectively. The randomness may affect the parameters of the equation such as the fluid viscosity, or the forcing term, initial or boundary conditions. Our goal is to find a low rank approximation of the velocity field.

The DLR approximate solution is sought in the constrained manifold of all S rank random fields with divergence free modes, and which satisfy the condition $\mathbf{u}_S = \mathbf{g}_M$ on the Dirichlet boundary, with $\mathbf{g}_M(x, t, \omega) = \sum_{i=1}^{M} \boldsymbol{v}_i(x, t)Z_i(\omega) \approx \mathbf{g}(x, t, \omega)$, i.e.:

$$
\mathcal{M}_{S,div}^{\mathbf{g}_M(t)} = \left\{ \mathbf{u}_S = \sum_{i=1}^{S} \mathbf{U}_i Y_i \text{ s.t. } \quad \mathbf{u}_{S|_{\Gamma_D}}(t) = \mathbf{g}_M(t), \text{ and } \mathbf{U}_i \in H_{div}^1(D), \right.
$$
$$
\left. \mathbb{E}[Y_i] = 0, \mathbb{E}[Y_i Y_j] = \delta_{ij}, \text{ rank}(\mathbf{M}) = R \right\}
$$

$$(1.22)$$

where $R = S - M$ and $\mathbf{M}$ is the full rank correlation matrix of the first R deterministic modes: $\mathbf{M}_{ij} = <\mathbf{U}_i, \mathbf{U}_j> = \sum_{k=1}^{d} <U_{i,k}, U_{j,k}>$.

We write the DLR variational principle for problem (1.21) in $\mathcal{M}_{S,div}^{\mathbf{g}_M(t)}$. In deriving the Dual DO dynamical system, the divergence free constraint is then imposed on each deterministic mode by introducing S Lagrange multipliers $p_1, ..., p_S$. This leads to the following reduced system:

$$
\begin{aligned}
&\dot{\mathbf{U}}_i + \nabla p_i = \mathbb{E}\big[\big(\nu\Delta\mathbf{u}_S - \mathbf{u}_S \cdot \nabla\mathbf{u}_S + \mathbf{f}\big)Y_i\big] \\
&\nabla \cdot \mathbf{U}_i = 0 & \forall i = 1, ..., S \\
&\mathbf{M}_{ik}\dot{Y}_k = <\mathbf{U}_i, \Pi_{<Y_1,...,Y_S>}^{\perp}\big[\nu\Delta\mathbf{u}_S - \mathbf{u}_S \cdot \nabla\mathbf{u}_S + \mathbf{f}\big]> & \forall i = 1, ..., R
\end{aligned}
$$

$$(1.23)$$

with initial conditions given by the best S rank approximation of $\mathbf{u}_0$ in $\mathcal{M}_{S,div}^{\mathbf{g}_M}$ and boundary
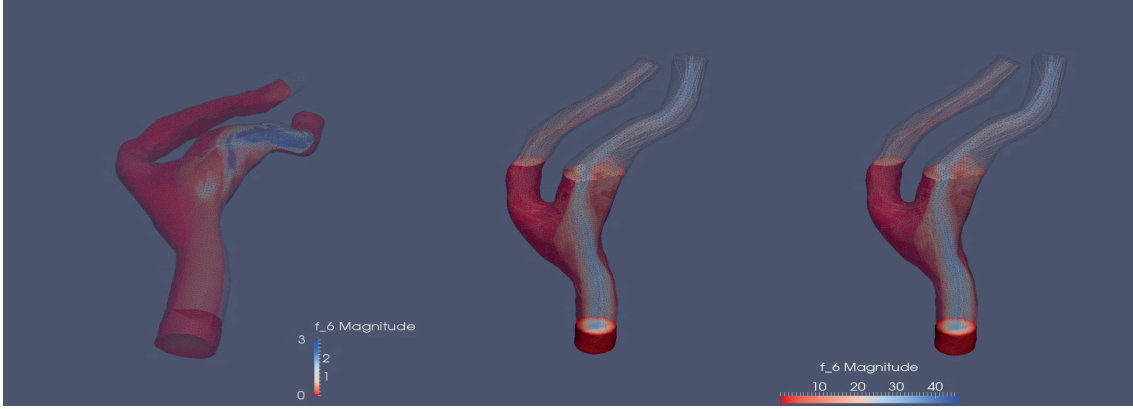
Figure 1.1 – The Dual DO method has been applied to simulate the blood flow in a carotid artery, where the uncertainty affects the flow rate imposed at the inlet. On the left the standard deviation of the solution at time $t = 1.6$ during the second heart beat simulated. On the right we compare the mean of the Dual DO approximate solution computed with 5 modes (right) ($S = 5$) to the mean of the reference solution at the same time. We observe that the approximate solution effectively describes the dynamics and allow to accurately quantify the variability of the solutions.

conditions given by:

$$
\begin{aligned}
\mathbf{U}_i(x,t) &= \boldsymbol{v}_i(x,t) && (\mathbf{x},t) \in \Gamma_D \times (0,T], \ \forall i = 1,...,R \\
\mathbf{U}_i(x,t) &= \mathbf{0} && (\mathbf{x},t) \in \Gamma_D \times (0,T], \ \forall i = R+1,...,S \\
v\partial_{\mathbf{n}}\mathbf{U}_i(\mathbf{x},t) - p_i(\mathbf{x},t)\cdot\mathbf{n} &= \mathbb{E}[\mathbf{h}(\mathbf{x},t,\cdot)Y_i] && (\mathbf{x},t) \in \Gamma_N \times (0,T], \ \forall i = 1,...,S.
\end{aligned}
$$

The Dual DO method has been tested on two fluid dynamics problems. In the first one, our goal is to test the performance of the Dual DO approximation in the challenging case in which the rank of the solution continues to increase in time. We consider the classical benchmark 2D problem of an incompressible viscous fluid flowing around a cylindrical obstacle in a channel at moderate Reynolds numbers.

In the second we address a hemodynamic problem for biomedical applications. In this context, simulations of blood flow using image-based models of the computational domain and computational fluid dynamics has found widespread application to quantifying hemodynamic factors relevant to the initiation and progression of cardiovascular diseases and for planning surgical interventions. In particular, numerical simulations of parameter dependent PDEs can be used as a virtual platform for the prediction of input/output response of biological values. To this aim, the speed up of the computational time is a crucial issue. We consider the problem of simulating blood flow in a realistic carotid artery reconstructed from MRI data, where the inflow boundary conditions are taken as random due to the uncertainty and large errors in Doppler measurements of the inflow velocity profile. See Figure 1.1 for an illustration of some numerical results.

1.4.4 Symplectic DLR for wave equations with random parameters

In Chapter 5 we extend and modify the DLR approach to the approximation of second order wave equations with random parameters.

A critical issue in the context of low-rank approximation of hyperbolic equations is how to construct reduced order systems which preserve the stability and the geometrical properties of the original problem. It has been reported in literature [109], that a POD-based reduced system may indeed become unstable even if the original hyperbolic systems was not. This highlights the need for developing "novel" reduced order techniques, built *ad hoc* to deal with hyperbolic problems. In Chapter 5 we propose a reduced order method with symplectic and dynamical deterministic basis which enjoys the conservation of energy, as the original problem. This approach, which we name Symplectic Dynamical Low Rank (Symplectic DLR) method, is based on recasting the governing wave equation in Hamiltonian form and is designed as a combination of:

- the DLR approximation, which is suited for the approximation of parabolic equations (first order time derivative),

- the symplectic reduced basis technique proposed in [107, 86] for the approximation of parametric Hamiltonian systems, in which the approximate solution is expanded over a set of symplectic deterministic bases, and the reduced system preserves the symplectic structure of the full order system.

The aim is both to preserve the Hamiltonian structure of the original problem and provide adaptivity for long time integration.

We consider the following initial boundary value problem:

$$\begin{cases} \ddot{u}(x,t,\omega) = \nabla \cdot \Big(c(x,\omega) \nabla u(x,t,\omega) \Big) + f(u(x,t,\omega),\omega) & x \in D, \ t \in (0,T], \ \omega \in \Omega, \\ u(x,0,\omega) = p_0(x,\omega) & x \in D, \ \omega \in \Omega, \\ \dot{u}(x,0,\omega) = q_0(x,\omega) & x \in D, \ \omega \in \Omega, \\ u(\sigma,t,\omega) = 0 & \sigma \in \partial D, \ t \in (0,T], \ \omega \in \Omega, \end{cases} \tag{1.24}$$

with homogeneous Dirichlet boundary conditions. Here the randomness may affect the wave speed c as well as the initial conditions p_0, q_0 and the (possibly non linear) source term f. We assume that c is bounded and uniformly coercive and the initial data satisfy: $q_0 \in L^2(\Omega, H_0^1(D))$, $p_0 \in L^2(\Omega, L^2(D))$. This guarantees the existence and uniqueness of the solution $u \in L^\infty((0,T), H_0^1(D) \otimes L^2(\Omega))$ having time derivative $\dot{u} \in L^\infty((0,T), L^2(D) \otimes L^2(\Omega))$ [118, 94].

By introducing the phase space variables $(p,q) = (u,\dot{u})$, problem (1.24) can be recast in

Hamiltonian form as follows:

$$
\begin{cases}
\dot{q}(x,t,\omega) = p(x,t,\omega) & x \in D, \ t \in (0,T], \ \omega \in \Omega, \\
\dot{p}(x,t,\omega) = \nabla \cdot \big(c(x,\omega)\nabla q(x,t,\omega)\big) - f(q(x,t,\omega),\omega) & x \in D, \ t \in (0,T], \ \omega \in \Omega, \\
q(x,t,\omega) = q_0(x,\omega) & x \in D, \ \omega \in \Omega, \\
p(x,0,\omega) = p_0(x,\omega) & x \in D, \ \omega \in \Omega, \\
q(x,t,\omega) = 0 & x \in \partial D, \ t \in (0,T], \ \omega \in \Omega,
\end{cases}
\tag{1.25}
$$

The Hamiltonian energy associated to (1.25) can be defined pointwise in ω as:

$$
H_\omega(q,p) = \frac{1}{2}\int_D \Big(|p|^2 + c(\omega)|\nabla q|^2 + F(q)\Big), \qquad F'(q) = f(q).
$$

Let $\nabla_q H_\omega, \nabla_p H_\omega$ denote functional derivatives of H_ω with respect to q and p respectively, i.e.

$$
\begin{aligned}
\langle \nabla_q H_\omega, \delta q \rangle &= \int_D c\nabla q \nabla \delta q + \int_D f(q)\delta q \\
&= \int_D \big(-\nabla\cdot(c\nabla\cdot q) + f(q)\big)\delta q,
\end{aligned}
\qquad \text{and} \qquad \langle \nabla_p H_\omega, \delta p \rangle = \int_D p\delta p.
$$

Then, for any $\delta q \in H_0^1(D)$, $\delta p \in L^2(D)$, the Hamiltonian system is written in canonical form with respect to $\mathbf{u} = (q,p)$ as:

$$
\begin{cases}
\dot{\mathbf{u}}(x,t,\omega) = \mathscr{J}_2 \nabla H_\omega\big(\mathbf{u}(x,t,\omega),\omega\big), \\
\mathbf{u}(x,0,\omega) = (q_0(x,\omega), p_0(x,\omega)))^T
\end{cases}
\tag{1.26}
$$

for almost every $x \in D$ and $\omega \in \Omega$.

The symplectic DLR approximate solution is sought in the following low rank manifold:

Definition 1.4.3. *We call symplectic manifold of rank S, denoted by $\mathcal{M}_S^{sym}$, the collection of all random fields $\mathbf{u}_S = (q_S, p_S)^T \in [H_0^1(D)]^2 \otimes L^2(\Omega)$ that can be written as: $\mathbf{u}_S = \mathbf{U}\mathbf{Y}$ where*

- *$\mathbf{U} \in U(S, [H^1(D)]^2)$ ($L^2(D)$-orthogonal symplectic basis),*

- *$\mathbf{Y} = Y_1, ..., Y_{2S}$ is a $2S$ dimensional vector of square integrable random variables $Y_i \in L^2(\Omega)$, such that $rank(\mathbb{E}[\mathbf{Y}\mathbf{Y}^T] + \mathbf{J}_{2S}^T\mathbb{E}[\mathbf{Y}\mathbf{Y}^T]\mathbf{J}_{2S}) = 2S$.*

We call symplectic S rank random field any function $\mathbf{u}_S \in \mathcal{M}_S^{sym}$. This can be written component-wise as follows:

$$
q_S(x,\omega) = \sum_{i=1}^S Q_i(x) Y_i(\omega) - \sum_{i=1}^S P_i(x) Y_{S+i}(\omega), \quad p_S(x,\omega) = \sum_{i=1}^S P_i(x) Y_i(\omega) + \sum_{i=1}^S Q_i(x) Y_{S+i}(\omega).
$$

$$\tag{1.27}$$

We impose the following:

Symplectic DLR Variational Principle. *At each $t \in (0, T]$, find $\mathbf{u}_S(t) \in \mathscr{M}_S^{sym}$ such that:*

$$\mathbb{E}\left[\left\langle \mathscr{J}_2 \dot{u}_S + \nabla H_\omega(\mathbf{u}_S, \cdot), \mathbf{v}\right\rangle\right] = 0, \qquad \forall \mathbf{v} \in \mathscr{T}_{\mathbf{u}_S(t)}\mathscr{M}_S^{sym}, \; \forall t \in (0, T], \tag{1.28}$$

with initial conditions given by the symplectic projection of the initial data into $\mathscr{M}_S^{sym}$.

The term $\mathbb{E}\left[\left\langle \nabla H_\omega(\mathbf{u}_S, \cdot), \mathbf{v}\right\rangle\right]$ in (1.28) is interpreted as $\frac{d}{dt}\big|_{t=0}\mathbb{E}[H_\omega(\boldsymbol{\gamma}_S(t))]$, i.e. directional derivative along the curve $\boldsymbol{\gamma}_S(t) \in \mathscr{M}_S^{sym}$ with $\boldsymbol{\gamma}_S(0) = \mathbf{u}_S$ and $\dot{\boldsymbol{\gamma}}_S(0) = \mathbf{v}$.

The Symplectic Variational Principle corresponds to a symplectic projection of the governing equation onto the tangent space to $\mathscr{M}_S^{sym}$ of rank S, along the trajectory of the approximate solution, and has the desirable property of conserving the mean Hamiltonian energy. In Chapter 5 we show that the Symplectic DLR coincides with the complex DRL approximation of Hamiltonian system (1.26) rewritten in complex variables. Thus, recast in complex setting, the symplectic DLR approximation is very closed to the MCTDH method. The necessary prerequisites of symplectic geometry are provided in Section 5.4.

We show that $\mathscr{M}_S^{sym}$ can be equipped with a structure of differential manifold and we parametrize the tangent space in term of dynamical constraints on the deterministic modes. This is achieved by identifying $\mathscr{M}_S^{sym}$ with the S rank manifold of complex valued functions. Then by means of the parametrization of the tangent space in real setting, we derive the dynamical equations for the evolution of the bases $\mathbf{U}, \mathbf{Y}$. Let us denote by $B^{sym}(2S, L^2(\Omega)) \subset [L^2(\Omega)]^{2S}$ the set of all $2S$-vectors $\mathbf{Y} = (Y_1, ..., Y_{2S}) \in [L^2(\Omega)]^{2S}$ which satisfy $\text{rank}(\mathbb{E}[\mathbf{Y}\mathbf{Y}^T] + \mathbf{J}_{2S}^T\mathbb{E}[\mathbf{Y}\mathbf{Y}^T]\mathbf{J}_{2S}) = 2S$, and $\tilde{H}_\omega$ the function:

$$\tilde{H}_\omega: \quad [L^2(\Omega)]^{2S} \to L^2(\Omega)$$
$$\mathbf{Y} \to H_\omega\left(\sum_{i=1}^{2S} \mathbf{U}_i Y_i, \omega\right),$$

then we have:

Proposition 1.4.2. *Let $(\mathbf{U}(t), \mathbf{Y}(t)) \in U(S, [H_0^1(D)]^2) \times B^{sym}(2S, L^2(\Omega))$ be a solution of the following system:*

$$\begin{cases} \dot{\mathbf{Y}} = \mathbf{J}_{2S}\nabla_{\mathbf{Y}}\tilde{H}_\omega(\mathbf{Y}) \\[2mm] \dot{\mathbf{U}}(\mathbb{E}[\mathbf{Y}\mathbf{Y}^T] + \mathbf{J}_{2S}^T\mathbb{E}[\mathbf{Y}\mathbf{Y}^T]\mathbf{J}_{2S}) = \mathscr{P}_{\mathbf{U}^+}^\perp\left[\mathbb{E}[\nabla H_\omega(\mathbf{u}_S)\mathbf{Y}^T\mathbf{J}_{2S}] + \mathbb{E}[\mathscr{J}\nabla H_\omega(\mathbf{u}_S)\mathbf{Y}^T]\right] \end{cases} \tag{1.29}$$

where $H_\omega \circ \phi_{\mathbf{U}}$ with initial conditions given by the complex singular value decomposition of the initial data (q_0, p_0). Then $u_S(t) = \mathbf{U}(t)\mathbf{Y}(t) \in \mathscr{M}_S^{sym}$ satisfies the DLR variational principle (1.28) at any $t \in (0, T]$.

The symplectic DO system (1.29) consists of $2S$ random ODEs coupled to $2S$ deterministic PDEs. However, exploiting the unitary structure of $\mathbf{U}$, we actually need to solve only S PDEs to completely characterize the deterministic basis at any time.

27

1.5   Outline of the thesis

- **Chapter 2**: we derive the geometrical construction of the approximation manifold $\mathcal{M}_S$ of S rank random fields. After recalling the necessary prerequisites of differential geometry, we discuss how to equip $\mathcal{M}_S$ with a structure of differential manifold.

- **Chapter 3**: we define the Dynamical Low-Rank variational principle for the approximation of PDEs with random data and we derive a theoretical bound for the approximation error of the $S-$terms DLR approximate solution by the corresponding S-terms best approximation. This chapter is based on the paper "Error Analysis of the Dynamically Orthogonal Approximation of Time Dependent Random PDEs", published in SIAM Journal on Scientific Computing, in January 2015.

- **Chapter 4**: we extend the Dynamical Low-Rank approach to parabolic PDEs, in particular, incompressible Navier Stokes equations, with random Dirichlet boundary conditions. We propose the Dual DO formulation which is more suited to deal with problems with non-homogeneous random Dirichlet boundary conditions and allows for a strong imposition of them, as well as the incompressibility constraint. We test the method on the classical benchmark of a laminar flow around a cylinder with random inflow velocity, and a biomedical application for simulating blood flow in realistic carotid artery reconstructed from MRI data with random inflow conditions coming from Doppler measurements. This chapter is based on the paper "Dual Dynamically Orthogonal approximation of incompressible Navier Stokes equations with random boundary conditions", appeared as Mathicse report n. 03.2017, in January 2017, and submitted for publications.

- **Chapter 5**: we propose a reduced order technique, Symplectic Dynamical Low-Rank method, for the approximation of wave equations with random parameters. We rewrite the governing equation in a Hamiltonian framework and we expand the approximate solution over a set of few symplectic and dynamical deterministic modes. This satisfies the symplectic projection of the governing Hamiltonian equations into the tangent space to the approximation manifold, along the approximate trajectory. By parametrizing the tangent space, we derive a reduced system of dynamical equations which enjoys the conservation of Hamiltonian energy.

2 Differential Manifolds

In this chapter, we set up the geometrical framework underling the Dynamical Low-Rank approximation. We start by recalling some standard definitions and results of Differential Geometry for Hilbert manifolds and principle fiber bundles. Our focus is the characterization of the tangent space to quotient and abstract manifolds. In particular, we detail how to equip the Grassmann $\mathcal{G}(S, \mathcal{H})$, i.e. the collection of all S dimensional subspaces of a Hilbert space $\mathcal{H}$, with a structure of differential quotient manifold. This construction allows to uniquely characterize the tangent space to the Grassmann in terms of variations in $[\mathcal{H}]^S$ and thus locally parametrize $\mathcal{G}(S, \mathcal{H})$ in terms of S dimensional orthonormal bases in $\mathcal{H}$. These tools are then used to equip the manifold of all fixed rank random fields, with a structure of differential quotient manifold. In particular, in Section 2.1, we recover the orthogonal constraints used [116], and employed hereafter, to derive the Dynamically Orthogonal approximation.

In a nutshell, a differential manifold $\mathcal{M}$ is a set endowed with a differentiable structure. Intuitively, a smooth manifold modeled on some space $\mathcal{H}$ can be regarded as a smoothly curved space, which locally *looks like $\mathcal{H}$*, but globally may have a much richer structure. This means that every point of $\mathcal{M}$ has a neighborhood that can be identified with (uniquely described by) a subset of $\mathcal{H}$ by means of bijective maps, called charts. One needs the ensemble of those charts to get a global description of the manifold. Lines and surfaces are the simplest examples manifolds, respectively of dimension one and two.
Differential manifolds are the suitable underlying framework for several numerical algorithms. In this thesis we consider only manifolds modeled on a (infinite dimensional) Hilbert space $\mathcal{H}$. For a more general setting we refer to [70, 74, 16].

Definition 2.0.1. *Let I be an indexing set and $\mathcal{H}$ a Hilbert space. A set $\mathcal{M}$ is called **manifold** of class C^p $(p \geq 0)$ modeled on $\mathcal{H}$ if it is equipped with a C^p-**atlas**, i.e. a collection of pairs $\{\mathcal{U}_i, \phi_i\}_{i \in I}$ which satisfy:*

- *$\mathcal{U}_i \subset \mathcal{M}$ for all $i \in I$ and the union of U_i covers $\mathcal{M}$, i.e.: $\mathcal{M} \subseteq \bigcup_{i \in I} \mathcal{U}_i$;*

- *$\phi_i(\mathcal{U}_i)$ is an open subset of $\mathcal{H}$ and the map $\phi_i : \mathcal{U}_i \to \phi_i(\mathcal{U}_i)$ is a bijection, for all $i \in I$.*
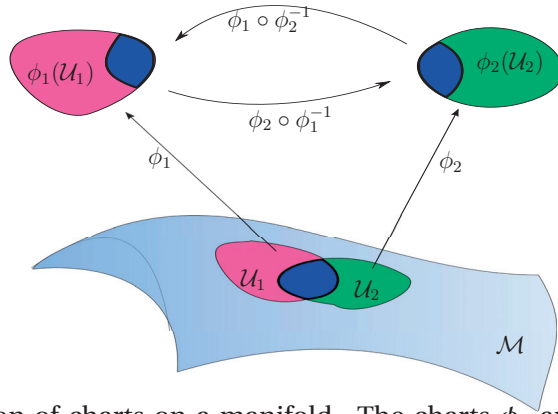
Figure 2.1 – Illustration of charts on a manifold. The charts ϕ_1 and ϕ_2 map $\mathcal{U}_1$ and $\mathcal{U}_2$ respectively in $\phi_1(\mathcal{U}_1)$ and $\phi_2(\mathcal{U}_2)$ and $\phi_1 \circ \phi_2^{-1}$, $\phi_2 \circ \phi_1^{-1}$ are the transition maps defined in $\mathcal{U}_1 \cap \mathcal{U}_2$.

- $\phi_i(\mathcal{U}_i \cap \mathcal{U}_j) \subset \mathcal{H}$ is open and the map:

$$\phi_j \circ \phi_i^{-1} : \phi_i(\mathcal{U}_i \cap \mathcal{U}_j) \to \phi_j(\mathcal{U}_i \cap \mathcal{U}_j)$$

  is a C^p-diffeomorphism (called **transition map**), for all $i, j \in I$.

It is easy to show that the atlas induces an unique topology in $\mathcal{M}$, according to which a subset $V \subset \mathcal{M}$ is open if and only if $\phi_i(V \cap U_i) \subset \mathcal{H}$ is open for all $i \in I$. With this in mind we can equivalently define a manifold $\mathcal{M}$ as a topological space equipped with a C^p-atlas $\{\mathcal{U}_i, \phi_i\}$ such that:

- $\mathcal{U}_i \subset \mathcal{M}$ is open for all $i \in I$ and $\mathcal{M} \subseteq \bigcup_i \mathcal{U}_i$;

- $\phi_i : \mathcal{U}_i \to \phi_i(\mathcal{U}_i)$ is a diffeomorphism into a open subspace of $\mathcal{H}$, for all $i \in I$;

- for all $i, j \in I$ the map:

$$\phi_j \circ \phi_i^{-1} : \phi_i(\mathcal{U}_i \cap \mathcal{U}_j) \to \phi_j(\mathcal{U}_i \cap \mathcal{U}_j)$$

  is a C^p-diffeomorphism.

As the definition shows, a manifold is always characterized by a set $\mathcal{M}$ and an atlas $\{\mathcal{U}_i, \phi_i\}$ that gives to $\mathcal{M}$ a manifold structure. We say that two atlas $\{\mathcal{U}_i, \phi_i\}$ and $\{\mathcal{V}_i, \psi_i\}$ are compatible if $\phi_j \circ \psi_i^{-1}$ is a transition map, for any i, j. The union of all compatible atlas is called maximal atlas. An arbitrary set $\mathcal{M}$ may admit more than one maximal atlas or do not admit any atlas, in this case we say that $\mathcal{M}$ can not be equipped with a manifold structure (see Figure 2.2 right).

Example 2.0.1. *Consider the real space $\mathbb{R}$ and the two maps $\phi_1, \phi_2 : \mathbb{R} \to \mathbb{R}$ defined as $\phi_1 : x \to x$ and $\phi_2 : x \to x^3$, which cover the whole space. One can easily see that the atlas generated by the two maps are not compatible between each others, since $\phi_1 \circ \phi_2^{-1}$ is not differentiable in the*

origin. It follows that ϕ_1, ϕ_2 generate to different maximal atlas. (This shows that $\mathbb{R}$ may admits more than one maximal atlas, However, in this case, the corresponding differential structures are isomorphic.)

Example 2.0.2. *The common intuitive idea of regarding manifolds as smoothly curved spaces may something be misleading. Consider for instance the curve $\alpha : \mathbb{R} \to \mathbb{R}^2$, $\alpha(t) = (t, |t|)$. We wonder if the image $\mathcal{M} := \{\alpha(t), t \in \mathbb{R}\}$ admits a manifold structure. Despite the corner in the origin we can equip $\mathcal{M}$ with a manifold structure isomorphic to $\mathbb{R}$, for instance by defining the chart $\phi(t, |t|) = t$. On the other hand, intuitively because of the "corner", $\mathcal{M}$ is not a submanifold of $\mathbb{R}^2$.*

In view of numerical approximations one may be interested in understanding if a given set (for example the set in which the approximate solution of the problem at hand is sought) admits or not a manifold structure and, if the answer is positive, which one is more natural, or convenient, for the problem under analysis. The following examples and definitions will be used for later developments.

2.0.1 Product of manifolds, sub-manifold, maps between manifolds

Remark 2.0.1. [*Product of manifolds*]
Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two manifolds modeled on $\mathcal{H}_1$ and $\mathcal{H}_2$ with atlas $\{\mathcal{U}_i, \phi_i\}_{i \in I_1}$ and $\{\mathcal{V}_i, \psi_i\}_{i \in I_2}$ respectively. The set $\mathcal{M}_1 \times \mathcal{M}_2$ admits a manifold structure, called product, when equipped with the following collection of charts:

$$
\begin{aligned}
\phi_i \times \psi_j : \mathcal{U}_i \times \mathcal{V}_j \quad &\to \mathcal{H}_1 \times \mathcal{H}_2 \\
(x_1, x_2) \quad &\mapsto (\phi_i(x_1), \psi_j(x_2)).
\end{aligned}
$$

for all $i \in I_1$ and $j \in I_2$

Definition 2.0.2. *Let $\mathcal{M}$ be a C^p manifold modeled on $\mathcal{H}$ with atlas $\{\mathcal{U}_i, \phi_i\}_{i \in I}$. A subset $\mathcal{N} \subset \mathcal{M}$ is called **sub-manifold** of $\mathcal{M}$ if there exists $\mathcal{H}_1$ linear subspace of $\mathcal{H}$ such that $\{\mathcal{U}_i \cap \mathcal{N}, \phi_i\}_{i \in I}$ forms an atlas for $\mathcal{N}$ modeled on $\mathcal{H}_1$, i.e. $\phi_i(\mathcal{U}_i \cap \mathcal{N}) \subset \mathcal{H}_1$ is open for all $i \in I$ and the union of $\mathcal{U}_i \cap \mathcal{N}$ covers $\mathcal{N}$. We say that $\{\mathcal{U}_i, \phi_i\}_{i \in I}$ induces an atlas on $\mathcal{N}$.*

For ease of notation, in the following we will generally omit the index in denoting atlas and charts. Namely we will write $\{\mathcal{U}, \phi\}$ to refer to the atlas $\{\mathcal{U}_i, \phi_i\}_{i \in I}$ and $(\mathcal{U}, \phi)$ when we consider an arbitrary chart of the atlas.

Let $\mathcal{M}_1, \mathcal{M}_2$ be two C^p−manifolds, with $p \geq 1$, modeled on $\mathcal{H}_1$ and $\mathcal{H}_2$ with atlas $\{\mathcal{U}, \phi\}$ and $\{\mathcal{V}, \psi\}$ respectively. The smoothness of a map $F : \mathcal{M}_1 \to \mathcal{M}_2$ is regarded in terms of charts as smoothness in the model spaces $\mathcal{H}_1$ and $\mathcal{H}_2$.

Definition 2.0.3. *A map $F : \mathcal{M}_1 \to \mathcal{M}_2$ is **differentiable** in $x \in \mathcal{M}_1$ if there exists a chart $(\mathcal{U}, \phi) \in$*
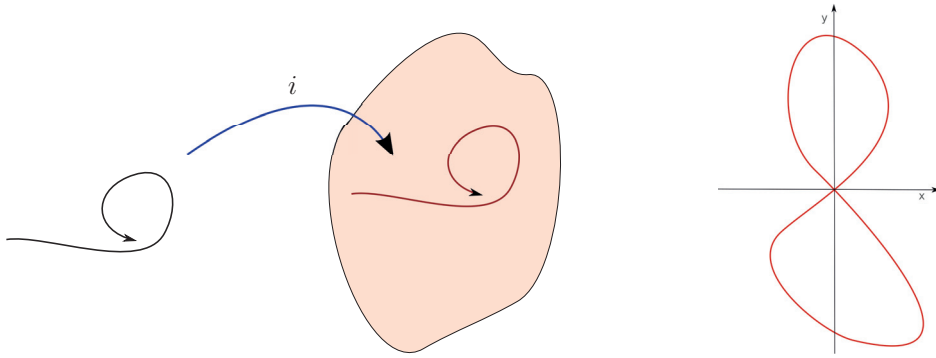
Figure 2.2 – Left: an example of an injective immersion that is not homeomorphic to a submanifold. The arrow means that the line approaches itself without touching. Right: the eight shape (red line) is not a manifold because it has a crossing point in the origin that is not locally homeomorphic to the Euclidean 1-space.

$\{\mathcal{U}, \phi\}$ *containing* x, *and a chart* $(\mathcal{V}, \psi) \in \{\mathcal{V}, \psi\}$ *containing* $F(x)$, *such that the map:*

$$\psi \circ F \circ \phi^{-1} : \phi(\mathcal{U}) \subset \mathcal{H}_1 \to \psi(\mathcal{V}) \subset \mathcal{H}_2$$

is differentiable in $\phi(x)$. *We say that* F *is differentiable if it is differentiable for any* $x \in \mathcal{M}_1$.

Observe that the map $\psi \circ F \circ \phi^{-1}$ goes from a subspace of $\mathcal{H}_1$ to $\mathcal{H}_2$ (Hilbert spaces), where the concepts of differentiability is defined in the standard way using the norms of $\mathcal{H}_1$ and $\mathcal{H}_2$. Definition 2.0.3 can be generalized to C^k differentiability for any $0 \le k \le p$. Consider the special case of real valued functions $f : \mathcal{M}_1 \to \mathbb{R}$. Following Definition 2.0.3, f is differentiable if for every $x \in \mathcal{M}_1$ there exists a smooth chart $(\mathcal{U}, \phi)$ such that $x \in \mathcal{U}$ and the composition $f \circ \phi^{-1} : \phi(\mathcal{U}) \to \mathbb{R}$ is smooth. We denote with $\mathcal{F}_x(\mathcal{M})$ the set of all smooth real valued functions defined in a neighborhood of $x \in \mathcal{M}$.

Definition 2.0.4. *A map* $F : \mathcal{M}_1 \to \mathcal{M}_2$ *is an **immersion** at* $x \in \mathcal{M}_1$ *if there exists an open neighborhood* $\mathcal{U} \subset \mathcal{M}_1$ *of* x *such that the restriction of* F *to* $\mathcal{U}$ *induces an homeomorphism*[*] *of* $\mathcal{U}$ *onto a submanifold of* $\mathcal{M}_2$. *We say that* F *is an immersion if it is an immersion for any* $x \in \mathcal{M}_1$.

If the immersion $F : \mathcal{M}_1 \to \mathcal{M}_2$ is injective, then $\mathcal{M}_1$ is homeomorphic to $F(\mathcal{M}_1) \subset \mathcal{M}_2$. Observe that an injective immersion is not necessarily homeomorphic to a submanifold. In other words an immersed subspace of $\mathcal{M}_2$ is not necessarily a submanifold of $\mathcal{M}_2$, see figure 2.2 (left) for an intuitive example. When this occurs we call it embedded submaniold, see subsection 2.0.3.

---

[*]an homeomorphism is a continuous bijective function between two topological spaces, whose inverse is continuous.

2.0.2 Tangent space

Let $\mathcal{M}$ be a C^p−manifold, with $p \geq 1$, modeled on $\mathcal{H}$ with atlas $\{\mathcal{U}, \phi\}$. We aim to equip $\mathcal{M}$ with a differential structure. To do so a tangent vectors ξ to $\mathcal{M}$ at x is represented by an equivalence class of triples $(\mathcal{U}, \phi, v)$ where $(\mathcal{U}, \phi)$ is a local chart of $\mathcal{M}$ such that $x \in \mathcal{U}$ and $v \in \mathcal{H}$ is the representative of ξ in this chart.

Let $(\mathcal{U}, \phi)$, $(\mathcal{V}, \psi)$ be two charts of $\mathcal{M}$ such that $x \in \mathcal{U} \cap \mathcal{V}$ and $v, w \in \mathcal{H}$, we say that $(\mathcal{U}, \phi, v)$ and $(\mathcal{V}, \psi, w)$ are equivalent representations of a tangent vector ξ to $\mathcal{M}$ at x if and only if $(\mathcal{V}, \psi, w) \iff w = D_{\phi(x)}(\psi \circ \phi^{-1})(v)$, where D is the usual directional derivative in $\mathcal{H}$. We say that v is the tangent vector ξ read in the local chart $(\mathcal{U}, \phi)$ whereas w is the tangent vector ξ read in the local chart $(\mathcal{V}, \psi)$. We denote with $\mathcal{T}_x\mathcal{M}$ the collection of all tangent vectors to $\mathcal{M}$ at x. Observe that the tangent space $\mathcal{T}_x\mathcal{M}$ can be equipped with a vector space structure induced by the model space. This is a remarkable property in view of the design of numerical approximation algorithms. Likewise the derivative to a real valued function provides a local linear approximation to the function, in same way we can think that the tangent space to a manifold gives a local vector space approximation of the manifold. Moreover, since the transition maps $\psi \circ \phi^{-1}$ are diffeomorphisms, the maps $D_{\phi(x)}(\psi \circ \phi^{-1})$ are isomorphisms from $\mathcal{H}$ to $\mathcal{H}$. Then, each chart $(\mathcal{U}, \phi)$ determines a bijection of $\mathcal{T}_x\mathcal{M}$ on $\mathcal{H}$, which allows to transport to the tangent space the Hilbert structure of the model space. It follows that for any $x \in \mathcal{M}$, $\mathcal{T}_x\mathcal{M}$ is a Hilbert space isomorphic to the model space $\mathcal{H}$ in which we can define a norm equivalent to the original norm of $\mathcal{H}$.

Example 2.0.3. *When $\mathcal{M}$ is a submanifold of a normed vector space $\mathcal{H}$, a tangent vector in x to $\mathcal{M}$ along a curve $\gamma : [0, T] \to \mathcal{M}$ can be simply identified with the classical derivative of γ at $t = 0$, i.e. $\xi = \dot{\gamma}(0) = \lim_{dt \to 0} \dfrac{\|\gamma(dt) - \gamma(0)\|_{\mathcal{H}}}{dt}$. In this case, the tangent space to $\mathcal{M}$ at x can be represented as the collection of the derivatives of all smooth curves passing by x:*

$$\mathcal{T}_x\mathcal{M} = \left\{\dot{\gamma}(0) \,|\, \gamma : \mathbb{R} \to \mathcal{M} \in C^1(\mathbb{R}, \mathcal{H}), \gamma(0) = x\right\}. \tag{2.1}$$

If now F is a differentiable map between two manifolds $\mathcal{M}_1$ and $\mathcal{M}_2$, by means of charts, we can interpret the derivative of F in a point $x \in \mathcal{M}_1$ as a map between the tangent space $\mathcal{T}_x\mathcal{M}_1$ and the tangent space $\mathcal{T}_{F(x)}\mathcal{M}_2$. More precisely we have the following definition:

Definition 2.0.5. *Let $\mathcal{M}_1, \mathcal{M}_2$ be two C^p−manifolds, with $p \geq 1$, modeled in $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, and equipped with atlas $\{\mathcal{U}, \phi\}$ and $\{\mathcal{V}, \psi\}$, and $F : \mathcal{M}_1 \to \mathcal{M}_2$ a differentiable function. We define **differential** (or **push-forward**) of F at x the linear map $D_xF : \mathcal{T}_x\mathcal{M}_1 \to \mathcal{T}_{F(x)}\mathcal{M}_2$ defined by:*

$$D_xF(\xi) = \mu \iff w = D_{\phi(x)}(\psi \circ F \circ \phi^{-1})(v) \tag{2.2}$$

where v is the tangent vector ξ read in the local chart $(\mathcal{U}, \phi)$ of $\mathcal{M}_1$ containing x and w is the tangent vector μ read in the local chart $(\mathcal{V}, \psi)$ of $\mathcal{M}_2$ containing $F(x)$. The differential of F may sometimes be denoted equivalently by F_*.

In other words, if a tangent vector $\xi \in \mathcal{T}_x \mathcal{M}_1$ is represented by $v \in \mathcal{H}_1$, then the tangent vector $D_x F(\xi) \in \mathcal{T}_{F(x)} \mathcal{M}_2$ is represented by $D_{\phi(x)}(\psi \circ F \circ \phi^{-1})(v) \in \mathcal{H}_2$. This can be represented by the following diagram:

$$
\begin{array}{ccc}
\mathcal{T}_x \mathcal{M}_1 & \xrightarrow{\ \sim\ } & \mathcal{H}_1 \\
{\scriptstyle D_x F}\downarrow & & \downarrow{\scriptstyle D_{\phi(x)}(\psi \circ F \circ \phi^{-1})} \\
\mathcal{T}_{F(x)} \mathcal{M}_2 & \xrightarrow{\ \sim\ } & \mathcal{H}_2
\end{array}
$$

Since the vector spaces $\mathcal{T}_x \mathcal{M}_1$ and $\mathcal{T}_{F(x)} \mathcal{M}_2$ are isomorphic to the model spaces, $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, the differential of F is a continuous map, for any F differentiable. Specifically, when $\mathcal{M}_1$ and $\mathcal{M}_2$ are submanifolds of normed (or metric) vector spaces and $\mathcal{T}_x \mathcal{M}_1$ is characterized as in (2.1), the differential to F at x can be interpreted as the map which associates to a tangent vector $v = \dot{\gamma}(0) \in \mathcal{T}_x \mathcal{M}_1$ the tangent vector to the curve $F \circ \gamma(t)$ at $t = 0$. Intuitively, one can think the push-forward map of F as the best local linear approximation of F. Based on the characterization of the differential, we can distinguish three categories of differentiable functions:

- F is a **submersion** at x if and only if $D_x F$ is surjective;

- F is a **immersion** at x if and only if $D_x F$ is injective and the image $D_x F(\mathcal{T}_x \mathcal{M}_1)$ is closed in $\mathcal{T}_{F(x)} \mathcal{M}_2$;

- F is a **diffeomorphism** at x if and only if $D_x F$ is bijective.

Observe that the definition of **immersion** given here coincides with Definition 2.0.4.

Definition 2.0.6. *Two manifolds $\mathcal{M}_1, \mathcal{M}_2$ are diffeomorphic, $\mathcal{M}_1 \simeq \mathcal{M}_2$, if there exists a diffeomorphism between them, i.e. a function $F : \mathcal{M}_1 \to \mathcal{M}_2$ such that the differential $D_x F$ is bijective at each $x \in \mathcal{M}_1$.*

Diffeomorphisms between manifolds will be used to characterize the tangent space of abstract manifolds in terms of tangent space of diffeomorphic manifolds with well know differential structures.

Definition 2.0.7. *The union of the tangent spaces at all elements of $\mathcal{M}$ is called **tangent bundle**:*

$$
\mathcal{T} \mathcal{M} := \bigcup_{x \in \mathcal{M}} \mathcal{T}_x \mathcal{M}
$$

Definition 2.0.8. *Let $\mathcal{M}$ be a C^p-manifold, with $p \geq 1$. We define C^k-**vector field**, for any $0 \leq k \leq p$, a map*

$$
\begin{aligned}
\Xi : \mathcal{M} &\to \mathcal{T} \mathcal{M} \\
x &\mapsto \Xi(x)
\end{aligned}
$$

*which assigns to any $x \in \mathcal{M}$ a tangent vector $\Xi(x) \in \mathcal{T}_x\mathcal{M}$. The **integral curve** of a vector field on the manifold $\mathcal{M}$ is a curve $c : [0, T] \to \mathcal{M}$ such that:*

$$\frac{d}{dt}c(t) = \Xi(c(t)), \quad \forall t \in [0, T]. \tag{2.3}$$

Namely $\Xi(c(t))$ is tangent to the curve at $c(t)$ at any time.

The time derivative in (2.3) is defined following Definition 2.0.5 as $\frac{d}{dt}(\psi \circ c)(t)$ read in the local chart $(\mathcal{V}, \psi)$ where $\mathcal{V}$ contains $c(t)$.

Definition 2.0.9. *Let g_x be a bilinear, symmetric and positive-definite form:*

$$\begin{aligned} g_x \quad &: \mathcal{T}_x\mathcal{M} \times \mathcal{T}_x\mathcal{M} \quad \to \mathbb{R} \\ &\quad (\xi, \nu) \qquad\qquad \mapsto g_x(\xi, \nu). \end{aligned}$$

*If this form varies smoothly over the tangent bundle, then g defines a **Riemannian metric**. A pair $(\mathcal{M}, g)$ consisting of a manifold $\mathcal{M}$ and a Riemannian metric g is called a **Riemannian manifold**.*

2.0.3 Embedded submanifold

We have seen that if a set can be equiped with a manifold structure, usually admits more than one manifold structure. However, a subset $\mathcal{N} \subset \mathcal{M}$, where $\mathcal{M}$ is a manifold, admits (at most) only one structure which makes it a submanifold of $\mathcal{M}$. Namely there exists only one maximal atlas which induces on $\mathcal{N}$ and on the tangent space to $\mathcal{N}$, the same topology induced by the ambient manifold. A special type of submanifolds are the, so called, **embedded submanifolds**, which are the image of injective immersions whose topology coincides with the (unique) subspace topology induced by $\mathcal{M}$.

Theorem 2.0.1. *[74] Let $F : \mathcal{N} \to \mathcal{M}$ be an injective immersion between the manifolds $\mathcal{N}$ and $\mathcal{M}$. If F is homeomorphic onto its image, then $F(\mathcal{N})$ is an (embedded) submanifold of $\mathcal{M}$ and $\mathcal{N}$ is isomorphic to $F(\mathcal{N})$.*

The following theorem is useful to recognize or construct embedded manifolds.

Theorem 2.0.2. *[76] Let $F : \mathcal{M}_1 \to \mathcal{M}_2$ be a differentiable function between two smooth manifolds $\mathcal{M}_1, \mathcal{M}_2$ and y be an element in the range of F. If the differential of F at x is surjective for all $x \in F^{-1}(y)$, then $F^{-1}(y)$ is an embedded manifold of $\mathcal{M}_1$.*

Stiefel manifold

Let $\mathcal{H}$ be a (possibly infinite dimensional) Hilbert space equipped with an inner product $<\cdot, \cdot>$ and let $S > 0$ be a natural number. We call **Stiefel manifold**, denoted with $St(S, \mathcal{H})$, the

collection of all orthonormal frames of S elements in $\mathcal{H}$:

$$St(S, \mathcal{H}) = \{\mathbf{V} = (V_1, ..., V_S) : V_i \in \mathcal{H} \text{ and } <V_i, V_j> = \delta_{ij} \; \forall i, j = 1, ..., S\}. \tag{2.4}$$

$St(S, \mathcal{H})$ is a smooth embedded submanifold of $[\mathcal{H}]^S$ [57]. This can be shown by applying Theorem 2.0.2 to the following differentiable map:

$$\begin{aligned} F : [\mathcal{H}]^S &\to \mathscr{S}_{sym}(S) \\ \mathbf{V} &\mapsto \langle\langle \mathbf{V}, \mathbf{V} \rangle\rangle - \mathbb{I}_S \end{aligned}$$

where $\mathbf{V}$ is a row vector, $\mathbb{I}_S$ is the identity in $\mathbb{R}^{S \times S}$, $\langle\langle \mathbf{V}, \mathbf{V} \rangle\rangle \in \mathbb{R}^{S \times S}$ is defined as $\langle\langle \mathbf{V}, \mathbf{U} \rangle\rangle_{ij} := \langle\langle V_j, V_i \rangle\rangle$ for all $i, j = 1, ..., S$ and $\mathscr{S}_{sym}(S) := \{\mathbf{B} \in \mathbb{R}^{S \times S} : \mathbf{B}^T = \mathbf{B}\}$. It is evident that $St(S, \mathcal{H}) = F^{-1}(0)$. We need to verify that the differential of F is surjective for all $\mathbf{V} \in St(S, \mathcal{H})$. We recall that the tangent space to $[\mathcal{H}]^S$ and $\mathscr{S}_{sym}(S)$ are respectively given by $\mathcal{T}_{\mathbf{U}}[\mathcal{H}]^S = [\mathcal{H}]^S$, $\mathcal{T}_{\mathbf{B}}\mathscr{S}_{sym}(S) = \mathscr{S}_{sym}(S)$, for any $\mathbf{U} \in [\mathcal{H}]^S$ and $\mathbf{B} \in \mathscr{S}_{sym}(S)$. In light of this, we do not need to read the differential of F by means of charts and we can write directly:

$$D_{\mathbf{V}}F(\mathbf{Z}) = \langle\langle \mathbf{V}, \mathbf{Z} \rangle\rangle + \langle\langle \mathbf{Z}, \mathbf{V} \rangle\rangle, \forall \mathbf{Z}, \mathbf{V} \in [\mathcal{H}]^S$$

This is surjective at $\mathbf{V} \in St(S, \mathcal{H})$ because $D_{\mathbf{V}}F(\frac{1}{2}\mathbf{V}\mathbf{B}) = \mathbf{B}$ for all $\mathbf{B} \in \mathscr{S}_{sym}(S)$. In particular $St(S, \mathcal{H})$ is a complete Riemannian manifold with the induced Riemannian metric given by $<\mathbf{V}, \mathbf{U}> = \sum_{i=1}^{S} <V_i, U_i>$.

The fact that $St(S, \mathcal{H})$ is an embedded manifold of $[\mathcal{H}]^S$ relieves us from using charts for reading the tangent vectors to $St(S, \mathcal{H})$. More precisely $St(S, \mathcal{H})$ admits a global chart $(St(S, \mathcal{H}), \phi)$ where ϕ is actually the inclusion $\phi(\mathbf{V}) = \mathbf{V}$, which implies that $\xi \in \mathcal{T}_{\mathbf{V}}St(S, \mathcal{H})$ is actually an element of $[\mathcal{H}]^S$ for any $\mathbf{V} \in St(S, \mathcal{H})$. This facilitates us to characterize the elements of $\mathcal{T}_{\mathbf{V}}St(S, \mathcal{H})$. Consider a smooth curve $\mathbf{V}(t)$ in $St(S, \mathcal{H})$; we have that:

$$\langle\langle \mathbf{V}(t), \mathbf{V}(t) \rangle\rangle = \mathbb{I}_S, \qquad \forall t. \tag{2.5}$$

If we differentiate relation (2.5) with respect to t we get:

$$\langle\langle \dot{\mathbf{V}}(t), \mathbf{V}(t) \rangle\rangle + \langle\langle \mathbf{V}(t), \dot{\mathbf{V}}(t) \rangle\rangle = \mathbf{0}, \qquad \forall t. \tag{2.6}$$

where the time derivative $\dot{\mathbf{V}}$ is thought in the ambient space, thanks to the fact that $St(S, \mathcal{H})$ is an embedded manifold of $[\mathcal{H}]^S$. Then observe that, since $\mathcal{T}_{\mathbf{V}}St(S, \mathcal{H}) = [\mathcal{H}]^S$, any tangent vector $\dot{\mathbf{V}}(t) \in \mathcal{T}_{\mathbf{V}}St(S, \mathcal{H})$ belongs actually to $[\mathcal{H}]^S$ and can be decomposed in:

$$\dot{\mathbf{V}}(t) = \mathbf{V}(t)\mathbf{\Omega} + \mathbf{Z}$$

where $\mathbf{\Omega} \in \mathbb{R}^{S \times S}$ and $\mathbf{Z} = (Z_1, ..., Z_S)$ is in the orthogonal complement to $\mathbf{V}(t)$ in $[\mathcal{H}]^S$, i.e. $<Z_i, V_j(t)> = 0 \; \forall i, j = 1, ..., S$. Then, for relation (2.6) to be satisfied, we get that $\mathbf{\Omega}$ is necessarily skew-symmetric, i.e.: $\mathbf{\Omega} = -\mathbf{\Omega}^T$. Finally we have that the tangent space to $St(S, \mathcal{H})$ at $\mathbf{V}$ can be

written as:

$$
\begin{aligned}
\mathcal{T}_\mathbf{V} St(S, \mathcal{H}) \quad &= \{\delta\mathbf{V} \in [\mathcal{H}]^S \text{ such that } <\delta V_i, V_j> + <V_i, \delta V_j> = 0, \forall i, j = 1, ..., S\} \\
&= \{\delta\mathbf{V} = \mathbf{V}\mathbf{\Omega} + \mathbf{Z} \in [\mathcal{H}]^S \,|\, \mathbf{\Omega} = -\mathbf{\Omega}^T \in \mathbb{R}^{S\times S}, \mathbf{Z} \in [\mathcal{H}]^S : <Z_i, V_j> = 0 \, \forall i, j = 1, ..., S\}
\end{aligned}
$$

In this thesis we consider both:

- $St(S, L^2(\Omega))$ defined as the collection of all vectors of S $L^2(\Omega)-$orthonormal random variables in a probability space $(\Omega, \mathcal{A}, \mathcal{P})$;

- $St(S, H^1(D))$ which denotes the collection of all $L^2(D)-$orthonormal frames of S vector functions in $H^1(D)$; with $D \subset \mathbb{R}^d$ an open bounded domain.

$$
St(S, H^1(D)) = \left\{\mathbf{V} = (V_1, ..., V_S) : V_i \in H^1(D) \text{ and } <V_i, V_j>_{L^2(D)} = \delta_{ij} \, \forall i, j = 1, ..., S\right\}
\tag{2.7}
$$

We emphasize that $St(S, H^1(D))$ is embedded in $[H^1(D)]^S$ while the orthonormality condition is required in the weaker norm $L^2(D)$. However this distinction does not affect the construction of the differential structure of $St(S, H^1(D))$ and we can proceed as previously described (where in (2.6) we consider the $L^2(D)$-inner product). In particular we have that for any $\mathbf{V} \in St(S, H^1(D))$ the tangent space is isomorphic to $[H^1(D)]^S$ and can be written as:

$$
\begin{aligned}
\mathcal{T}_\mathbf{V} St(S, H^1(D)) = \{\delta\mathbf{V} = \mathbf{V}\mathbf{\Omega} + \mathbf{Z} \in [H^1(D)]^S \quad &|\, \mathbf{\Omega} = -\mathbf{\Omega}^T \in \mathbb{R}^{S\times S}, \mathbf{Z} \in [H^1(D)]^S : \\
&<Z_i, V_j>_{L^2(D)} = 0 \, \forall i, j = 1, ..., S\}
\end{aligned}
\tag{2.8}
$$

where for convenience $\delta\mathbf{V} \in [H^1(D)]^S$ is decomposed in the part belonging to the sub-space spanned by $\mathbf{V}$ and the part in the orthogonal complement to $\mathbf{V}$ in $[H^1(D)]^S$ with respect to the $L^2(D)$-norm. In the following, we always denote by $St(S, H^1(D))$ the set in (2.7). This observation can be generalized to any Sobolev space $H^p(D)$ for any $p > 1$.

2.0.4 Quotient spaces of manifolds

Intuitively quotient spaces can be imagined as the result of an equivalence classing in which all equivalent elements are "contracted" in only one point. When equipped with a suitable manifold structure they provide the proper mathematical framework for several numerical/-computational applications. Think, for instance, to numerical algorithms involving finite dimensional subspaces which, for computational applications, strictly need to be represented in terms of bases (matrices at the discrete level). It is clear that there are infinitely many "equivalent" bases spanning the same subspace, but only one have to be chosen. The questions are how to properly choose only one "representative element" per class and how to make this choice depend smoothly (in a suitable sense) on the element that has to be represented. A possible solution can be obtained by equipping the quotient space with a differential manifold structure. We start by recalling under which conditions this can be achieved. Let $\mathcal{M}$ be a man-

ifold and ~ an equivalence relation defined for all elements of $\mathcal{M}$. We call **fiber** (or *equivalent class*) containing $x \in \mathcal{M}$ the set of all elements which are equivalent to x:

$$[x] := \{y \in \mathcal{M} \text{ such that } y \sim x\}$$

If we regard all equivalent elements as a unique element, we get what is called **quotient space** of $\mathcal{M}$ by $\sim$, namely the set of all equivalence classes of $\sim$ in $\mathcal{M}$, i.e.:

$$\mathcal{M} / \sim := \big\{[x] : x \in \mathcal{M}\big\}$$

Any element of the quotient space is a fiber, and so it corresponds to a subset of $\mathcal{M}$. The map which associates to any element its fiber, is called **canonical map** and is defined as

$$\begin{aligned} \pi : \mathcal{M} &\to \mathcal{M} / \sim \\ x &\mapsto [x] \end{aligned}$$

It is evident that $\pi(x) = \pi(y)$ if and only if $x \sim y$. Any fiber $[x]$ can be seen as the inverse of the canonical map in the point $y = \pi(x)$, namely $[x] = \pi^{-1}(y)$.

Proposition 2.0.1. *[2, 16] Let $\mathcal{M} / \sim$ be a quotient space equipped with the structure of quotient manifold of $\mathcal{M}$, and let π denote the canonical projection map. Each equivalent class $[x] = \pi^{-1}(\pi(x))$, is an embedded manifold of $\mathcal{M}$, $\forall x \in \mathcal{M}$.*

In general quotients of smooth manifolds are not necessarily smooth manifolds themselves. Nice quotient structures can be derived by the action of Lie groups, under some additional conditions which will be summarized in the next section.

Quotient manifold by Lie group action

Definition 2.0.10. *A **Hilbert Lie group** G modeled on a Hilbert space $\mathcal{G}$ is a C^∞ manifold modeled on $\mathcal{G}$, with group operation, given by the multiplication:*

$$\begin{aligned} G \times G &\to G \\ (g, h) &\mapsto gh \end{aligned}$$

and inverse map:

$$\begin{aligned} G &\to G \\ g &\mapsto g^{-1}. \end{aligned}$$

which are both smooth.

We denote with e the neutral element of G, i.e. the only element of G such that $ge = g$ for all $g \in G$. A **right-action** θ of Lie group G on a smooth manifold $\mathcal{M}$ is a map:

$$\begin{aligned} \theta : \mathcal{M} \times G &\to \mathcal{M} \\ (x, g) &\mapsto \theta(x, g) =: R_g(x) \end{aligned}$$

which satisfies:

- $\theta(x, e) = x$ for all $x \in \mathcal{M}$,

- $\theta(x, gh) = \theta(\theta(x, h), g)$, for all $x \in \mathcal{M}$ and $g, h \in G$. (Equivalently written as $R_{gh} = R_g \circ R_h$)

The action is smooth if the map θ is smooth. For convenience of notation, right actions will be often denoted as $\theta(x, g) = x \cdot g$ in what follows. We recall that the G-**orbit** of a point $x \in \mathcal{M}$ is the set of elements in $\mathcal{M}$ to which x can be moved to, by the Lie action of G. This concept can be used to define an equivalence relation in $\mathcal{M}$: we say that two elements $x, y \in \mathcal{M}$ are equivalent if they can be moved one towards the other by the action of an element of G. More precisely:

$$x \sim y \iff \exists g \in G : \theta(x, g) = y$$

According to this definition, the G-orbits correspond to the fibers of $\mathcal{M}$ by the equivalence relation $\sim$ induced by the G action. The quotient space $\mathcal{M}/G := \mathcal{M}/\sim$ is the collection of all orbits. We recall that:

- an action θ of G on a smooth manifold $\mathcal{M}$ is **free** if has not fixed points. This means that any element $g \in G$ different from the neutral element moves any point $x \in \mathcal{M}$, i.e.:

$$\text{if } \exists x \in \mathcal{M} : \theta(x, g) = x \Rightarrow g = e$$

- an action θ of G on a smooth manifold $\mathcal{M}$ is **proper** if the graph of θ:

$$\begin{aligned} \Theta : \quad G \times \mathcal{M} &\to \mathcal{M} \times \mathcal{M} \\ (g, x) &\mapsto (\theta(x, g), x) \end{aligned}$$

  is proper, i.e. preimages of compact sets have compact closure. Roughly speaking G acts properly if each compact subset is moved away from itself by most elements of the group.

- an action θ is **isometric** if it leaves the metric, given by the inner product, invariant on the fibers:

$$\langle D_x R_g[v], D_x R_g[\xi] \rangle = \langle v, \xi \rangle, \quad \forall v, \xi \in \mathcal{T}_x\mathcal{M}, \forall x \in \mathcal{M}$$

Let us define the following map, ϑ_x, that sends an element of G to an element of the orbit containing x, here denoted by F_x:

$$\begin{aligned} \vartheta_x : \quad G &\to F_x \\ g &\mapsto \theta(g, x). \end{aligned}$$

The action θ is free if and only if ϑ_x is injective for all $x \in \mathcal{M}$.

Theorem 2.0.3. *[Quotient Manifold theorem, [76, 74]] Let G be a Hilbert Lie group acting on the right on a Riemannian Hilbert manifold $\mathcal{M}$ via an isometric action θ which is smooth, free and proper. Provided that the tangent map $D_e\vartheta_x$ has a closed range, then:*

- *the orbits are closed submanifolds of $\mathcal{M}$ and ϑ_x is a diffeomorphism,*

- *the quotient space $\mathcal{M}/G$ has a smooth Hilbert structure,*

Grassmann manifold as quotient space

Let $\mathcal{H}$ be a general Hilbert space and S a positive real number, we denote by $\mathcal{G}(S,\mathcal{H})$ the Grassmann manifold of dimension S that consists of all the $S-$dimensional linear subspaces of $\mathcal{H}$. The Grassmann $\mathcal{G}(S,\mathcal{H})$ can be identified with the Stiefel manifold $St(S,\mathcal{H})$, defined in (2.4), quotiented by the equivalence relation ~ defined as follows:

$$\mathbf{V} \sim \mathbf{U} \iff \text{span}(\mathbf{V}) = \text{span}(\mathbf{U}) \tag{2.9}$$

It is straightforward to verify that ~ is an equivalence relation as it is reflexive, symmetric and transitive. Directly from the definition of ~ we easily see that there is a one-to-one correspondence between $St(S,\mathcal{H})/\sim$ and $\mathcal{G}(S,\mathcal{H})$, since two elements $\mathbf{V},\mathbf{U} \in St(S,\mathcal{H})$ are equivalent if and only if they span the same subspace . This implies that, if $St(S,\mathcal{H})/\sim$ admits a manifold structure, this can be naturally reflected to $\mathcal{G}(S,\mathcal{H})$.

Observe that two elements $\mathbf{V},\mathbf{U} \in St(S,\mathcal{H})$ are equivalent if and only if $\mathbf{U} = \mathbf{VO}$ for some orthogonal matrix $\mathbf{O}$. Let us denote by $\mathcal{O}_S$ the subspace of all the orthogonal matrices of dimension S: $\mathcal{O}_S = \{O \in \mathbb{R}^{S\times S} : \mathbf{O}^T\mathbf{O} = \mathbf{OO}^T = \mathbb{I}\}$. It is straightforward to verify that $\mathcal{O}_S$ is a (compact) Lie group modelled on $\mathbb{R}^{S\times S}$ where the group operation is given by matrix multiplication. Let us define the following smooth right action of $\mathcal{O}_S$ in $St(S,\mathcal{H})$:

$$\theta : St(S,\mathcal{H}) \times \mathcal{O}_S \rightarrow St(S,\mathcal{H})$$
$$(\mathbf{U},\mathbf{O}) \mapsto \mathbf{V} = \mathbf{UO} =: R_{\mathbf{O}}(\mathbf{U}) \iff V_k = \sum_{i=1}^{S} U_i O_{ik} \quad \forall k = 1,...,S$$

Since the equivalence relation induced by the action θ coincides with ~ in (2.9), $St(S,\mathcal{H})/\sim$ coincides with $St(S,\mathcal{H})/\mathcal{O}_S$. For any $\mathbf{U} \in St(S,\mathcal{H})$, the equivalent class containing $\mathbf{U}$, which is given by $\mathbf{U}\mathcal{O}_S = \{\mathbf{UW} : \mathbf{W} \in \mathcal{O}_S\}$, is identified by the subspace $\mathcal{U} = \text{span}(U_1,...,U_S) \in \mathcal{G}(S,\mathcal{H})$. One can verify that θ is an isometric action, indeed:

$$< D_{\mathbf{Z}}R_{\mathbf{O}}[\mathbf{U}], D_{\mathbf{Z}}R_{\mathbf{O}}[\mathbf{W}] > = < \mathbf{UO},\mathbf{WO} > = < \mathbf{U},\mathbf{W} > \qquad \forall \mathbf{U},\mathbf{W} \in \mathcal{T}_{\mathbf{Z}}St(S,\mathcal{H}), \forall \mathbf{Z} \in St(S,\mathcal{H})$$

which acts freely and properly in $St(S,\mathcal{H})$. Hence Theorem 2.0.3 implies that $\mathcal{G}(S,\mathcal{H}) = St(S,\mathcal{H})/\mathcal{O}_S$ admits a (unique) structure of (Hilbert) quotient manifold.

2.0.5 Principal fiber bundle

The structure of principal fibre-bundle is the appropriate mathematical formulation underlying the *gauge theory*, first developed in physics to describe the dynamics of all non-gravitational interactions, and used afterwards in several other applications.

Definition 2.0.11. *Let $(\mathcal{M}, \mathcal{B}, \pi, G)$ be a geometrical structure, where:*

- *$\mathcal{M}$ is a manifold, called total space;*

- *G is a Lie group which acts freely on $\mathcal{M}$ on the right:*

$$
\begin{aligned}
\mathcal{M} \times G &\to \mathcal{M} \\
\theta(x, g) &\mapsto x \cdot g,
\end{aligned}
$$

- *$\mathcal{B}$, called base, is the quotient manifold of $\mathcal{M}$ by the action of G,*

- *$\pi : \mathcal{M} \to \mathcal{B}$ is the projection map, i.e. the surjective continuous map that associates to each point in $\mathcal{M}$ the G-orbit containing x.*

*We say that $(\mathcal{M}, \mathcal{B}, \pi, G)$ forms a **principle fiber bundle** if it satisfies the condition of local triviality, i.e. if there exists a family of charts $\{U_i, \phi_i\}$ where $\{U_i\}$ are open subsets covering $\mathcal{B}$ and $\phi_i : \pi^{-1}(U_i) \to U_i \times G$ are homeomorphic maps, such that the following diagram commutes:*

$$
\begin{array}{ccc}
\pi^{-1}(U_i) & \xrightarrow{\varphi} & U_i \times G \\
{\scriptstyle \pi} \downarrow & \swarrow {\scriptstyle proj_1} & \\
U_i & &
\end{array}
$$

where $proj_1$ is the projection into the first component. Namely there exists a G-equivariant[†] map $g_i : \pi^{-1}(U_i) \to G$ which is a fibrewise diffeomorphism and such that $\phi_i(p) = (\pi(p), g_i(p))$. The collection $\{U_i, \phi_i\}$ is called local trivialization of the bundle.

Definition 2.0.12. *We call **local section** any right inverse of π in $U_i \subset \mathcal{B}$, namely any smooth map $\sigma : U_i \to \pi^{-1}(U_i)$ such that $\pi \circ \sigma = id$, i.e. $\sigma(x) = y$ if and only if $y \in \pi^{-1}(x)$.*

Roughly speaking, section maps are functions that assign to each point (equivalent class) $x \in U_i$ a unique "representative" point y on the fiber corresponding to x. Observe that the local triviality condition guarantees the existence of local sections.

Lemma 2.0.1. *[122] Let $\pi : \mathcal{M} \to \mathcal{B}$ be a surjective submersion and let G be a Lie group which acts freely on $\mathcal{M}$ from the right such that the orbits of the action are exactly the fibers $\pi^{-1}(y)$ for any $y \in \mathcal{B}$. Then $(\mathcal{M}, \mathcal{B}, \pi, G)$ is a principal G-bundle.*

[†]A map ϕ defined on $\pi^{-1}(U_i)$ is G-equivariant if $\phi(pg) = \phi(p)g$ for all $g \in G$ and $p \in \pi^{-1}(U_i)$.

In other words, assumed that G acts freely on $\mathcal{M}$, a necessary and sufficient condition for $(\mathcal{M}, \mathcal{M}/G, \pi, G)$ to form a principle fiber bundle is that the mapping π admit local sections. Combining Lemma 2.0.1 with Proposition 2.0.1 we have that any $(\mathcal{M}, \mathcal{M}/G, \pi, G)$ where $\mathcal{M}/G$ is a quotient manifold, forms a principal fiber bundle.

The fiber bundle associate to the Grassmann manifold

In Section 2.0.4 we have seen that $\mathcal{G}(S, \mathcal{H}) = St(S, \mathcal{H})/\mathcal{O}_S$ admits a structure of quotient manifold, which implies, by means of Proposition 2.0.1, that the canonical projection map π : $St(S, \mathcal{H}) \to \mathcal{G}(S, \mathcal{H})$ is a submersion and the fibers: $\mathbf{U}\mathcal{O}_S = \{\mathbf{U}\mathbf{W} : \mathbf{W} \in \mathcal{O}_S\}$ are embedded submanifolds of $St(S, \mathcal{H})$. As a consequence of Lemma 2.0.1, we conclude that $(St(S, \mathcal{H}), \mathcal{G}(S, \mathcal{H}), \pi, \mathcal{O}_S)$ form a principle fiber bundle.

2.0.6   Tangent space to the base manifold

Consider a principle fiber bundle $(\mathcal{M}, \mathcal{B}, \pi, G)$; the differential structure of the total space $\mathcal{M}$ is typically well known and used to characterize the tangent vectors to the base space $\mathcal{B}$. Let $\xi \in \mathcal{T}_y\mathcal{B}$ be a tangent vector to $\mathcal{B}$ at y and x be an element in the preimage of y, i.e. $x \in \pi^{-1}(y) \subset \mathcal{M}$; any tangent vector $v \in \mathcal{T}_x\mathcal{M}$ which satisfies $D_x\pi(v) = \xi$ can be considered a suitable representation of ξ. The drawback of this approach is that $D_x\pi$ is not injective and so the representation of ξ is not uniquely determined. The uniqueness is recovered by decomposing $\mathcal{T}_x\mathcal{M}$ into the subset of directions tangent to the fiber $\pi^{-1}(y)$ and its complementary space. The latter, which consists in all directions that do not induce displacements along the fiber, provides a suitable representation of the tangent space to the base manifold.

Definition 2.0.13. *The **vertical space** at $x \in \mathcal{M}$, denoted by V_x, is the vector subspace of $\mathcal{T}_x\mathcal{M}$ consisting of all the tangent vectors which are tangent to fiber of x.*

We call **vertical distribution** the map that assigns to each element $x \in \mathcal{M}$ the vertical space $V_x \subset \mathcal{T}_x\mathcal{M}$. Any vertical distribution is G invariant. We remind that for any x, the fiber containing x is an embedded submanifold of $\mathcal{M}$, which coincides with the image of the inverse of the canonical map in the point $y = \pi(x)$. Since π is constant along the fibers, the differential of π at x, $D_x\pi : \mathcal{T}_x\mathcal{M} \to \mathcal{T}_{\pi(x)}\mathcal{B}$, vanishes along the fiber. So the vertical space V_x can be defined as the kernel of $D_x\pi$, i.e.

$$V_p = \{v \in \mathcal{T}_x\mathcal{M} : D_x\pi(v) = 0\}.$$

Intuitively this means that movements in the vertical direction make no changes in the quotient space. This motivates the choice of representing tangent vectors in the tangent space to $\mathcal{B}$ at $\pi(x)$ by means of tangent vectors in the complement of the vertical space V_x.

Definition 2.0.14. *A **connection** on a principle fiber bundle $(\mathcal{M}, \mathcal{B}, \pi, G)$ is a smooth splitting:*

$$\mathcal{T}_x\mathcal{M} = V_x \oplus H_x, \qquad \forall x \in \mathcal{M}$$
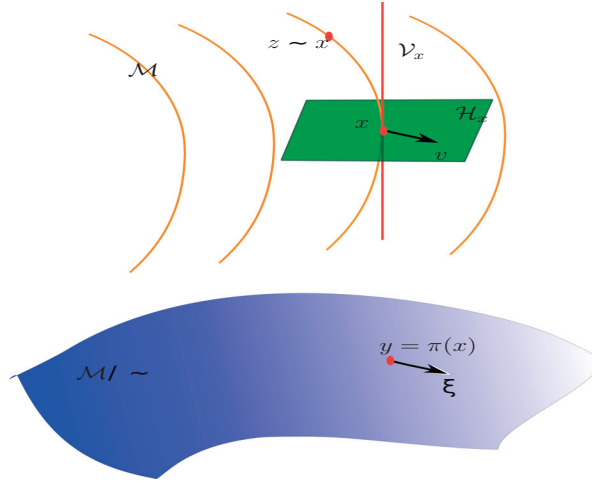
Figure 2.3 – The tangent vector $\xi \in \mathcal{T}_y(\mathcal{M}/\sim)$ to the quotient manifold $\mathcal{M}/\sim$ is represented by the tangent vector v to the total manifold $\mathcal{M}$. The vector v belongs to the horizontal space $\mathcal{H}_x \subset \mathcal{T}_x\mathcal{M}$ at $x \in \mathcal{M}$ such that $y = \pi(x)$.

*satisfying the G-invariance property: $H_{xg} = D_x R_g[H_x]$, $\forall g \in G$, $\forall x \in \mathcal{M}$. H_x is called **horizontal space** to $\mathcal{M}$ at x.*

The map that assigns to each element $x \in \mathcal{M}$ the horizontal space $H_x \subset \mathcal{T}_x\mathcal{M}$ is named **horizontal distribution** and is G invariant. The existence (but not uniqueness) of a horizontal distribution is always guaranteed in our context of Hilbert manifolds; the same conclusion does not generally apply to manifold modelled on Banach spaces. Observe that a horizontal distribution turns the map $D_x\pi : H_x \to \mathcal{T}_{\pi(x)}\mathcal{B}$ into an ismorphism for any $x \in \mathcal{M}$.

Definition 2.0.15. *Let $\Xi : \mathcal{B} \to \mathcal{T}\mathcal{B}$ be a smooth vector field on $\mathcal{B}$. We call **horizontal lift** of Ξ the vector field $\Psi : \mathcal{M} \to H$ for which $D_x\pi(\Psi(x)) = \Xi(\pi(x))$ for all $x \in \mathcal{M}$.*

We emphasize that in the absence of any extra structure there is no natural complement to V_x in $\mathcal{T}_x\mathcal{M}$.

Remark 2.0.2. *If $\mathcal{M}$ is a Riemannian manifold equipped with a G-invariant metric, we can simply define the horizontal space as the orthogonal complement (with respect to the metric) of V_x in $\mathcal{T}_x\mathcal{M}$, i.e. $H_x = V_x^\perp$.*

The choice of the horizontal space can be equivalently seen as the definition of a linear map $D_{\pi(x)}\sigma : \mathcal{T}_{\pi(x)}\mathcal{B} \to \mathcal{T}_x\mathcal{M}$ so that:

- $D_x\pi \circ D_{\pi(x)}\sigma = i_d$ the identity in $\mathcal{T}_{\pi(x)}\mathcal{B}$.

- $D_{\pi(x)}\sigma$ is G-invariant.

As the symbol suggests, the map $D_{\pi(x)}\sigma$ can be interpreted as the differential of a local section σ at $\pi(x)$.

Definition 2.0.16. *Let $\Xi : \mathcal{B} \to \mathcal{T}\mathcal{B}$ be a smooth vector field and $\gamma : [0, T] \to \mathcal{B}$ the integral curve of Ξ passing by y at $t = 0$. We define **lift** of γ through $x \in \pi^{-1}(y)$ a curve $\tilde{\gamma} : [0, T] \to \mathcal{M}$ such that $\tilde{\gamma}(0) = x$ and $\pi(\tilde{\gamma}(t)) = \gamma(t)$ for any $t \in [0, T]$. A lift of γ is **horizontal** if is the integral curve of the horizontal lift of Ξ.*

A rigorous treatment of these concepts is given for instance in [25, 63]. In particular the following classical result of fiber bundle theory tells us that horizontal distributions provide unique parametrization of the tangent space to a base manifold.

Theorem 2.0.4. *[25] Let $(\mathcal{M}, \mathcal{B}, \pi, G)$ be a principal fiber bundle and $\mathcal{H}_x : \mathcal{M} \to H$, $x \mapsto H_x$ a horizontal distribution. For any $y \in \mathcal{B}$ and any smooth curve $\gamma : \mathbb{R} \to \mathcal{B}$ such that $\gamma(0) = y$, there exists a unique horizontal lift $\tilde{\delta} : \mathbb{R} \to \mathcal{M}$ with $\tilde{\delta}(0) = x$, for any $x \in \mathcal{M}$ such that $\pi(x) = y$.*

In other words a horizontal distribution turns the map $D_x\pi : H_x \to \mathcal{T}_x\mathcal{B}$ into an isomorphism. This leads to a unique representation of the tangent space to $\mathcal{B}$ at $\gamma(t)$ and hence a local parametrization of the base manifold by means of pull back.

Tangent space to the Grassmann manifold

We continue in the analysis of the principle fiber bundle $(St(S, \mathcal{H}), \mathcal{G}(S, \mathcal{H}), \pi, \mathcal{O}_S)$. To equip $\mathcal{G}(S, \mathcal{H})$ with a differential structure we need to fix a horizontal distribution in $St(S, \mathcal{H})$. We have seen that for any $\mathbf{V} \in St(S, \mathcal{H})$ the tangent space to $St(S, \mathcal{H})$ at $\mathbf{V}$ is given by:

$$\mathcal{T}_{\mathbf{V}}St(S, \mathcal{H}) = \{\delta\mathbf{V} = \mathbf{V}\mathbf{\Omega} + \mathbf{Z} \in [\mathcal{H}]^S \,|\, \mathbf{\Omega} = -\mathbf{\Omega}^T \in \mathbb{R}^{S \times S}, \mathbf{Z} \in [\mathcal{H}]^S : <Z_i, V_j> = 0 \; \forall i, j = 1, ..., S\}$$

Any fiber $\mathbf{U}\mathcal{O}_S$ is an embedded submanifold of $St(S, \mathcal{H})$ thus it inherits the differential structure of the ambient space $St(S, \mathcal{H})$. Take an arbitrary curve along the fiber $\mathbf{U}\mathcal{O}_S$, this is written as $\alpha : \mathbb{R} \to \mathbf{U}\mathcal{O}_S$, $t \mapsto \alpha(t) = \mathbf{U}\mathbf{O}(t)$, which implies that the tangent vector to α at time t has to be written as $\mathbf{U}\dot{\mathbf{O}}(t)$. Since $\mathbf{U}\dot{\mathbf{O}}(t) \in \mathcal{T}_{\mathbf{V}(t)}St(S, \mathcal{H})$, it is evident that $\dot{\mathbf{O}}(t) \in \mathbb{R}^{S \times S}$ is skew-symmetric. In conclusion, for any $\mathbf{U} \in St(S, \mathcal{H})$ the vertical space at $\mathbf{U}$ is given by:

$$V_{\mathbf{U}} = \{\mathbf{U}\mathbf{\Omega} : \mathbf{\Omega} = -\mathbf{\Omega}^T \in \mathbb{R}^{S \times S}\}.$$

Depending on the choice of the horizontal distribution we can have different parametrizations of the tangent bundle of $\mathcal{G}(S, \mathcal{H})$. However, since $St(S, \mathcal{H})$ is a Riemannian manifold with metric defined as $<\mathbf{V}, \mathbf{U}> = \sum_{i=1}^{S} <V_i, U_i>$, a natural choice consists in defining the horizontal space as the orthogonal complement to the vertical space with respect to this metric. This procedure reduces to consider $\mathcal{G}(S, \mathcal{H})$ as a quotient submanifold of $St(S, H)$. Precisely the

horizontal space at **U** reads:

$$
\begin{aligned}
H_{\mathbf{U}} &= \{\delta\mathbf{V} \in \mathcal{T}_{\mathbf{U}} St(S, \mathcal{H}) \text{ such that } <\delta\mathbf{V}, \mathbf{U}\boldsymbol{\Omega}> = 0, \forall \boldsymbol{\Omega} \in \mathbb{R}^{S \times S} : \boldsymbol{\Omega} = -\boldsymbol{\Omega}^T\} \\
&= \{\mathbf{Z} \in [\mathcal{H}]^S \text{ such that } <Z_i, U_j>_{\mathcal{H}} = 0, \forall i, j = 1, ..., S\}
\end{aligned}
\tag{2.10}
$$

Observe that the horizontal space consists of only variations in $St(S, \mathcal{H})$ which modify the span. Finally the tangent space to $\mathcal{G}(S, \mathcal{H})$ at $\mathcal{U} = \pi(\mathbf{U}) = \text{span}(\mathbf{U})$ is represented by $H_{\mathbf{U}}$. In particular if $\gamma : [0, T] \mapsto \mathcal{G}(S, \mathcal{H})$ is a curve passing by $\mathcal{U} = \pi(\mathbf{U}_0)$ with $\mathbf{U}_0 \in St(S, \mathcal{H})$, the unique horizontal lift of γ through $\mathbf{U}_0$ in $St(S, H)$ is defined as $t \mapsto \mathbf{U}(t)$ and satisfies: $\mathbf{U}(0) = \mathbf{U}_0$ and $<\dot{U}_i(t), U_j>_{\mathcal{H}} = 0, \forall i, j = 1, ..., S$. In the following we refer to the last condition as **dynamically orthogonal constraint**.

2.1 Manifold of S rank random fields

In this thesis we are interested in approximating the solutions of differential equations with random parameters by a linear combination of S (and not less than S) linearly independent deterministic modes combined with S linearly independent stochastic modes. More precisely the approximate solution is sought in the manifold of S rank random fields, defined as follows:

Definition 2.1.1. *We define $\mathcal{M}_S \subset \mathcal{H} \otimes L^2(\Omega)$ the manifold of all S rank random fields, i.e.:*

$$
\mathcal{M}_S = \Big\{ u_S \in \mathcal{H} \otimes L^2(\Omega) : u_S = \sum_{i=1}^S U_i Y_i \mid \quad \begin{aligned} &span(U_1, ..., U_S) \in \mathcal{G}(S, \mathcal{H}), \\ &span(Y_1, ..., Y_S) \in \mathcal{G}(S, L^2(\Omega)) \end{aligned} \Big\}
\tag{2.11}
$$

where $\mathcal{H} \subset L^2(D)$ is a Hilbert space. We use the tools of differential geometry discussed in the previous Sections to show that $\mathcal{M}_S$ is actually a differential manifold. Hence we derive the parametrizations of the tangent space which will be used in the following Chapters.
First of all we emphasize that there are two levels of flexibility in the parametrization of $\mathcal{M}_S$:

- the choice of the atlas to equip $\mathcal{M}_S$ with a manifold structure. Practically, when we define an atlas, we write each element in $\mathcal{M}_S$ as a linear combination of S terms, each one written in separable form. The choice of the atlas consists in fixing in which and how many components each term is separated and setting the functional spaces to which each component belongs.

- the choice of a horizontal distribution to equip $\mathcal{M}_S$ with a differential structure and parametrize the tangent space.

In the context of dynamical low rank approximation, the second step leads to the dynamic constraints, known in physics as gauge constraints, which are used to derive the reduced order system. Let $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ be an inner product (non necessarily the one of $\mathcal{H}$), we consider

the following representation:

$$u_S = \sum_{i=1}^{S} Y_i U_i = \mathbf{U}\mathbf{Y} \tag{2.12}$$

where:

- $\mathbf{U} = (U_1, ..., U_S) \in [\mathcal{H}]^S$ is a S dimensional (row) vector of functions $U_i \in \mathcal{H}$ which are orthonormal with respect to the $\langle \cdot, \cdot \rangle$-product, i.e. $\langle U_i, U_j \rangle = \delta_{ij}$ for all $i, j = 1, ..., S$.

- $\mathbf{Y}$ is a row vector of S linearly independent random variables, hence with full rank covariance matrix $\mathbf{C} = \mathbb{E}[\mathbf{Y}\mathbf{Y}^T]$.

In practice, $\mathcal{H}$ will be same Sobolev space $H^s(D)$ for $s \geq 0$, and $\langle \cdot, \cdot \rangle$ will be the $L^2(D)$-inner product.

Remark 2.1.1. *Here we have assumed that $\langle \cdot, \cdot \rangle$ is an inner product, i.e. a symmetric and positive bilinear form in $\mathcal{H}$. Actually this condition can be relaxed by asking $\langle \cdot, \cdot \rangle$ to be a non degenerate bilinear form in $\mathcal{H}$, i.e $\langle U, V \rangle = 0$ for all $V \in \mathcal{H}$ if and only if $U = 0$. The non-degeneracy is a sufficient condition to equip $\mathcal{M}_S$ with a structure of differentiable quotient manifold. This type of construction will be used in Chapter 5, where we take $\langle \cdot, \cdot \rangle$ to be a symplectic (i.e. non degenerate and antisymmetric) form in $\mathcal{H}$.*

Hereafter we refer to (2.12) as DO decomposition, as used in [116, 117]. Observe that the deterministic modes $\mathbf{U}$ belong to the Stiefel manifold $St(S, \mathcal{H})$ defined in (2.7), that, we recall, is an embedded submanifold of $[\mathcal{H}]^S$ (see Section 2.0.3). Let us denote by $\mathrm{B}(S, L^2(\Omega)) \subset [L^2(\Omega)]^S$ the sub-manifold of all $L^2(\Omega)$-linearly independent random vectors of dimension S; we claim that $\mathcal{M}_S$ admits a structure of differential manifold based on the isomorphism with $\big(\mathcal{G}(S, \mathcal{H}) \times \mathrm{B}(S, L^2(\Omega))\big)$ where the Grassmannian manifold $\mathcal{G}(S, \mathcal{H}) = St(S, \mathcal{H})/\mathcal{O}_S$ is interpreted as the Riemanniann quotient manifold of $St(S, \mathcal{H})$. The proof is based on the fact that horizontal distributions in $St(S, \mathcal{H})$ imply the existence of local section maps $\sigma^{\mathbf{U}} : \mathcal{B}_{\pi(\mathbf{U})} \subset \mathcal{G}(S, \mathcal{H}) \to St(S, \mathcal{H})$, defined in a neighborhood $\mathcal{B}_{\pi(\mathbf{U})}$ of $\pi(\mathbf{U})$, for all $\mathbf{U} \in St(S, \mathcal{H})$.

Proposition 2.1.1. *$\mathcal{M}_S$ admits a manifold structure diffeomorphic to $\big(\mathcal{G}(S, \mathcal{H}) \times \mathrm{B}(S, L^2(\Omega))\big)$ according to which, for any $u_S \in \mathcal{M}_S$, the tangent space to $\mathcal{M}_S$ at u_S can be parametrized as follows:*

$$\mathcal{T}_{u_S}\mathcal{M}_S = \Big\{ \delta v = \sum_{i=1}^{S} \big(\delta U_i Y_i + U_i \delta Y_i \big) \in \mathcal{H} \otimes L^2(\Omega) \quad \text{with } \delta Y_i \in L^2(\Omega) \text{ and } \delta U_i \in \mathcal{H},$$
$$\text{s.t. } \langle \delta U_i, U_j \rangle = 0 \; \forall i, j = 1, ..., S \Big\} \tag{2.13}$$

Proof. We start by considering the Grassmannian manifold $\mathcal{G}(S, \mathcal{H})$, that is the collection of all S dimensional subspaces of $\mathcal{H}$. Following the discussion in Section 2.0.4, we have that

$\mathcal{G}(S,\mathcal{H})$ is diffeomorphic to $St(S,\mathcal{H})/\mathcal{O}_S$ and admits a quotient manifold structure. Moreover $(St(S,\mathcal{H}),\mathcal{G}(S,\mathcal{H}),\pi,\mathcal{O}_S)$ forms a principle fiber bundle, where the projection map π corresponds to the span operation (see Section 2.0.5). We consider in $St(S,\mathcal{H})$ the L^2 norm defined as: $\langle \mathbf{V},\mathbf{U}\rangle = \sum_{i=1}^{S}\langle V_i, U_i >_{L^2(D)}$ and we choose the horizontal distribution by means of L^2 projection on the orthogonal complement to the vertical space. Then for any $\mathbf{U} \in St(S,\mathcal{H})$ the horizontal space is written as:

$$H_{\mathbf{U}} = \{\mathbf{Z} \in [\mathcal{H}]^S \text{ such that } \langle Z_i, U_j\rangle = 0, \forall i, j = 1,...,S\} \tag{2.14}$$

We recall that a horizontal distribution turns the map $D_{\mathbf{U}}\pi : H_{\mathbf{U}} \to \mathcal{T}_{\pi(\mathbf{U})}\mathcal{G}(S,\mathcal{H})$ into an isomorphism, for any $\mathbf{U} \in St(S,\mathcal{H})$. Namely, by fixing the horizontal space $H_{\mathbf{U}}$, we implicitly define a linear map $D_{\pi(\mathbf{U})}\sigma^{\mathbf{U}} : \mathcal{T}_{\pi(\mathbf{U})}\mathcal{G}(S,\mathcal{H}) \to \mathcal{T}_{\mathbf{U}}St(S,\mathcal{H})$, which is the right inverse of the differential $D_{\mathbf{U}}\pi$, and which coincides to the push-forward of the local section map $\sigma^{\mathbf{U}}$ induced by the horizontal distribution, defined in a neighborhood of $\pi(\mathbf{U})$. We now define the following maps:

$$\begin{aligned} s: \quad \left(St(S,\mathcal{H}) \times \mathrm{B}(S,L^2(\Omega))\right) &\to \mathcal{M}_S \\ (\mathbf{U},\mathbf{Y}) &\mapsto \sum_{i=1}^{S} U_i Y_i \end{aligned} \tag{2.15}$$

$$\begin{aligned} \tilde{\sigma} := s \circ (\sigma^{\mathbf{U}} \times i_d): \quad \left(\mathcal{B}_{\mathcal{U}} \times \mathrm{B}(S,L^2(\Omega))\right) &\to \mathcal{M}_S \\ (\mathcal{V},\mathbf{Y}) &\mapsto \sigma^{\mathbf{U}}(\mathcal{V})\mathbf{Y} = \sum_{i=1}^{S}\left(\sigma^{\mathbf{U}}(\mathcal{V})\right)_i Y_i \end{aligned} \tag{2.16}$$

where $\sigma^{\mathbf{U}}$, defined in a neighborhood $\mathcal{B}_{\mathcal{U}} \subset \mathcal{G}(S,\mathcal{H})$ of $\mathcal{U} = \pi(\mathbf{U})$, is the local section map associated to the quotientation of $St(S,\mathcal{H})$. The map $\tilde{\sigma}$ is a diffeomorphism, hence the tangent space to $\mathcal{M}_S$ can be represented in terms of the differential of $\tilde{\sigma}$: for any $v_S = \tilde{\sigma}(\mathcal{V},\mathbf{Y})$ and $\delta v_S \in \mathcal{T}_{v_S}\mathcal{M}_S$

$$\begin{aligned} \delta v_S &= D_{(\mathcal{V},\mathbf{Y})}\tilde{\sigma}(\delta\mathcal{V},\delta\mathbf{Y}) \\ &= \sum_{i=1}^{S}\left(D_{\mathcal{V}}\sigma^{\mathbf{U}}(\delta\mathcal{V})\right)_i Y_i + \sum_{i=1}^{S}\left(\sigma^{\mathbf{U}}(\mathcal{V})\right)_i \delta Y_i \\ &= \delta\mathbf{V}\mathbf{Y} + \mathbf{V}\delta\mathbf{Y} \quad \text{such that } \mathbf{V} = \sigma^{\mathbf{U}}(\mathcal{V}), \text{ and } \langle V_i, \delta V_j\rangle = 0, \forall i, j = 1,...,S. \end{aligned} \tag{2.17}$$

Here we explicitly used the horizontal distribution defined by (2.14). $\qquad\square$

This isomorphism allows us to recast variational problems, defined in the abstract manifold $\mathcal{M}_S$, as systems of equations defined respectively in $\mathcal{H}$ and $L^2(\Omega)$.
Observe that we can retrace the same construction as before by inverting the role of the deterministic and stochastic bases and get the Dual DO decomposition, used in Chapter 4, in which the stochastic modes are now orthonormal and subject to orthogonal dynamic constraints.

2.1.1 Some alternative parametrizations

Alternative parametrizations of low rank manifolds in separable form have been investigated in literature and used to derive reduced order methods or rank-constrained geometrical optimization algorithms. We have already mentioned the DyBO (or BO) method introduced in [30] and used in [31, 36], consisting in a representation in three fields:

$$u_S(x, t, \omega) = \sum_{i=1}^{S} \lambda_i(t) U_i(x, t) Y_i(t, \omega)$$

which maintains both the deterministic and the stochastic modes orthogonal (but non orthonormal). This technique is included in the class of Dynamical Low Rank method, i.e. the DyBO-approximate solution satisfies the DLR variational principle (1.4.1) and is equivalent the (Dual) Dynamical approximate solution, provided that no crossing between the eigenvalues $\lambda_1, ..., \lambda_S$ occurs. The equivalence between the DO and DyBO reduced systems is shown in [35]. For deterministic Schrödinger equations, the same result in variational setting has been derived in [12] (Section 3.2 and Section 3.2) and can be recast into the setting of PDEs with random coefficients by following the discussion in Chapter 1.

Other different geometrical constructions of manifold with fixed rank in finite dimensional setting, have been recently introduced in [93], for machine learning applications. In this framework $\mathcal{M}_S$ is the manifold of all matrices with rank S, and the different types of representations of S rank random fields recalled above correspond to different matrix factorizations. In [93], the authors proposed three geometrical interpretations underling the full-rank factorization (corresponding to the Cholesky-type decomposition), the polar factorization (corresponding to the SVD decomposition) and the subspace-projection factorization (corresponding to the QR decomposition). The former is the finite dimensional analogue of the DO representation for S rank random fields. However the differential structure proposed in [93] (which is summarized hereafter) is different to the one proposed here and, as a result, leads to a different parametrization of the tangent space. The parametrization proposed in [93] appears quite natural but leads to more involved dynamic constraints. It is based on considering $\mathcal{M}_S$ as the collection of the equivalence classes defined as follow:

$$[(\mathbf{U}, \mathbf{Y})] = \left\{ (\mathbf{U}\mathbf{O}, \mathbf{O}^T\mathbf{Y}), \ \forall \mathbf{O} \in \mathcal{O}_S \right\}$$

More precisely, the principle fiber bundle associated to $\mathcal{M}_S$ is given by:

- the total space $\big(St(S, \mathcal{H}), B(S, L^2(\Omega)) \big)$;

- the base space $\big(St(S, \mathcal{H}), B(S, L^2(\Omega)) \big) / \mathcal{O}_S$ defined with respect to the following equivalence relation:
  $$(\mathbf{U}, \mathbf{Y}) \sim (\mathbf{V}, \mathbf{Z}) \iff \exists \mathbf{O} \in \mathcal{O}_S : (\mathbf{V}, \mathbf{Z}) = (\mathbf{U}\mathbf{O}, \mathbf{O}^T\mathbf{Y}),$$
  for all $\mathbf{U}, \mathbf{V} \in St(S, \mathcal{H})$ and $\mathbf{Y}, \mathbf{Z} \in B(S, L^2(\Omega))$;

- the Lie group $\mathcal{O}_S$ with right action defined as:

$$
\begin{aligned}
\left(St(S,\mathcal{H}),B(S,L^2(\Omega))\right) \times \mathcal{O}_S &\;\rightarrow\; \left(St(S,\mathcal{H}),B(S,L^2(\Omega))\right) \\
((\mathbf{U},\mathbf{Y}),\mathbf{O}) &\;\mapsto\; (\mathbf{U}\mathbf{O},\mathbf{O}^T\mathbf{Y})
\end{aligned}
$$

- the projection map π which associates to any pair $(\mathbf{U},\mathbf{Y}) \in \left(St(S,\mathcal{H}),B(S,L^2(\Omega))\right)$ the corresponding equivalence class.

and $\mathcal{M}_S$ is identified with the base space $\left(St(S,\mathcal{H}),B(S,L^2(\Omega))\right)/\mathcal{O}_S$. Moreover, the total space is equipped with the following $\mathcal{O}_S$-invariant Riemannian metric:

$$
g\left((\mathbf{V},\mathbf{Z}),(\mathbf{W},\mathbf{X})\right) = \sum_{i=1}^{S} < V_i, W_i > +\operatorname{Trace}\left(\mathbb{E}[\mathbf{Z}\mathbf{X}^T]\mathbb{E}[\mathbf{Y}\mathbf{Y}^T]^{-1}\right).
$$

For any $(\mathbf{U},\mathbf{Y}) \in \left(St(S,\mathcal{H}),B(S,L^2(\Omega))\right)$, the corresponding fiber is written as:

$$
\pi^{-1}[(\mathbf{U},\mathbf{Y})] = \left\{(\mathbf{U}\mathbf{O},\mathbf{O}^T\mathbf{Y}),\ \forall \mathbf{O} \in \mathcal{O}_S\right\},
$$

and the vertical space at $(\mathbf{U},\mathbf{Y})$ is given by:

$$
V_{(\mathbf{U},\mathbf{Y})} = \left\{(\mathbf{U}\boldsymbol{\Omega},\boldsymbol{\Omega}^T\mathbf{Y}),\ \forall \boldsymbol{\Omega} \in \mathbb{R}^{S\times S} : \boldsymbol{\Omega}^T = -\boldsymbol{\Omega}\right\}.
$$

When endowed with the metric g, the base space $\left(St(S,\mathcal{H}),B(S,L^2(\Omega))\right)/\mathcal{O}_S$ becomes the Riemannian quotient manifold of $\left(St(S,\mathcal{H}),B(S,L^2(\Omega))\right)$ by $\mathcal{O}_S$. In other words, the tangent space to $\mathcal{M}_S$, seen as $\left(St(S,\mathcal{H}),B(S,L^2(\Omega))\right)/\mathcal{O}_S$, is parametrized in terms of the horizontal distribution induced by the metric g, and is written as:

$$
\begin{aligned}
\mathcal{T}_{u_S}\mathcal{M}_S \;=\; \Big\{ \delta v = \textstyle\sum_{i=1}^{S} \left(\delta U_i Y_i + U_i \delta Y_i\right) \;\in \mathcal{H}_1 \otimes L^2(\Omega) \text{ with } \delta Y_i \in L^2(\Omega) \text{ and } \delta U_i \in \mathcal{H}_1, \\
\text{s.t. } \ll \mathbf{U},\delta\mathbf{U} \gg +\left(\mathbb{E}[\mathbf{Y}\delta\mathbf{Y}^T]\mathbb{E}[\mathbf{Y}\mathbf{Y}^T]^{-1}\right) \text{ is simmetric} \Big\}
\end{aligned}
$$

$$(2.18)$$

at any $u_S = \mathbf{U}\mathbf{Y} \in \mathcal{M}_S$.

3 Dynamical Low-Rank approximation for parabolic PDEs with random data

This Chapter is mainly based on the paper [96] with respect to which we have done minor changes in the notation (i.e. $B_{[L_0^2(\Omega)]^s}$ has been replaced with $B(S, L_0^2(\Omega))$) and added a numerical test considering a parabolic equation with non linear reaction term. In particular, in Section 3.2 we introduce the Dynamically Orthogonal method, we show the analogy with the Dynamically Double Orthogonal method and we formalize the Dynamical Low-Rank variational principle for parabolic equations with random parameters which establishes the link between the DO(DDO) method and the DLR approximation. Afterwards, in Section 3.2.3, we analyze same properties of the approximation manifold which are then used in Section 3.3.1 to derive a theoretical bound for the approximation error of the $S-$terms DLR approximate solution by the corresponding S-terms best approximation, for linear parabolic equations. After studying the main properties of the DO approximations on simple cases of deterministic equations with random initial data in Section 3.4, we conclude in Section 3.5 with some numerical tests which confirm the theoretical bound and show potentials and limitations of the DLR approach.

Introduction

Many physical and engineering problems can be properly described by mathematical models, typically of differential type. However, in many situations, the input parameters may be affected by uncertainty due e.g. to measurement errors, limited data availability or intrinsic variability of the phenomenon itself. A convenient way to characterize uncertainty consists in describing the uncertain parameters as random variables or space and/or time varying random fields. Starting from a suitable Partial Differential Equation (PDE) model, the aim of the Uncertainty Quantification is to assess the effects of the uncertainty by computing the statistics of the solutions or of some quantities of interest. Several approaches have been proposed and analyzed in the last decades. We name the Monte Carlo method [46, 21], Quasi Monte Carlo [51, 98] and the corresponding Multilevel versions [37], or the approaches based on deterministic approximations of the parameters-to-solution map (response function) such as the generalized Polynomial Chaos [38, 136, 131, 75] in its Galerkin [6, 89, 49] and collocation

versions [11, 5, 101, 133, 135].

In this work we focus on a reduced basis method to approximate the solution. We consider a general type of time dependent PDE with random data of the form:

$$
\begin{cases}
\frac{\partial u(x,t,\omega)}{\partial t} = \mathscr{L}(u(x,t,\omega),\omega), & x \in D,\ t \in \mathscr{T},\ \omega \in \Omega, \\
\mathscr{B}(u(\sigma,t,\omega)) = h(\sigma,t), & \sigma \in \partial D,\ t \in \mathscr{T},\ \omega \in \Omega, \\
u(x,t=0,\omega) = u_0(x,\omega), & x \in D,\ \omega \in \Omega,
\end{cases}
\tag{3.1}
$$

where $x \in D \subset \mathbb{R}^d$ is the spatial coordinate, t is the time variable in $\mathscr{T} \equiv [0,T]$ and ω is the random elementary event in the complete probability space $(\Omega, \mathscr{A}, \mathscr{P})$. In addition $\mathscr{L}$ is a general (linear or non linear) differential operator and $\mathscr{B}$ is an operator defining the boundary conditions. Here the randomness can appear in the operator $\mathscr{L}$ as a random parameter or forcing term as well as in the initial datum. A possible approach to approximate the solution consists in expanding u on a deterministic (Proper Orthogonal Decomposition- POD [23, 13, 62]) or stochastic (gPC [136, 139, 100]) set of orthogonal basis functions, performing a Galerkin projection and computing the coefficients at any time step. Specifically, the POD method requires a set of pre-computed snapshots for different parameter values and time instants. However, since the dependence of the solution on the random parameters may significantly vary in time, any approximation which makes use of time fixed basis functions (either deterministic or stochastic), necessarily requires during the evolution an increasing number of terms to maintain a proper level of accuracy and, in general, needs a very high computational effort. Several adaptive and greedy type techniques have been proposed in the literature to (partially) overcome this problem, e.g. time-dependent gPC [48, 59] and Generalized POD [104, 106, 33]. On the other hand, in many cases, the collection of all solutions at a given time corresponding to all possible outcomes of the input random processes can still be well approximated in a low-dimensional subspace, which however, will change at each time instant.

It is well known that the optimal S-dimensional subspace, in L^2 sense, is the one which is spanned by the first S terms of the Karhunen-Loève decomposition of the solution [75, 49, 78]. The main practical difficulty is that such subspace is, in general, not easy to characterize a priori and might significantly change in time. Therefore the idea of the approach proposed here is to approximate the solution on an evolving subspace, exploiting the structure of the differential equation. In other words, the approximate solution is expanded on a dynamical deterministic orthonormal basis with stochastic coefficients which evolve in time as well, i.e.:

$$
u_S(x,t,\omega) = \bar{u}_S(x,t) + \sum_{i=1}^{S} Y_i(t,\omega) U_i(x,t),
\tag{3.2}
$$

Here $\bar{u}_S(x,t) \simeq \mathbf{E}[u(x,t,\cdot)]$ is the approximated expected value, $U_1, ..., U_S$ are $L^2(D)-$orthonormal deterministic basis functions and $Y_1, ..., Y_S$ are zero-mean stochastic variables. The approximate solution (3.2) is obtained by suitably projecting the residual of the differential equation (non linear Galerkin projection) and aims to be close enough to the

Karhunen-Loève decomposition, even if it does not coincide, in general, with it.

This approach is not new; it has been introduced in [116] and named "Dynamically Orthogonal" approximation (DO) and applied in [117, 127] to the approximation of fluid equations with random initial data. Analogous formulations are also used in [30, 31, 34]. On the other hand, similar ideas have been developed in a quite different context, namely in chemistry and quantum dynamics, for the approximation of the deterministic Schrödinger equations by the Multi-Configuration time-dependent Hartree method (MCTDH, [67, 39]) and the Dirac-Frenkel Variation principal [42, 47]. There, the goal is to look for an approximate solution written in separable form as a product of functions depending on one space variable only, whereas, in the DO approach presented here, we aim at separating the space variables from the stochastic ones. The discrete analogue of the MCTDH method consists in looking for a Dynamical Low-Rank approximation of a deterministic evolution matrix or tensor equation [66, 68, 83]. A few theoretical results are available on the accuracy and error estimates for either the MCTDH approximation of Schrödinger equations or Dynamical Low-Rank approximation of matrix equations [39, 66].

Our first goal in this paper is to establish a precise link between the DO approach (as proposed in [116]) and the Dynamical Low-Rank approximation analyzed e.g. in [66]. This allows us to "import" some of the theoretical results developed in [39, 66] to our situation of a parabolic equation with random parameters. In particular, we reinterpret the DO equations given in [116, 117, 127] as a Galerkin projection onto the tangent space to the manifold of the rank S functions of the form (3.2). Using curvature bounds for such manifold, given in [39], we show that the DO approximation error for a linear parabolic equation with random input data can be bounded in terms of the best rank S approximation of the solution (Karhunen-Loève expansion), at each time instant. The bound is applicable on the largest time interval in which the best $S-$terms approximation is continuously differentiable in time. This request on time differentiability is actually unavoidable and corresponds to asking that certain eigenvalues of the Karhunen-Loève decomposition do not cross in time. By means of simple examples with a deterministic linear operator and random initial datum, we highlight how and when the crossing of the eigenvalues negatively effects the DO approximation. In particular we show in which cases, for a deterministic operator, the DO solution is exact and on the other hand, when the DO error can not be properly bounded by the best approximation error. Finally, we describe the numerical method that we have adopted in this work and the technique utilized to deal with singular covariance matrices. We conclude with some numerical examples in which we specifically address: i) a deterministic linear parabolic equation with random initial condition, ii) a linear parabolic equation with stochastic coefficient and deterministic initial datum, iii) a non-linear parabolic equation of reaction-diffusion type.

The outline of the paper is as follows: in Section 2 we introduce the mathematical problem and basic notations; in Section 3 we describe the DO approximation, we show the analogy with the dynamical low-rank approach and we give a variational interpretation of it. In Section 4 the DO approximation is applied to a linear stochastic parabolic equation and an analysis of the DO approximation error is provided. In Section 5 we analyze the case of a linear deterministic operator. Finally, in Section 6 we describe the numerical discretization of the DO equations

and we present several numerical test cases that will show when the DO approximation is effective and when is not.

3.1 Problem setting

Let $D \subset \mathrm{R}^d$, $1 \le d \le 3$, be an open bounded domain and $(\Omega, \mathscr{A}, \mathscr{P})$ a complete probability space, where Ω is the set of outcomes, $\mathscr{A}$ a σ-algebra and $P : \mathscr{A} \to [0,1]$ a probability measure. The problem considered in this paper is the following time dependent stochastic PDE:

$$\frac{\partial u(x,t,\omega)}{\partial t} = \mathscr{L}(u(x,t,\omega),\omega), \quad x \in D, \ t \in \mathscr{T}, \ \omega \in \Omega, \tag{3.3}$$

where $\mathscr{L}$ is a general (linear or non-linear) differential operator, $x \in D$ is the spatial coordinate and t is the time variable in $\mathscr{T} \equiv [0, T]$. Additionally, the initial state of the problem is described by

$$u(x, t = 0, \omega) = u_0(x, \omega), \quad x \in D, \ \omega \in \Omega, \tag{3.4}$$

and the (deterministic) boundary condition is given by

$$\mathscr{B}(u(\sigma, t, \omega)) = h(\sigma, t), \quad \sigma \in \partial D,$$

where $\mathscr{B}$ is a linear differential or algebraic operator. We specifically address the parabolic case in which $\mathscr{L}$ is an elliptic second order differential operator in the space variable x. For a random function $v(x, t, \omega)$, we define its mean value as

$$\bar{v}(x, t) = \mathbb{E}[v(x, t, \cdot)] = \int_\Omega v(x, t, \omega) d\mathscr{P}(\omega),$$

as well as the L^2 inner product in the physical space

$$\langle u(\cdot, t, \omega), v(\cdot, t, \omega) \rangle = \int_D u(x, t, \omega) v(x, t, \omega) dx.$$

In what follows we use the notation

$$u^*(x, t, \omega) = u(x, t, \omega) - \mathbb{E}[u(x, t, \cdot)]$$

We assume that all the random fields considered in this paper are square integrable for any $t \in \mathscr{T}$, that is,

$$\int_D \mathbb{E}[u^2(x, t, \cdot)] dx < +\infty \quad \forall t \in \mathscr{T}.$$

As the approaches considered in this work have a strong relationship with the Karhunen-Loève expansion, we review some basic properties of the latter. To begin with, let us define the

covariance function of a space-dependent random field $u(x,\omega)$ as

$$\mathrm{Cov}_u(x,y) = \mathbb{E}\left[u^*(x,\cdot)u^*(y,\cdot)\right], \quad x,y \in D.$$

It is well known that any second order random field $u(x,\omega)$, with continuous and positive definite covariance function $\mathrm{Cov}_u : \overline{D \times D} \to \mathbb{R}$, can be represented as an infinite sum of random variables, by means of the Karhunen-Loève expansion [49]. To this end, we introduce the compact and self-adjoint operator $T_u : L^2(D) \to L^2(D)$, which is defined by

$$T_u v(\cdot) = \int_D \mathrm{Cov}_u(x,\cdot)v(x)dx, \quad \forall v \in L^2(D). \tag{3.5}$$

Then, consider the sequence of non-negative decreasing eigenvalues of T_u, $\{\mu_i\}_{i=1}^\infty$, and the corresponding sequence of orthonormal eigenfunctions, $\{Z_i\}_{i=1}^\infty$, satisfying

$$T_u Z_i = \mu_i Z_i, \quad \langle Z_i, Z_j \rangle = \delta_{ij} \qquad \forall i,j \in \mathbb{N}^+, \tag{3.6}$$

where δ_{ij} is the Kronecker symbol. In addition, define the mutually uncorrelated real random variables

$$\gamma_i(\omega) := \frac{1}{\sqrt{\mu_i}} \int_D u^*(x,\omega)Z_i(x)dx \qquad \forall i \in \mathbb{N}^+, \tag{3.7}$$

with zero mean and unit variance, i.e. $\mathbb{E}[\gamma_i] = 0$ and $\mathbb{E}[\gamma_i \gamma_j] = \delta_{ij}$ for $i,j \in \mathbb{N}^+$. Then, the truncated Karhunen-Loève expansion of the stochastic function u, which we denote by z_S, is defined by

$$z_S(x,\omega) = \bar{u}(x) + \sum_{i=1}^{S} \sqrt{\mu_i}\gamma_i(\omega)Z_i(x), \ \forall S \in \mathbb{N}^+. \tag{3.8}$$

By Mercer's theorem [78], it follows that

$$\lim_{S \to \infty} \sup_{x \in D} \mathbb{E}[(u(x,\cdot) - z_S(x,\cdot))^2] = \lim_{S \to \infty} \sup_{x \in D} \sum_{i=S+1}^{\infty} \mu_i Z_i^2(x) = 0.$$

Observe that the S random variables in (3.7), describing the approximate random function z_S (3.8), are weighted differently due to the decay of the eigenvalues of the Karhunen-Loève expansion. The decay properties of eigenvalues and eigenvectors have been investigated e.g. in the works [49, 119].

In the case of a time-varying random filed $u(x,t,\omega)$, the truncated Karhunen-Loève expansion at each fixed $t \in \mathscr{T}$ would read

$$z_S(x,\omega,t) = \bar{u}(x,t) + \sum_{i=1}^{S} \sqrt{\mu_i(t)}\gamma_i(t,\omega)Z_i(x,t), \ \forall S \in \mathbb{N}^+ \tag{3.9}$$

with $\langle Z_i(\cdot,t), Z_j(\cdot,t) \rangle = \delta_{ij}, \ \forall t \in \mathscr{T}$.

3.2 Dynamically Orthogonal approximation

Several approaches have been proposed in the literature to numerically compute the random solution $u(x, t, \omega)$ of PDEs with stochastic input data. For instance, in a generalized Polynomial Chaos (gPC) approach (see e.g. [75, 136, 131]), after parameterizing the probabilistic space by a sequence of random variables $\{\eta_i\}_{i \geq 1}$, the solution is expanded on a fixed basis of orthogonal polynomials in η_i with space and time varying coefficients:

$$v_S(x, t, \omega) = \bar{v}_S(x, t) + \sum_{i=1}^{S} V_i(x, t) \Phi_i(\eta_1(\omega), \eta_2(\omega), \ldots) = \bar{v}_S(x, t) + \sum_{i=1}^{S} V_i(x, t) \tilde{\Phi}_i(\omega),$$

and $\mathbb{E}[\tilde{\Phi}_i \tilde{\Phi}_j] = \delta_{ij}$.

Unlike the gPC approach, the Dynamically Orthogonal (DO) approximation, first introduced in [116], utilizes a more general expansion

$$u_S(x, t, \omega) = \bar{u}_S(x, t) + \sum_{i=1}^{S} U_i(x, t) Y_i(t, \omega). \tag{3.10}$$

Namely, both the spatial basis $\{U_i(x, t)\}_{i=1}^{S}$ and the random basis $\{Y_i(t, \omega)\}_{i=1}^{S}$ are time dependent and either U_i or Y_i are kept orthogonal at all times, thus aiming to mimic the Karhunen-Loève expansion (3.9). Note that the above approximations are finite sums where the index S represents the approximation level.

In what follows we focus on the case where the spatial basis $\{U_i\}_{i=1}^{S}$ is kept orthogonal at all times. The uniqueness of the DO approximation (3.10) is guaranteed by the following *dynamically orthogonal* conditions [116],[117],[127]:

$$\mathbb{E}[Y_i(t, \cdot)] = 0, \quad \langle U_i(\cdot, t), U_j(\cdot, t) \rangle = \delta_{ij},$$
$$\langle \frac{\partial U_i(\cdot, t)}{\partial t}, U_j(\cdot, t) \rangle = 0, \ 1 \leq i, j \leq S, \ \forall t \in \mathcal{T}. \tag{3.11}$$

Given problem (3.3), by using together the Galerkin projection onto the subspaces spanned by the basis functions in (3.10) and the DO conditions (3.11), one gets the following DO system[116, 117, 127]:

$$\begin{cases} \dfrac{\partial \bar{u}_S(x, t)}{\partial t} = \mathbb{E}[\mathcal{L}(u_S(x, t, \cdot))] & \text{(3.12)} \\[2ex] \displaystyle\sum_{i=1}^{S} \mathbf{C}_{ij}(t) \dfrac{\partial U_i(x, t)}{\partial t} = \Pi_{\mathcal{U}}^{\perp} \mathbb{E}[\mathcal{L}(u_S(x, t, \cdot)) Y_j(t, \cdot)] & j = 1, \cdots, S & \text{(3.13)} \\[2ex] \dfrac{\partial Y_i(t, \omega)}{\partial t} = \langle \mathcal{L}^*(u_S(\cdot, t, \omega), \omega), U_i(\cdot, t) \rangle & i = 1, \cdots, S & \text{(3.14)} \end{cases}$$

where

$$\mathbf{C}_{ij}(t) = \mathbb{E}[Y_i(t, \cdot) Y_j(t, \cdot)], \qquad \forall i, j = 1, \cdots, S,$$
$$\mathcal{L}^*(u(x, t, \omega), \omega) = \mathcal{L}(u(x, t, \omega), \omega) - \mathbb{E}[\mathcal{L}(u(x, t, \cdot))]$$

and $\Pi_{\mathscr{U}}^{\perp}$ is the projection operator from the space $L^2(D)$ to the orthogonal complement of the S dimensional subspace $\mathscr{U} = \mathrm{span}\{U_1, \cdots, U_S\}$, namely,

$$\Pi_{\mathscr{U}}^{\perp}[v] = v - \Pi_{\mathscr{U}}[v] = v - \sum_{i=1}^{S} \langle v, U_i \rangle U_i, \quad \forall v \in L^2(D).$$

The associated boundary conditions have the form

$$\begin{aligned}
\mathscr{B}(\bar{u}_S(\sigma, t)) &= h(\sigma, t), & \sigma &\in \partial D \\
\sum_{i=1}^{S} \mathbf{C}_{ij}(t) \mathscr{B}(U_i(\sigma, t)) &= 0, & \sigma &\in \partial D, \; j = 1, \cdots, S
\end{aligned} \tag{3.15}$$

and the corresponding initial conditions are given by

$$\bar{u}_S(x, 0) = \bar{u}_0(x) = \mathbb{E}[u_0(x, \cdot)], \quad U_i(x, 0) = Z_{i0}(x), \quad Y_i(0, \omega) = \langle u_0(\cdot, \omega) - \bar{u}_0, Z_{i0} \rangle, \tag{3.16}$$

where $\{Z_{i0}(x)\}_{i=1}^{S}$ are the spatial basis functions appearing in the Karhunen-Loève expansion of $u_0(x, \omega)$. Note that the DO equations (3.12)-(3.14) are coupled together, in general. By solving this system, one easily gets the approximation of the mean and of the total variance of the solution:

$$\mathbb{E}[u(x, t, \cdot)] \approx \mathbb{E}[u_S(x, t, \cdot)] = \bar{u}_S(x, t), \qquad \mathrm{Var}_T[u](t) \approx \mathrm{Var}_T[u_S](t) = \sum_{i=1}^{S} \mathbb{E}[Y_i^2(t)].$$

where the total variance is defined as $\mathrm{Var}_T[u](t) = \int_D \mathbb{E}[(u(x, t, \cdot) - \bar{u}(x, t, \cdot))^2] dx$. Concerning the numerical approximation of the DO system (3.12)-(3.14), many approaches can be followed, among which the Finite Elements or the Finite Difference methods for spacial discretization and the Stochastic Collocation [5, 101, 135], gPC [137, 136] or (Quasi) Monte Carlo [46, 21] methods for the stochastic discretization. Any time splitting scheme can be adopted for the time derivative discretization, but care should be taken in respecting exactly or with good accuracy, the DO conditions (3.11) at each time step.

3.2.1 Dynamically Double Orthogonal approximation

The DO conditions (3.11) in the derivation of the DO approach are somehow unsymmetric as only the deterministic fields $\{U_i\}_{i=1}^{S}$ are required to be orthogonal. An alternative approach consists in considering a double orthogonal basis $\{\tilde{U}_i\}_{i=1}^{S}$ and $\{\tilde{Y}_i\}_{i=1}^{S}$ and the general formulation:

$$u(x, t, \omega) \approx \bar{u}_S(x, t, \omega) = \tilde{\bar{u}}_S(x, t) + \sum_{i,j=1}^{S} A_{ij}(t) \tilde{U}_i(x, t) \tilde{Y}_j(t, \omega) = \tilde{\bar{u}}_S + \tilde{\mathbf{U}}^T \mathbf{A} \tilde{\mathbf{Y}}, \tag{3.17}$$

with notations
$$\tilde{\mathbf{U}} = (\tilde{U}_1, ..., \tilde{U}_S)^T, \quad \mathbf{A} = \left(A_{ij} \right)_{i,j=1}^{S}, \quad \tilde{\mathbf{Y}} = (\tilde{Y}_1, ..., \tilde{Y}_S)^T.$$

Here we require that both $\{\tilde{U}_i\}_{i=1}^S$ and $\{\tilde{Y}_i\}_{i=1}^S$ are dynamically orthonormal, or rather:

$$
\begin{cases}
\mathbb{E}[\tilde{Y}_i(t,\cdot)] = 0, & \forall\, 1 \leq i \leq S, & (3.18)\\[2mm]
\langle \tilde{U}_i(\cdot,t), \tilde{U}_j(\cdot,t)\rangle = \delta_{ij}, \quad \mathbb{E}\left[\tilde{Y}_i(\cdot,t),\tilde{Y}_j(\cdot,t)\right] = \delta_{ij}, & \forall\, 1 \leq i,j \leq S, & (3.19)\\[2mm]
\langle \dfrac{\partial \tilde{U}_i(\cdot,t)}{\partial t}, \tilde{U}_j(\cdot,t)\rangle = 0, \quad \mathbb{E}\left[\dfrac{\partial \tilde{Y}_i(\cdot,t)}{\partial t},\tilde{Y}_j(\cdot,t)\right] = 0, & \forall\, 1 \leq i,j \leq S. & (3.20)
\end{cases}
$$

Analogously to what has been done in the DO approximation, one can easily derive the following dynamically double orthogonal (DDO) system:

$$
\begin{cases}
\dfrac{\partial \tilde{u}_S(x,t)}{\partial t} = \mathbb{E}[\mathscr{L}(\tilde{u}_S(x,t,\cdot))], & (3.21)\\[3mm]
\dfrac{\mathrm{d}}{\mathrm{dt}}\mathbf{A}(t) = \mathbb{E}\left[\langle \mathscr{L}^*(\tilde{u}_S(\cdot,t,\cdot)), \tilde{\mathbf{U}}^T(\cdot,t)\rangle \tilde{\mathbf{Y}}(\cdot,t)^T\right], & (3.22)\\[3mm]
\mathbf{A}^T(t)\dfrac{\partial \tilde{\mathbf{U}}(x,t)}{\partial t} = \Pi_{\tilde{\mathscr{U}}}^{\perp}\mathbb{E}\left[\tilde{\mathbf{Y}}(\cdot,t)\mathscr{L}^*(\tilde{u}_S(\cdot,t,\cdot))\right] & (3.23)\\[3mm]
\mathbf{A}(t)\dfrac{\partial \tilde{\mathbf{Y}}(t,\omega)}{\partial t} = \Pi_{\tilde{\mathscr{Y}}}^{\perp}\langle \mathscr{L}^*(\tilde{u}_S(\cdot,t,\omega),\omega), \tilde{\mathbf{U}}(\cdot,t)^T\rangle, & (3.24)
\end{cases}
$$

where $\mathscr{L}^*(u) = \mathscr{L}(u) - \mathbb{E}[\mathscr{L}(u)]$ and $\Pi_{\tilde{\mathscr{Y}}}^{\perp}$ is the projection operator from the space $L^2(\Omega)$ to the orthogonal complement of the S dimension subspace $\tilde{\mathscr{Y}} = \mathrm{span}(\tilde{Y}_1,...,\tilde{Y}_S)$. The related initial and boundary conditions can be obtained by the same way as in (3.15) and (3.16). The decomposition (3.17) and the corresponding system (3.21)-(3.24) have been proposed in [66, 68] for a dynamically low rank approximation of a time dependent differential matrix/tensor equation. An analogous formulation in infinite dimensional setting is derived in [67, 39], related to the multi-configuration time-dependent Hartree approach (MCTDH), in the quantum dynamics framework. We remark that for time dependent SPDEs, a Dynamically bi-orthogonal method (DyBO), which has a close relation with the DDO approximation, has been introduced in [30, 31]. As our error analysis relies on the symmetric property of the DDO approach, we will show in the following the equivalence between the DDO and the DO approximations. Note that in the DDO system (3.21)-(3.24), the equation for the mean function coincides with equation (3.12) in the DO system. Furthermore, letting $\mathbf{Y} = \mathbf{A}\tilde{\mathbf{Y}}$, it is easy to show that the approximation $\tilde{u}_S = \bar{u}_S + \tilde{\mathbf{U}}^T\mathbf{Y}$ satisfies the DO system. Indeed, by using together equations (3.22) and (3.24), we have

$$
\begin{aligned}
\frac{\partial \mathbf{Y}}{\partial t} &= \frac{\mathrm{d}\mathbf{A}}{\mathrm{dt}}\tilde{\mathbf{Y}} + \mathbf{A}\frac{\partial \tilde{\mathbf{Y}}}{\partial t}\\
&= \mathbb{E}\left[\langle \mathscr{L}^*(\tilde{u}_S), \tilde{\mathbf{U}}^T\rangle \tilde{\mathbf{Y}}^T\right]\tilde{\mathbf{Y}} + \Pi_{\tilde{\mathscr{Y}}}^{\perp}\langle \mathscr{L}^*(\tilde{u}_S), \tilde{\mathbf{U}}^T\rangle\\
&= \mathbb{E}\left[\langle \mathscr{L}^*(\tilde{u}_S), \tilde{\mathbf{U}}^T\rangle \tilde{\mathbf{Y}}^T\right]\tilde{\mathbf{Y}} + \langle \mathscr{L}^*(\tilde{u}_S), \tilde{\mathbf{U}}^T\rangle - \mathbb{E}\left[\langle \mathscr{L}^*(\tilde{u}_S), \tilde{\mathbf{U}}^T\rangle \tilde{\mathbf{Y}}^T\right]\tilde{\mathbf{Y}}\\
&= \langle \mathscr{L}^*(\tilde{u}_S), \tilde{\mathbf{U}}^T\rangle,
\end{aligned}
\tag{3.25}
$$

which coincides with equation (3.14) in the DO system. Moreover, by multiplying both sides of (3.23) by $\mathbf{A}$ we obtain

$$
\mathbf{A}\mathbf{A}^T\frac{\partial \tilde{\mathbf{U}}}{\partial t} = \Pi_{\tilde{\mathscr{U}}}^{\perp}\mathbb{E}\left[\mathbf{Y}\mathscr{L}^*(\tilde{u}_S)\right].
\tag{3.26}
$$

Note that the covariance matrix of $\mathbf{Y}$ is

$$\mathbf{C} = \mathbb{E}[\mathbf{Y}\mathbf{Y}^T] = \mathbb{E}[\mathbf{A}\tilde{\mathbf{Y}}(\mathbf{A}\tilde{\mathbf{Y}})^T] = \mathbf{A}\mathbf{A}^T. \tag{3.27}$$

Thus, the equation (3.26) coincides with (3.13) in the DO system. Using similar techniques, one can show that the corresponding initial and boundary conditions for the DO system and the DDO system also coincide. On the other hand, if $u_S = \bar{u}_S + \mathbf{U}^T\mathbf{Y}$ is a solution of the DO system (3.12)-(3.14), then defining $\mathbf{A}$ as the square root of $\mathbf{C}$ one can show by the same arguments as above that $u_S = \bar{u}_S + \mathbf{U}^T\mathbf{A}\tilde{\mathbf{Y}}$ with $\tilde{\mathbf{Y}} = \mathbf{A}^{-1}\mathbf{Y}$ is a solution of the DDO system (3.21)-(3.24). In particular, $\tilde{\mathbf{Y}}$ is a vector of orthonormal random variables in $L^2(\Omega)$. We thus conclude that the DO and the DDO formulations produce the same approximate solution.

3.2.2 An equivalent Variational Formulation

Let $\mathcal{H} \subset L^2(D)$ be a suitable Hilbert space and $\mathcal{H}'$ its dual. We assume that equation (3.3) can be set in $\mathcal{H}' \otimes L^2(\Omega)$ and it admits a unique solution $u(t) \in \mathcal{H} \otimes L^2(\Omega)$ for any $t \in \mathcal{T}$. Denoted by $L_0^2(\Omega) \subset L^2(\Omega)$ the subspace of all the square integrable random variables with zero mean, let us define:

$$\tilde{\mathscr{F}}_S = \mathcal{O}_{[\mathcal{H}]^S} \times M_{S \times S} \times \mathcal{O}_{[L_0^2(\Omega)]^S} \tag{3.28}$$

with:

$$M_{S \times S} = \{\mathbf{A} \in \mathbb{R}^{S \times S} : \text{rank}(\mathbf{A}) = S\},$$
$$\mathcal{O}_{[\mathcal{H}]^S} = \{\tilde{\mathbf{U}} = (\tilde{U}_1, ..., \tilde{U}_S) : \tilde{U}_i \in \mathcal{H} \text{ and } <\tilde{U}_i, \tilde{U}_j> = \delta_{ij} \; \forall i, j = 1, ..., S\},$$
$$\mathcal{O}_{[L_0^2(\Omega)]^S} = \{\tilde{\mathbf{Y}} = (\tilde{Y}_1, ..., \tilde{Y}_S) : \tilde{Y}_i \in L_0^2(\Omega) \text{ and } \mathbb{E}[\tilde{Y}_i\tilde{Y}_j] = \delta_{ij} \; \forall i, j = 1, ..., S\}$$

and the map:

$$\begin{array}{rll} \pi : & \tilde{\mathscr{F}}_S & \to \mathcal{M}_S \\ & (\tilde{\mathbf{U}}, \mathbf{A}, \tilde{\mathbf{Y}}) & \mapsto \pi(\tilde{\mathbf{U}}, \mathbf{A}, \tilde{\mathbf{Y}}) = \sum_{j=1}^S \sum_{i=1}^S \mathbf{A}_{ij} \tilde{U}_i \tilde{Y}_j \end{array} \tag{3.29}$$

The image of π is the manifold:

$$\mathcal{M}_S = \left\{ u_S^* = \sum_{j=1}^S \sum_{i=1}^S \mathbf{A}_{ij} \tilde{U}_i \tilde{Y}_j : (\tilde{\mathbf{U}}, \mathbf{A}, \tilde{\mathbf{Y}}) \in \tilde{\mathscr{F}}_S \right\} \subset \mathcal{H} \otimes L_0^2(\Omega) \tag{3.30}$$

Observe that the subspace $\tilde{\mathscr{F}}_S$ is isomorphic to

$$\mathscr{F}_S = \mathcal{O}_{[\mathcal{H}]^S} \times B(S, L_0^2(\Omega)), \tag{3.31}$$

with $B(S, L_0^2(\Omega)) = \{\mathbf{Y} = (Y_1, ..., Y_S) : Y_i \in L_0^2(\Omega) \text{ and } \mathbb{E}[Y_iY_j] = \lambda_i\delta_{ij} \; \forall i, j = 1, ..., S\}$, i.e. the set of all the pairs $(\mathbf{U}, \mathbf{Y})$ such that $\mathbf{U} \in \mathcal{O}_{[\mathcal{H}]^S}$ and $\mathbf{Y} = (Y_1, ... Y_S)$ is a vector of, not necessary independent, zero mean, square integrable random variables with full rank covariance matrix

$\mathbf{C} = \mathbb{E}[\mathbf{Y}^T \mathbf{Y}]$. Therefore the manifold $\mathcal{M}_S$ can be equivalently defined as

$$\mathcal{M}_S = \left\{ u_S^* = \sum_{i=1}^S U_i Y_i : (\mathbf{U}, \mathbf{Y}) \in \mathcal{F}_S \right\} \tag{3.32}$$

with the associated map $\pi' : \mathcal{F}_S \to \mathcal{M}_S$ such that $\pi'(\mathbf{U}, \mathbf{Y}) = \sum_{i=1}^S U_i Y_i$.

Definition 3.2.1. *We define a S-rank random field as a function $u_S = \bar{u}_S + u_S^* \in \mathcal{H} \otimes L^2(\Omega)$ such that $u_S^* \in \mathcal{M}_S$, and we call $\mathcal{M}_{\mathcal{S}}(\mathcal{H} \otimes L_0^2(\Omega))$, or simply $\mathcal{M}_{\mathcal{S}}$ if no ambiguity arises on the functional spaces, the manifold of all S-rank zero mean random fields $u_S^* \in \mathcal{H} \otimes L_0^2(\Omega)$.*

The DO approximate solution is a S rank function at each time, provided that the covariance matrix is not singular. However, observe that the map π, analogously π', is not injective, i.e. the representation of a stochastic field $u_S^* \in \mathcal{M}_S$ in $\tilde{\mathcal{F}}_S$ (respectively $\mathcal{F}_S$) is not unique: for any orthogonal matrices $\Theta, \tilde{\Theta} \in M_{S \times S}$ we have that $\pi(\tilde{\mathbf{U}}, \mathbf{A}, \tilde{\mathbf{Y}}) = \pi(\tilde{\mathbf{U}}\Theta, \Theta^T \mathbf{A}\tilde{\Theta}, \tilde{\mathbf{Y}}\tilde{\Theta})$ and conversely for any $(\tilde{\mathbf{U}}, \mathbf{A}, \tilde{\mathbf{Y}})$ and $(\tilde{\mathbf{V}}, \mathbf{B}, \tilde{\mathbf{Z}})$ such that $\pi(\tilde{\mathbf{U}}, \mathbf{A}, \tilde{\mathbf{Y}}) = \pi(\tilde{\mathbf{V}}, \mathbf{B}, \tilde{\mathbf{Z}})$, there exists a unique orthogonal matrix $\Theta \in M_{S \times S}$ such that $(\tilde{\mathbf{V}}, \mathbf{B}, \tilde{\mathbf{Z}}) = (\tilde{\mathbf{U}}\Theta, \Theta^T \mathbf{A}\tilde{\Theta}, \tilde{\mathbf{Y}}\tilde{\Theta})$. In terms of differential geometry $\tilde{\mathcal{F}}_S$, $\mathcal{M}_S$ and π define a fiber bundle with fiber given by the group $\mathcal{O}_S$ of the orthogonal matrix of dimension S. In particular $\tilde{\mathcal{F}}_S / \mathcal{O}_S$ is isomorphic to $\mathcal{M}_S$. This construction allows us to equip $\mathcal{M}_S$ with a manifold structure and define the tangent space to $\mathcal{M}_S$ at $u_S^* = \pi(\tilde{\mathbf{U}}, \mathbf{A}, \tilde{\mathbf{Y}}) = \pi'(\mathbf{U}, \mathbf{Y})$.

Proposition 3.2.1. *The tangent space $\mathcal{T}_{u_S^*} \mathcal{M}_{\mathcal{S}}$ consists of the elements $\delta u_S^* \in \mathcal{H} \otimes L^2(\Omega)$ of the form:*

$$
\begin{aligned}
\mathcal{T}_{u_S^*} \mathcal{M}_{\mathcal{S}} \quad &= \Big\{ \delta u = \quad \sum_{i=1}^S \sum_{j=1}^S \big(\delta \mathbf{A}_{ij} \tilde{U}_i \tilde{Y}_j + \mathbf{A}_{ij} \delta \tilde{U}_i \tilde{Y}_j + \mathbf{A}_{ij} \tilde{U}_i \delta \tilde{Y}_j \big) \in \mathcal{H} \otimes L_0^2(\Omega) \\
&\qquad\qquad :< \delta \tilde{U}_i, \tilde{U}_j >= 0, \mathbb{E}[\delta \tilde{Y}_i] = 0, \mathbb{E}[\delta \tilde{Y}_i \tilde{Y}_j] = 0 \; \forall i, j = 1, ..., S \Big\} \\
&= \Big\{ \delta u = \quad \sum_{i=1}^S \big(\delta Y_i \tilde{U}_i + \delta \tilde{U}_i Y_i \big) :< \delta \tilde{U}_i, \tilde{U}_j >= 0 \, \mathbb{E}[\delta Y_i] = 0, \; \forall i, j = 1, ..., S \Big\}
\end{aligned} \tag{3.33}
$$

where in the last line we use notations in (3.32) with $\mathbf{Y} = \mathbf{A}\tilde{\mathbf{Y}}$. (The two different notations correspond to the DDO and the DO formulation respectively).

Observe that the tangent space does not depend on the choice of coordinates $(\tilde{\mathbf{U}}, \mathbf{A}, \tilde{\mathbf{Y}})$ but only on the point u_S^*.

Remark 3.2.1. *The DO approximate solution describes a path from $\mathcal{T}$ to $\mathcal{M}_S$ defined as $t \to u_S^*(t) = u_S(t) - \bar{u}_S(t)$. On the other hand the same path can be parametrized by infinitely many different continuous flows $t \to (\tilde{\mathbf{U}}(t), \mathbf{A}(t), \tilde{\mathbf{Y}}(t))$ foliating the fibers $\tilde{\mathcal{F}}_S$, satisfying $\pi(\tilde{\mathbf{U}}(t), \mathbf{A}(t), \tilde{\mathbf{Y}}(t)) = u_S^*(t)$ at each $t \in \mathcal{T}$, and so-called gauge transformations [7, 80] allow to continuously pass from one flow to an equivalent one. A generic gauge constrain is defined as follows:*

$$
\begin{aligned}
&< \dot{\tilde{U}}_i, \tilde{U}_j >=< \mathcal{G}_U \tilde{U}_i, \tilde{U}_j > \\
&\mathbb{E}[\dot{\tilde{Y}}_i \tilde{Y}_j] = \mathbb{E}[\mathcal{G}_Y \tilde{Y}_i \tilde{Y}_j]
\end{aligned} \tag{3.34}
$$

where $\mathcal{G}_U$ and $\mathcal{G}_Y$ are arbitrary self-adjoint operators in $L^2(D)$ and $L^2(\Omega)$ respectively. Any of these gauge constrains leads to an evolution system for one of the possible parametrization $(\tilde{\mathbf{U}}(t), \mathbf{A}(t), \tilde{\mathbf{Y}}(t)) \in \tilde{\mathcal{F}}_S$ of the same path $t \to u_S^*(t) \in \mathcal{M}_S$. Specifically, in the DDO formulation, we assume $\mathcal{G}_U \equiv 0$ and $\mathcal{G}_Y \equiv 0$, which leads to system (3.22)-(3.24). On the contrary, the gauge constrain that keeps $\mathbf{A}(t)$ diagonal at all times leads to the DyBO method proposed in [30, 31]. As shown in [35], the DO and the DyBO method provide, indeed, the same approximate solution, as long as the covariance matrix is not singular. In other words the two approaches lead to two different parameterizations in $\tilde{\mathcal{F}}_S$ of the same path $u_S^*(t)$ in $\mathcal{M}_S$. In fact, the analogy between the two methods can be alternatively proved by following the analysis on [7], properly adopted to our context.

According to the DO approach, the tangent space to $\mathcal{M}_S$, defined in (3.33), is parametrized by imposing condition in (3.11). Then the following proposition holds, suitably adapted to our framework from [66, 39].

Proposition 3.2.2. *For all $v \in \mathcal{H} \otimes L^2(\Omega)$ and $u_S^* \in \mathcal{M}_{\mathcal{S}}$, the orthogonal projection $P_{u_S^*}$ onto the tangent space $\mathcal{T}_{u_S^*} \mathcal{M}_{\mathcal{S}}$ of v is given by*

$$P_{u_S^*}(v) = P_{u_S^*}(v^*) = \mathbf{U}^T \langle v^*, \mathbf{U}^T \rangle + (\Pi_{\tilde{\mathcal{U}}}^\perp \{\mathbb{E}[v^* \mathbf{Y}^T]\} \mathbf{C}^{-1})^T \mathbf{Y},$$

where $\mathbf{C}^{-1}$ is the inverse of the covariance matrix $\mathbf{C} = \mathbb{E}[\mathbf{Y} \mathbf{Y}^T]$, that has full rank, by definition of S-rank function. We denote with $P_{u_S^*}^\perp v = P_{u_S^*}^\perp v^* = (\mathbb{I} - P_{u_S^*})(v^*)$ the complementary projection.

Furthermore we observe that the governing equation (3.3) can be formulated as:

$$
\begin{aligned}
\frac{\partial \bar{u}(x,t)}{\partial t} &= \mathbb{E}[\mathcal{L}(u(x,t,\cdot))], &&\text{in} \quad \mathcal{H}' \\
\frac{\partial u^*(x,t,\omega)}{\partial t} &= \mathcal{L}^*(u(x,t,\omega),\omega) &&\text{in} \quad \mathcal{H}' \otimes L^2(\Omega)
\end{aligned}
\tag{3.35}
$$

with $u^* = u - \bar{u}$. Finally we have the following variational formulation for the DO approach:

Proposition 3.2.3. *Let $(\bar{u}_S(t), \mathbf{Y}(t), \mathbf{U}(t)) \in \mathcal{H} \times \mathcal{F}_S$ be the strong solution of system (3.12)-(3.14) at each $t \in \mathcal{T}$, then $u_S(t) = \bar{u}_S(t) + u_S^*(t) = \bar{u}_S(t) + \pi'(\mathbf{U}(t), \mathbf{Y}(t)) \in \mathcal{H} \times \mathcal{M}_S$ satisfies:*

$$\mathbb{E}\left[\langle \frac{\partial u_S(\cdot,t,\cdot)}{\partial t} - \mathcal{L}(u_S(\cdot,t,\cdot)), v \rangle \right] = 0, \quad \forall v = \bar{v} + v^*, \ (\bar{v}, v^*) \in \mathcal{H} \times \mathcal{T}_{u_S^*(t)} \mathcal{M}_{\mathcal{S}} \tag{3.36}$$

at each $t \in \mathcal{T}$, which can be equivalently written as

$$\frac{\partial u_S(x,t,\omega)}{\partial t} = \mathbb{E}[\mathcal{L}(u_S(x,t,\cdot))] + P_{u_S^*(t)}(\mathcal{L}^*(u_S(x,t,\omega)))$$

with $\mathcal{L}^*(u_S) = \mathcal{L}(u_S) - \mathbb{E}[\mathcal{L}(u_S)]$.

Remark 3.2.2. *Observe that the tangent space $\mathcal{T}_{u_S^*(t)} \mathcal{M}_{\mathcal{S}}$ is time dependent and depends only*

on $u_S(t)$ and not on the parametrization. It follows that the variational formulation in (3.36) is valid for the DDO as well as DyBO approach proposed in [30, 31].

Proposition (3.2.3) emphasizes that the approximate solution $u_S^* = u_S - \bar{u}_S$ is forced to belong to the S dimensional manifold $\mathcal{M}_{\mathcal{S}}$ at all times. We point out that the DO solution (3.36) does not coincide, in general, with the best S-rank approximation (denoted by z_S in (3.9)) which instead minimizes the approximation error in L^2 sense at each time instant, i.e.

$$z_S(t) = \bar{u}(\cdot, t) + \mathrm{argmin}_{w \in \mathcal{M}_{\mathcal{S}}} \quad \mathbb{E}\left[\|u^*(\cdot, t, \cdot) - w(\cdot, \omega)\|_{L^2(D)}^2 \right], \ \forall t \in \mathcal{T}. \tag{3.37}$$

It is well known that the best S-rank approximation corresponds indeed to the truncated Karhunen-Loève expansion, with S terms. Observe that in the best S-rank approximation (3.37) the solution u^* of the equation (3.35) is projected onto the manifold $\mathcal{M}_S$, whereas in (3.36) the residual of the equation (3.35) is projected onto the tangent space $\mathcal{T}_{u_S^*(t)}\mathcal{M}_{\mathcal{S}}$. However, the DO formulation takes inspiration from the Karhunen-Loève decomposition. It aims at developing an analogous type of approximation without directly computing the Karhunen-Loève decomposition. In fact the DO method evolves a dynamically low rank approximation and adapts at each time instant the spatial basis as well the stochastic variables to what best describes the structure of the solution. This makes the method numerically accessible and effective in terms of approximation error at any time instant for long time integration.

3.2.3 Properties of the manifold

In this subsection, we shall discuss some properties of the manifold $\mathcal{M}_{\mathcal{S}}$, which will play an important role in the next section when analyzing the convergence properties of the DO approach. Given the equivalence between the DO and DDO formulations, shown in Section 3.1 here we will use either formalism depending on what is more convenient for the presentation.

Definition 3.2.2. *Denoted with $\mathbf{A}$ the square root of the covariance matrix $\mathbf{C} = \mathbb{E}[\mathbf{YY}^T]$, the singular values of a S-rank function $u_S = \bar{u}_S + \mathbf{U}^T\mathbf{Y}$ are defined as the singular values of $\mathbf{A}$:*

$$\sigma(u_S) := \sigma(\mathbf{A}) = \sqrt{eig(\mathbf{C})}. \tag{3.38}$$

Equivalently for the DDO formulation, $u_S = \bar{u}_S + \mathbf{U}^T\mathbf{A}\tilde{\mathbf{Y}}$, the singular values of u_S are by definition the singular values of $\mathbf{A}$.

In the following, we denote with $\|\cdot\|_0 := \|\cdot\|_{L^2(D) \otimes L^2(\Omega)}$ the norm in $L^2(D) \otimes L^2(\Omega)$. The norm for a function vector $\mathbf{U}$ is defined as usual, namely, $\|\mathbf{U}\|_0 = (\sum_i \|U_i\|_{L^2(D)}^2)^{1/2}$. We also denote with $\|\cdot\|_{\mathcal{F}}$ and $\|\cdot\|_2$ the Frobenius and the spectral norm of a matrix, respectively. Note that, with such definition we have $\|u_S^*\|_0 = \|\mathbf{A}\|_{\mathcal{F}}$, for $u_S^* = \tilde{\mathbf{U}}^T\mathbf{A}\tilde{\mathbf{Y}} \in \mathcal{M}_S$.

We introduce now a useful lemma concerning the properties of the operator $P_{u_S^*}$ and the curvature estimates for the manifold $\mathcal{M}_S$. This lemma is taken from [39] with just small

adjustments to the notations and settings used here. We skip the proof as it would follow very closely the one in [39]. Analogous results are achieved in [66], where the authors considered a very similar approach for matrix equations in finite dimensional spaces.

Lemma 3.2.1. *Consider the manifold* $\mathscr{M}_S(L^2(D) \otimes L^2(\Omega))$. *Let* $u_S^* = \mathbf{U}^T \mathbf{Y} \in \mathscr{M}_S$ *such that the smallest nonzero singular value satisfies* $\sigma_s(u_S^*) \geq \rho > 0$, *and let* $v_S^* = \mathbf{V}^T \mathbf{Z} \in \mathscr{M}_S$ *with* $\|u_S^* - v_S^*\|_0 \leq \frac{1}{8}\rho$. *Then,* $\forall w \in L^2(D) \otimes L^2(\Omega)$, *the following bounds hold:*

$$\|(P_{u_S^*} - P_{v_S^*})w\|_0 \leq 8\rho^{-1} \|u_S^* - v_S^*\|_0 \cdot \|w\|_0, \tag{3.39}$$

$$\|P_{u_S^*}^{\perp}(u_S^* - v_S^*)\|_0 \leq 4\rho^{-1} \|u_S^* - v_S^*\|_0^2. \tag{3.40}$$

Further we observe that any linear deterministic bounded operator applied to a S-rank function, does not increases its rank.

Proposition 3.2.4. *Let* $\mathcal{V}_1$ *and* $\mathcal{V}_2$ *be two Hilbert spaces such that* $\mathcal{V}_2 \subseteq \mathcal{V}_1 \subseteq L^2(D)$ *and* $\mathscr{B} : \mathcal{V}_1 \to \mathcal{V}_2$ *a linear bounded operator. For any* $u_S = \bar{u}_S + u_S^*$ *with* $(\bar{u}_S, u_S^*) \in \mathcal{V}_1 \times \mathscr{M}_S(\mathcal{V}_1 \otimes L^2(\Omega))$, *we have that* $\mathscr{B} \otimes \mathbb{I} u_S = \mathscr{B}\bar{u}_S + \mathscr{B} \otimes \mathbb{I} u_S^*$ *with* $(\mathscr{B}\bar{u}_S, \mathscr{B} \otimes \mathbb{I} u_S^*) \in \mathcal{V}_2 \times \mathscr{M}_S(\mathcal{V}_2 \otimes L^2(\Omega))$.

Proof. It is enough to observe that $(\mathscr{B} \otimes \mathbb{I}) u_S^* = \sum_{i=1}^S (\mathscr{B}U_i) Y_i$ and it can be expanded as:

$$\begin{aligned}
(\mathscr{B} \otimes \mathbb{I}) u_S^* &= \sum_{i=1}^S \Pi_{\mathbf{U}}(\mathscr{B}U_i) Y_i + \sum_{i=1}^S \Pi_{\mathscr{U}}^{\perp}(\mathscr{B}U_i) Y_i \\
&\sum_{i=1}^S \left(\sum_{j=1}^S \langle \mathscr{B}U_i, U_j \rangle U_j \right) Y_i + \sum_{i=1}^S \Pi_{\mathscr{U}}^{\perp}(\mathscr{B}U_i) Y_i \\
&\sum_{j=1}^S \left(\sum_{i=1}^S \langle \mathscr{B}U_i, U_j \rangle Y_i \right) U_j + \sum_{i=1}^S \Pi_{\mathscr{U}}^{\perp}(\mathscr{B}U_i) Y_i = \mathbf{U}^T \delta\mathbf{Y} + \delta\mathbf{U}^T \mathbf{Y}
\end{aligned} \tag{3.41}$$

where $\delta\mathbf{Y} = \langle \mathscr{B}(\mathbf{U}^T)\mathbf{Y}, \mathbf{U} \rangle$ and $\delta\mathbf{U} = \Pi_{\mathscr{U}}^{\perp}(\mathscr{B}\mathbf{U})$ is orthogonal to $\mathbf{U}$ by construction. $\square$

3.3 Application to stochastic parabolic equations

In this section, we consider the DO approach for the following linear stochastic parabolic equation:

$$\begin{cases}
\dfrac{\partial u(x,t,\omega)}{\partial t} - \nabla \cdot (a(x,\omega)\nabla u(x,t,\omega)) = f(x,t,\omega), & x \in D, t \in \mathcal{T}, \omega \in \Omega, \quad (3.42) \\
u(\sigma,t,\omega) = 0 & \sigma \in \partial D, t \in \mathcal{T}, \omega \in \Omega, \quad (3.43) \\
u(x,0,\omega) = u_0(x,\omega), & x \in D, \omega \in \Omega, \quad (3.44)
\end{cases}$$

where $a(x,\omega) : D \times \Omega \to \mathbb{R}$ is a random field with continuous and bounded covariance function and D is an open, bounded and Lipschitz domain. We say that u is a weak solution of problem (3.42)-(3.44) if it satisfies the initial condition $u = u_0$ at $t = 0$ and if, at any $t \in \mathcal{T}$,

$u(\cdot, t, \cdot) \in H_0^1(D) \otimes L^2(\Omega)$ and

$$\mathbb{E}[\langle \frac{\partial u(\cdot, t, \cdot)}{\partial t}, v \rangle] + \mathbb{E}[\langle a\nabla u(\cdot, t, \cdot), \nabla v \rangle] = \mathbb{E}[\langle f(\cdot, t, \cdot), v \rangle] \quad \forall v \in H_0^1(D) \otimes L^2(\Omega). \tag{3.45}$$

A sufficient condition to guarantee the existence and uniqueness of the solution u consists in assuming that $f \in L^2(\mathcal{T}, L^2(D) \otimes L^2(\Omega))$, $u_0 \in L^2(D) \otimes L^2(\Omega)$ and the diffusion coefficient $a(x, \omega)$ is bounded and uniformly coercive almost surely, i.e.

$$\exists a_{min}, a_{max} \in (0, +\infty) : P\big(\omega \in \Omega : a(x, \omega) \in [a_{min}, a_{max}], \forall x \in \overline{D}\big) = 1. \tag{3.46}$$

Then by standard arguments applied for almost every $\omega \in \Omega$ (see also [100]), it is straightforward to show that there exists a unique solution $u \in L^2(\mathcal{T}, H_0^1(D) \otimes L^2(\Omega))$ with $\frac{\partial u}{\partial t} \in L^2(\mathcal{T}, H^{-1}(D) \otimes L^2(\Omega))$ and by standard energy estimates the following a priori bound holds $\forall T \in \mathcal{T}$:

$$\|u(T)\|_{L^2(D) \otimes L^2(\Omega)}^2 + a_{min}\|u\|_{L^2(\mathcal{T}, H_0^1(D) \otimes L^2(\Omega))}^2 \leq \quad \|u_0\|_{L^2(D) \otimes L^2(\Omega)}^2 \\ + \frac{c_p^2}{a_{min}}\|f\|_{L^2(\mathcal{T}, L^2(D) \otimes L^2(\Omega))}^2, \tag{3.47}$$

where c_p denotes the constant appearing in the Poincaré inequality. For the error analysis of the DO method that will be presented in the next section, we need some extra regularity on the exact solution u as well as its DO approximation u_S. We make the following assumption: (For simplicity of notation we denote with $\dot{u}$ the time derivative of u)

Assumption 1.

- $u, u_S \in L^2(\mathcal{T}, H^2(D) \cap H_0^1(D) \otimes L^2(\Omega))$

- $\dot{u}, \dot{u}_S \in L^2(\mathcal{T}, L^2(D) \otimes L^2(\Omega))$

We give here an informal discussion on why this assumption is reasonable under mild extra requirements on the data of the problem (3.42)-(3.44). In particular, while regularity results on the exact solution u can be proved by standard techniques, it is not obvious whether analogous results should hold for the DO solution u_S, because of the projection on the tangent manifold. Consider the pure Neumann problem $\partial u_n = 0$ on $\partial\Omega$ and look first at the exact problem (3.42)-(3.44) (with Neumann boundary conditions instead of Dirichlet ones). Under the assumption that $\nabla a \in L^\infty(D \times \Omega)$ and $\nabla u(0) \in L^2(D) \otimes L^2(\Omega)$, by taking $v = -\Delta u$ in (3.45) and integrating in time we get:

$$\|\nabla u(T)\|_{L^2(D) \otimes L^2(\Omega)}^2 + a_{min}\|\Delta u\|_{L^2(\mathcal{T}, L^2(D) \otimes L^2(\Omega))}^2 \\ \leq \frac{2}{a_{min}}\|\nabla a\|_{L^\infty(D \times \Omega)}\|\nabla u\|_{L^2(\mathcal{T}, L^2(D) \otimes L^2(\Omega))}^2 + \frac{2}{a_{min}}\|f\|_{L^2(\mathcal{T}, L^2(D) \otimes L^2(\Omega))}^2 \\ + \|\nabla u(0)\|_{L^2(D) \otimes L^2(\Omega)}^2, \tag{3.48}$$

which implies, in light of (3.47), that u is bounded in $L^2(\mathcal{T}, H^2(D) \otimes L^2(\Omega))$. In order to derive a bound on the time derivative of u, let us now take $v = \dot{u}$ in (3.45) and integrate in time. We

get the following a priori estimate:

$$\|\dot{u}\|^2_{L^2(\mathcal{T},L^2(D)\otimes L^2(\Omega))} + a_{min}\|\nabla u(T)\|^2_{L^2(D)\otimes L^2(\Omega)} \leq \quad a_{max}\|\nabla u_0\|^2_{L^2(D)\otimes L^2(\Omega)} \\ + \|f\|^2_{L^2(\mathcal{T},L^2(D)\otimes L^2(\Omega))} \tag{3.49}$$

which shows that $\dot{u} \in L^2(\mathcal{T}, L^2(D)\otimes L^2(\Omega))$. Therefore the regularity properties in Assumption 1 on u are sound provided that $\nabla a \in L^\infty(D \times \Omega)$ and $u(0) \in H^1(D)\otimes L^2(\Omega)$. Observe that, since the truncated Karhunen-Loève expansion inherits the spatial regularity of u [119], estimates (3.48) and (3.49) are valid for z_S as well, for any $S \in \mathbb{N}$. By following the same approach as before, we investigate now the regularity of the DO solution u_S. The weak formulation of the DO method reads: At each time $t \in \mathcal{T}$, find $u_S = \bar{u}_S + u_S^*$ with $(\bar{u}_S, u_S^*) \in H^1(D) \times \mathcal{M}_{\mathcal{S}}(H^1(D)\otimes L^2(\Omega))$ such that

$$\mathbb{E}[\langle \dot{u}_S(\cdot, t, \cdot), v \rangle] + \mathbb{E}[\langle a\nabla u_S(\cdot, t, \cdot), \nabla v \rangle] = \mathbb{E}[\langle f(\cdot, t, \cdot), v \rangle], \tag{3.50}$$

$\forall v = \bar{v} + v^*, (\bar{v}, v^*) \in H^1(D) \times \mathcal{T}_{u_S^*(t)}\mathcal{M}_{\mathcal{S}}$.

We now take as before $v = -\Delta u_S$ in (3.50). The key now is to observe that thanks to the Proposition 3.2.4, $v^* = -\Delta u_S^* \in \mathcal{T}_{u_S^*}\mathcal{M}_{\mathcal{S}}$ so that it is a suitable test function. By the same argument we can take $v = \dot{u}_S$ as a test function. Then, proceeding as before, one can derive the same bounds (3.48) and (3.49) for the DO solution as well. This shows that the regularity assumption (Assumption 1) are also sound for the DO solution u_S under the same conditions on the data: $\nabla a \in L^\infty(D \times \Omega)$ and $u(0) \in H^1(D)\otimes L^2(\Omega)$.

Remark 3.3.1. *The informal arguments that we have used to derive the bounds (3.48) and (3.49) for the exact solution as well as its DO approximation u_S can be made rigorous e.g. by using the so called Faedo-Galerkin method that consists on working with a sequence of Galerkin approximations of the solution u, which satisfy the governing equation projected in finite dimensional subspaces, and weakly converge to u (see e.g. [115, 45]).*

3.3.1 Analysis of DO approximation error

We are now ready to prove the convergence result for the DO approximation of the stochastic parabolic equation (3.42)-(3.44). The proof will follow closely the one by Lubich et al. in [66] for the error analysis of the Dynamical Low Rank approximation of time dependent data matrices. For notation simplicity, we denote

$$\mathcal{L}(u) := \nabla \cdot (a\nabla u) + f, \quad \mathcal{L}^*(\cdot) = \mathcal{L}(\cdot) - \mathbb{E}[\mathcal{L}(\cdot)]. \tag{3.51}$$

We suppose that the problem (3.42)-(3.44) admits a unique solution u in $L^2\big(\mathcal{T}, H^2(D) \cap H_0^1(D)\otimes L^2(\Omega)\big)$ and that there exists a continuously differentiable best S-rank approximation $z_S = \bar{z} + z_S^*$ of the solution u at any $t \in \mathcal{T}$. Observe that the covariance function Cov_u is equal to zero on the boundary and then each mode Z_i in (3.9) and, as a result, the truncate Karhunen-Loève expansion z_S, satisfy the homogeneous Dirichlet boundary conditions. The assumptions on the data can be summarized as follows.

Assumption 2. • $f \in L^2(\mathcal{T}, L^2(D) \otimes L^2(\Omega))$,

- $a(x,\omega)$ *bounded and uniformly coercive a. s.,*

- $\nabla a \in L^\infty(D \times \Omega)$,

- $u(0) \in H^1(D) \otimes L^2(\Omega)$.

In light of what discussed in the previous section we can argue that under Assumptions 2 the exact solution as well as the truncated Karhunen-Loève expansion z_S and the DO approximate solution u_S belong to $L^2\big(\mathcal{T}, H^2(D) \cap H_0^1(D) \otimes L^2(\Omega)\big)$ and, in particular, the quantities $\|\mathcal{L}(u)\|_0, \|\mathcal{L}(z_S)\|_0$ and $\|\mathcal{L}(u_S)\|_0$ will be bounded, which is a necessary condition for our proof of the quasi-optimality of the DO approximate solution. We will estimate the error of the DO approximate solution in terms of the best approximation error $\|u - z_S\|_{H^1(D) \otimes L^2(D)}$ as long as this remains small enough compared with the smallest singular value of z_S.

Theorem 3.3.1. *Suppose that, under Assumption 2, a continuously differentiable best S-rank approximation $z_S(t)$ of the exact solution $u(t)$ of (3.42)-(3.44) exists in $\big(H^2(D) \cap H_0^1(D)\big) \otimes L^2(\Omega)$ for $0 \le t \le \bar{t}$ and the smallest singular value of $z_S(t)$ is uniformly bounded from below, with lower bound $\sigma(z_S(t)) \ge \rho > 0, \forall t \in [0, \bar{t}]$. Then there exists $0 < \hat{t} \le \bar{t}$ such that the approximation error of the DO solution $u_S = \bar{u}_S + u_S^*$ with initial value $u_S(0) = z_S(0)$ is bounded by*

$$\|u_S(t) - z_S(t)\|_0^2 + a_{min} \int_0^t |u_S(\tau) - z_S(\tau)|_1^2 d\tau \le 2\alpha e^{2\beta(t)} \int_0^t \|z_S(\tau) - u(\tau)\|_1^2 d\tau, \qquad (3.52)$$

for all $0 < t \le \hat{t}$, with

$$\beta(t) = 4\rho^{-1} \int_0^t \big(4\|\mathcal{L}^*(z_S(\tau))\|_0 + \|\mathcal{L}^*(u(\tau))\|_0 + \|\mathcal{L}^*(u_S(\tau))\|_0 + \|\dot{z}_S^*(\tau)\|_0^2\big) d\tau,$$
$$\alpha = \max\Big\{\frac{a_{max}^2}{2a_{min}}, 4\rho^{-1}\Big\},$$

where $\|.\|_1, |.|_1$ denote respectively the norm and semi-norm in $H^1(D) \otimes L^2(\Omega)$, provided that all the terms in (3.52) are well defined.

Proof. Thanks to the assumptions of boundedness of $\dot{u}$ and $\dot{z}_S$ and being $u_S(0) = z_S(0)$, we have that for any $t \in [0, \bar{t}]$

$$\begin{aligned}\|u_S(t) - z_S(t)\|_0^2 &= \|\int_0^t \big(\dot{u}_S(\tau) - \dot{z}_S(\tau)\big) d\tau\|_0^2 \le t \int_0^t \|\dot{u}_S(\tau) - \dot{z}_S(\tau)\|_0^2 d\tau \\ &\le 2t \underbrace{\big(\|\dot{u}_S\|_{L^2([0,t], L^2(D) \otimes L^2(\Omega))}^2 + \|\dot{z}_S\|_{L^2([0,t], L^2(D) \otimes L^2(\Omega))}^2\big)}_{A(t)}\end{aligned}$$

therefore, for $\hat{t} = \min\big(\bar{t}, \frac{\rho^2}{2 \cdot 8^2 A(\hat{t})}\big)$ the distance between u_S and z_S remains bounded by $\frac{1}{8}\rho$, as required in Lemma 3.2.1.

For the best approximation z_S it must hold that $\mathbb{E}[z_S] = \mathbb{E}[u]$ and $\mathbb{E}[\dot{z}_S] = \mathbb{E}[\mathcal{L}(u)]$. Moreover $(z_S - \mathbb{E}[z_S]) - (u - \mathbb{E}[u])$ must be orthogonal to the tangent space $\mathcal{T}_{z_S^*}\mathcal{M}_S$, that is:

$$P_{z_S^*}\big((z_S - \mathbb{E}[z_S]) - (u - \mathbb{E}[u])\big) = P_{z_S^*}(z_S - u) = 0 \qquad (3.53)$$

For $z_S^* \in \mathcal{M}_S$, we denote with $\mathbf{D}_{z_S^*} P[\delta z_S^*]$ the Gateaux derivative of the projection operator in z_S^*, i.e.

$$\mathbf{D}_{z_S^*} P[\delta z_S^*] = \lim_{\varepsilon \to 0} \frac{P_{z_S^* + \varepsilon \delta z_S^*} - P_{z_S^*}}{\varepsilon}. \tag{3.54}$$

Observe that $\frac{d}{dt} P_{z_S^*(t)} = \mathbf{D}_{z_S^*} P[\dot{z}_S^*]$. We differentiate the relation (3.53) with respect to t and we then obtain:

$$P_{z_S^*}(\dot{z}_S - \dot{u}) + \mathbf{D}_{z_S^*} P[\dot{z}_S^*](z_S - u) = 0,$$

Since we have $P_{z_S^*}(\dot{z}_S) = P_{z_S^*}(\dot{z}_S^*) = \dot{z}_S^* = \dot{z}_S - \mathbb{E}[\dot{z}_S]$ the above equation becomes

$$\begin{aligned}
\dot{z}_S &= \mathbb{E}[\dot{z}_S] + P_{z_S^*}(\dot{u}) - \mathbf{D}_{z_S^*} P[\dot{z}_S^*](z_S - u) \\
&= \mathbb{E}[\mathscr{L}(u)] + P_{z_S^*}(\mathscr{L}^*(u)) - \mathbf{D}_{z_S^*} P[\dot{z}_S^*](z_S - u)
\end{aligned} \tag{3.55}$$

Since the DO solution satisfies

$$\dot{u}_S = \mathbb{E}[\mathscr{L}(u_S)] + P_{u_S^*}(\mathscr{L}^*(u_S)), \tag{3.56}$$

by subtracting equations (3.55) and (3.56) we get

$$\dot{u}_S - \dot{z}_S = \mathbb{E}[\mathscr{L}(u_S)] - \mathbb{E}[\mathscr{L}(u)] + P_{u_S^*}(\mathscr{L}^*(u_S)) - P_{z_S^*}(\mathscr{L}^*(u)) + \mathbf{D}_{z_S^*} P[\dot{z}_S^*](z_S - u).$$

By adding and subtracting $(P_{u_S^*} - P_{z_S^*})(\mathscr{L}^*(z_S))$ we obtain

$$\begin{aligned}
\dot{u}_S - \dot{z}_S &= \mathbb{E}[\mathscr{L}(u_S)] - \mathbb{E}[\mathscr{L}(u)] + (P_{u_S^*} - P_{z_S^*})(\mathscr{L}^*(z_S)) + P_{z_S^*}(\mathscr{L}^*(z_S) - \mathscr{L}^*(u)) \\
&\quad + [I - P_{u_S^*}^\perp](\mathscr{L}^*(u_S) - \mathscr{L}^*(z_S)) + \mathbf{D}_{z_S^*} P[\dot{z}_S^*](z_S - u).
\end{aligned}$$

and then

$$\begin{aligned}
\dot{u}_S - \dot{z}_S &= (P_{u_S^*} - P_{z_S^*})(\mathscr{L}^*(z_S)) + P_{z_S^*}(\mathscr{L}^*(z_S) - \mathscr{L}^*(u)) + \mathbf{D}_{z_S^*} P[\dot{z}_S^*](z_S - u) \\
&\quad + (\mathscr{L}(u_S) - \mathscr{L}(z_S)) - P_{u_S^*}^\perp(\mathscr{L}^*(u_S) - \mathscr{L}^*(z_S)) + \mathbb{E}[\mathscr{L}(z_S)] - \mathbb{E}[\mathscr{L}(u)].
\end{aligned}$$

By taking the inner product with $u_S - z_S$, on both sides, we obtain

$$\begin{aligned}
\mathbb{E}[\langle \dot{u}_S - \dot{z}_S, u_S - z_S \rangle] &= \overbrace{\mathbb{E}[\langle (P_{u_S^*} - P_{z_S^*})(\mathscr{L}^*(z_S)), u_S - z_S \rangle]}^{T_1} \\
&+ \overbrace{\mathbb{E}[\langle \mathscr{L}(u_S) - \mathscr{L}(z_S), u_S - z_S \rangle]}^{T_2} \\
&+ \overbrace{\mathbb{E}[\langle P_{z_S^*}(\mathscr{L}^*(z_S) - \mathscr{L}^*(u)), u_S - z_S \rangle] + \mathbb{E}[\langle \mathbb{E}[\mathscr{L}(z_S)] - \mathbb{E}[\mathscr{L}(u)], u_S - z_S \rangle]}^{T_3} \\
&+ \underbrace{\mathbb{E}[\langle \mathbf{D}_{z_S^*} P[\dot{z}_S^*](z_S - u), u_S - z_S \rangle]}_{T_4} + \underbrace{\mathbb{E}[\langle -P_{u_S^*}^\perp(\mathscr{L}^*(u_S) - \mathscr{L}^*(z_S)), u_S - z_S \rangle]}_{T_5}
\end{aligned} \tag{3.57}$$

We now estimate separately each term on the right hand side of (3.57). Lemma 3.2.1 implies

that:

$$T_1 : \mathbb{E}\big[\langle (P_{u_S^*} - P_{z_S^*})(\mathscr{L}^*(z_S)), u_S - z_S\rangle\big] \quad = \mathbb{E}\big[\langle \mathscr{L}^*(z_S), (P_{u_S^*} - P_{z_S^*})(u_S - z_S)\rangle\big]$$
$$\leq 8\rho^{-1}\|\mathscr{L}^*(z_S)\|_0 \|u_S - z_S\|_0^2 \tag{3.58}$$

$$T_2 : \mathbb{E}\big[\langle \mathscr{L}(u_S) - \mathscr{L}(z_S), u_S - z_S\rangle\big] \leq -a_{min}\|\nabla u_S - \nabla z_S\|_0^2 \leq -a_{min}|u_S - z_S|_1^2$$

For the term T_3, since

$$\mathbb{E}\big[\langle P_{z_S^*}\big(\mathscr{L}^*(z_S) - \mathscr{L}^*(u)\big), u_S - z_S\rangle\big] = \quad \mathbb{E}\big[\langle \mathscr{L}^*(z_S) - \mathscr{L}^*(u), u_S - z_S\rangle\big]$$
$$-\mathbb{E}\big[\langle \mathscr{L}^*(z_S) - \mathscr{L}^*(u), P_{z_S^*}^\perp(u_S - z_S)\rangle\big]$$

we have

$$T_3 : \mathbb{E}\big[\langle P_{z_S^*}\big(\mathscr{L}^*(z_S) - \mathscr{L}^*(u)\big), u_S - z_S\rangle\big] + \mathbb{E}\big[\langle \mathbb{E}[\mathscr{L}(z_S)] - \mathbb{E}[\mathscr{L}(u)], u_S - z_S\rangle\big]$$
$$T_3 = \mathbb{E}\big[\langle \mathscr{L}(z_S) - \mathscr{L}(u), u_S - z_S\rangle\big] - \mathbb{E}\big[\langle \mathscr{L}^*(z_S) - \mathscr{L}^*(u), P_{z_S^*}^\perp(u_S - z_S)\rangle\big]$$

and then

$$T_3 \leq a_{max}|z_S - u|_1|u_S - z_S|_1 + 4\rho^{-1}(\|\mathscr{L}^*(z_S)\|_0 + \|\mathscr{L}^*(u)\|_0)\|u_S - z_S\|_0^2$$

Analogously

$$T_5 : \mathbb{E}\big[\langle P_{u_S^*}^\perp\big(\mathscr{L}^*(u_S) - \mathscr{L}^*(z_S)\big), u_S - z_S\rangle\big] \quad \leq \|\mathscr{L}^*(u_S) - \mathscr{L}^*(z_S)\|_0\|P_{u_S^*}^\perp(u_S - z_S)\|_0$$
$$\leq 4\rho^{-1}(\|\mathscr{L}^*(z_S)\|_0 + \|\mathscr{L}^*(u_S)\|_0)\|u_S - z_S\|_0^2. \tag{3.59}$$

Also we have:

$$\|\mathbf{D}_{z_S^*}P[\dot{z}_S^*](z_S - u)\|_0 = \lim_{dt\to 0} \frac{P_{z_S^* + dt\dot{z}_S^*} - P_{z_S^*}}{dt}(z_S - u) \leq 8\rho^{-1}\|\dot{z}_S^*\|_0\|z_S - u\|_0,$$

and hence:

$$T_4 : \mathbb{E}\big[\langle \mathbf{D}_{z_S^*}P[\dot{z}_S^*](z_S - u), u_S - z_S\rangle\big] \leq 8\rho^{-1}\|\dot{z}_S^*\|_0\|z_S - u\|_0\|u_S - z_S\|_0. \tag{3.60}$$

Finally by combining (3.58)-(3.60) and denoting $\epsilon = u_S - z_S$, we obtain

$$\frac{1}{2}\frac{d}{dt}\|\epsilon\|_0^2 + \frac{1}{2}a_{min}|\epsilon|_1^2 \quad \leq \big\{16\rho^{-1}\|\mathscr{L}^*(z_S)\|_0 + 4\rho^{-1}\|\mathscr{L}^*(u)\|_0 + 4\rho^{-1}\|\mathscr{L}^*(u_S)\|_0$$
$$+ 4\rho^{-1}\|\dot{z}_S^*\|^2\big\}\|\epsilon\|_0^2 + \frac{a_{max}^2}{2a_{min}}|z_S - u|_1^2 \tag{3.61}$$
$$+ 4\rho^{-1}\|z_S - u\|_0^2$$

The result now follows using the Gronwall inequality. $\qquad\square$

Remark 3.3.2. *The derived error bound applies as well to the DDO and DyBO solutions as long as they remain full rank. Indeed, the proof is based on the variational formulation (3.36) and*

does not make use of the parametrization of the manifold.

Remark 3.3.3. *Improved upper bounds can be investigated under stronger assumptions as in [66]. Smaller errors over longer time intervals can be obtained if, not only the error $u - z_S$, but also its derivative is small.*

3.4 Deterministic equation with stochastic initial datum

To have a better understanding of the DO approximation, let us have now a closer look at the following simple problem

$$\begin{cases} \dot{u}(x,t,\omega) - \Delta u(x,t,\omega) = 0 & x \in D, \, t \in \mathcal{T}, \, \omega \in \Omega, & (3.62) \\ u(\sigma,t,\omega) = 0 & \sigma \in \partial D, \, t \in \mathcal{T}, \, \omega \in \Omega, & (3.63) \\ u(x,0,\omega) = u_0(x,\omega) & x \in D, \, \omega \in \Omega & (3.64) \end{cases}$$

For sake of simplicity we assume $\mathbb{E}[u_0] = 0$. However observe that in case of a deterministic linear operator the equation for the mean in the DO system (3.12)-(3.14) is completely decoupled from the others, which implies that nothing changes in the following analysis for any $\bar{u}_0 \neq 0$.

3.4.1 Case I: exactness of the DO approximation

We assume that the initial datum u_0 is in the manifold $\mathcal{M}_S$. According to the Karhunen-Loève decomposition, u_0 can be expanded as:

$$u_0(x,\omega) = \sum_{i=1}^{S} \sqrt{\mu_i} \gamma_i(\omega) Z_i(x) \tag{3.65}$$

Let $\{\lambda_i\}_{i=1}^{\infty}$ and $\{\Phi_i\}_{i=1}^{\infty}$ be respectively the eigenvalues and the eigenfunctions of the Laplace operator, then the exact solution of problem (3.62) with initial datum (3.65) is simply given by:

$$u(x,t,\omega) = \sum_{i=1}^{S} \sqrt{\mu_i} \gamma_i(\omega) \Big[\sum_{k=1}^{\infty} < Z_i, \Phi_k > e^{-\lambda_k t} \Phi_k(x) \Big]. \tag{3.66}$$

Observe that u is in the manifold $\mathcal{M}_S$ and the truncated Karhunen-Loève expansion of rank S is actually exact for all times. The exact solution belongs indeed to $\mathcal{M}_S$ at any time instant, hence it coincides with its best S-rank approximation z_S. We show here that the DO approximate solution is exact as well. First of all we have that the time derivative $\dot{u}$ is in manifold $\mathcal{M}_S$:

$$\dot{u}(x,t,\omega) = -\sum_{i=1}^{S} \sqrt{\mu_i} \gamma_i(\omega) \Big[\sum_{k=1}^{\infty} < Z_i, \Phi_k > \lambda_k e^{-\lambda_k t} \Phi_k(x) \Big].$$

69

Moreover we observe that, in light of Proposition 3.2.4, Δu belongs to $\mathcal{T}_u \mathcal{M}_S$ at each time instant, indeed:

$$\Delta u = \sum_{i=1}^{S} <\Delta u, U_i> U_i + \sum_{i=1}^{S} \mathbb{E}[\Pi_{\mathcal{U}}^{\perp}(\Delta u) Y_i] Y_i$$

which implies that the projection of Δu onto the tangent space $\mathcal{T}_u \mathcal{M}_S$ is actually equal to Δu itself. In particular, since the projection of the governing equation (3.62) onto the tangent space $\mathcal{T}_u \mathcal{M}_S$ coincides with the governing equation, we have that the DO solution u_S satisfies the exact equation (3.62). Finally the fact that $u_0 = z_S(0) = u_S(0)$ ensures that the three solutions coincide at each time. Formally the same conclusion can be achieved by looking at the evolution equations of z_S and u_S. As shown in (3.55) and (3.56) we have that:

$$\begin{aligned}
\dot{z}_S &= P_{z_S}(\mathcal{L}(u)) - \mathbf{D}_{z_S} P[\dot{z}_S](z_S - u) \\
\dot{u}_S &= P_{u_S}(\mathcal{L}(u_S))
\end{aligned}$$

with initial condition $u_0 = z_S(0) = u_S(0)$. Since $u(t) = z_S(t)$ at each time, the second term on the right side of the equation (3.4.1) is equal to zero, i.e.:

$$\begin{aligned}
\dot{z}_S &= P_{z_S}(\mathcal{L}(z_S)) = \mathcal{L}(z_S) \\
\dot{u}_S &= P_{u_S}(\mathcal{L}(u_S)) = \mathcal{L}(u_S).
\end{aligned}$$

The two functions satisfy the same evolution equation with equal initial condition which implies that they are equal at each times.

Remark 3.4.1. *More generally, if the differential operator $\mathcal{L}(\cdot)$ in (3.3) is a linear deterministic operator and the initial condition u_0 is in $\mathcal{M}_S$, then u belongs to $\mathcal{M}_S$ and the DO approximate solution (with rank equal to S) coincides with the exact solution at each time instant.*

Proposition 3.4.1. *If the initial condition $u_0 \in \mathcal{M}_S$ is a linear combination of S eigenfunctions $\Phi = (\Phi_1, ..., \Phi_S)^T$ of the Laplace operator, then the DO method coincides to the Proper Orthogonal Decomposition method (see e.g. [75] chapter 2) in which the governing equation is projected in the fixed (time independent) subspace spanned by Φ. Indeed the deterministic basis functions do not evolve in time and the DO solution u_S is given by $u_S(x, t, \omega) = \mathbf{U}^T(x, t)\mathbf{Y}(t, \omega)$ with:*

$$\mathbf{U}(x, t) = \mathbf{U}(x, 0), \quad \mathbf{Y}(t, \omega) = \Upsilon(0) e^{-\Lambda t} \Upsilon(0)^{\top} \mathbf{Y}(0, \omega), \tag{3.67}$$

where Λ is the diagonal matrix of the eigenvalues of the Laplace operator associated to Φ, i.e. $-\Delta\Phi = \Lambda\Phi$, and $\Upsilon(t)$ is the transformation matrix $\Upsilon(t)_{i,j} = \langle U_i(\cdot, t), \Phi_j \rangle$ between the basis of modes $\mathbf{U}$ and the basis of eigenfunctions Φ.

Proof. First of all we recall that, since $u_0 \in \mathcal{M}_S$, the exact solution is in $\mathcal{M}_S$. Moreover it is in the span of the S eigenfunctions $\Phi = (\Phi_1, ..., \Phi_S)^T$ at any time instant. Indeed, being $u_0(x, \omega) = \Phi(x)^T \Upsilon(0)\mathbf{Y}(0, \omega)$, the exact solution is given by:

$$u(x, t, \omega) = \Phi(x)^T e^{-\Lambda t} \Upsilon(0)^T \mathbf{Y}(0, \omega)$$

As previously discussed, we have that the DO solution coincides to the exact solution. Then it is easy to verify that $u = \mathbf{U}^T \mathbf{Y}$ and the couple $(\mathbf{U}, \mathbf{Y})$ in (3.67) satisfies the DO system. To this end, observe that the covariance matrix of the solution can be explicitly calculated as follows:

$$
\begin{aligned}
\mathbf{C}(t) \quad &= \Upsilon(0)e^{-\Lambda t}\Upsilon(0)^T \mathbb{E}[\mathbf{Y}(0)\mathbf{Y}(0)^T]\Upsilon(0)e^{-\Lambda t}\Upsilon(0)^T \\
&= \Upsilon(0)e^{-\Lambda t}\Upsilon(0)^T \mathbf{C}(0)\Upsilon(0)e^{-\Lambda t}\Upsilon(0)^T
\end{aligned}
\tag{3.68}
$$

The initial covariance matrix $\mathbf{C}(0)$ is assumed to have full rank and since $\Upsilon(0)e^{-\Lambda t}\Upsilon(0)^T$ is strictly positive definite, $\mathbf{C}(t)$ remains invertible at any $t \in \mathcal{T}$. This implies that $\mathbf{C}(t)$ can be simplified in (3.13) and then DO system is reduced to:

$$
\begin{cases}
\dfrac{d\Upsilon(t)}{dt} = \left[\Upsilon(t)\Lambda\,\Upsilon(t)^\top\Upsilon(t) - \Upsilon(t)\Lambda\right] = 0 \\[4mm]
\dfrac{\partial \mathbf{Y}(t,\omega)}{\partial t} = -\Upsilon(t)\Lambda\Upsilon(t)^\top\mathbf{Y}(t,\omega).
\end{cases}
\tag{3.69}
$$

where we use that $\Upsilon(t)$ is a square orthogonal matrix. By integrating in time we get the result. $\qquad\square$

3.4.2 Case II: effect of truncation - z_S is continuously time differentiable

We now consider an initial datum $u_0 \notin \mathcal{M}_S$. Assuming $u_0 \in L^2(D) \otimes L^2(\Omega)$ it can be expanded according to the Karhunen-Loève decomposition as:

$$
u_0(x,\omega) = \sum_{i=1}^{\infty} \sqrt{\mu_i}\gamma_i(\omega)Z_i(x)
\tag{3.70}
$$

Analogously the exact solution u of problem (3.62) with initial condition (3.70) can be in general decomposed at each time t as:

$$
u(x,t,\omega) = \sum_{i=1}^{\infty} \sqrt{\mu_i'(t)}\gamma_i'(\omega,t)Z_i'(x,t).
\tag{3.71}
$$

In order to apply the DO method, the initial datum is approximated by the first S terms of the series (3.70), whose sum $z_S(0)$ corresponds to the best rank-S approximation of u_0 in norm $L^2(D) \otimes L^2(\Omega)$ (S-rank truncated Karhunen-Loève expansion). In the same way, we denote with $z_S(t)$ the best rank-S approximation of the exact solution at time $t > 0$, i.e.:

$$
z_S(x,t,\omega) = \sum_{i=1}^{S} \sqrt{\mu_i'(t)}\gamma_i'(\omega,t)Z_i'(x,t)
\tag{3.72}
$$

where we assumed that the coefficient μ_i' are ordered in decreasing order at each time t:

$$
\mu_1'(t) \geq \mu_2'(t) \geq \ldots \geq \mu_S'(t).
$$

In other words the triplet $(\gamma_i'(\omega, t), \mu_i'(t), Z_i'(x, t))$ is the one with the ith biggest coefficient $\mu_i'(t)$ at time t. In addition, we denote with $\mu_i(t)$ the trajectory of the ith term of the Karhunen-Loève decomposition of u_0 or, rather, the evolution of the term that at $t = 0$ has the ith biggest variance. Observe that the function

$$u_S(x, t, \omega) = \sum_{i=1}^{S} \sqrt{\mu_i(t)} \gamma_i(\omega, t) Z_i(x, t) \tag{3.73}$$

is the exact solution of the problem (3.62) with initial condition $z_S(0)$ (the best S-rank approximation of u_0):

$$u_{0\,S}(x, \omega) = z_S(x, 0, \omega) = \sum_{i=1}^{S} \sqrt{\mu_i(0)} \gamma_i(\omega, 0) Z_i(x, 0) \in \mathcal{M}_S. \tag{3.74}$$

and differs, in general, from $z_S(t)$. Moreover, from what previously discussed, since $u_{0\,S}$ is in a S dimensional manifold the DO solution coincides to (3.73). This shows that the DO method, differently to the best S-rank approximation, may be affected by the truncation of the initial datum. Indeed the DO solution of problem (3.62) with initial condition u_0 will be always equal to the exact solution of the same problem with initial datum $u_{0\,S}$, that is generally different to u and z_S as well.

We consider first the case in which the best S-rank approximation z_S is continuously differentiable in time, as in the hypothesis of Theorem 4.1. For the problem we are analyzing this regularity assumption implies that the Sth eigenvalue of the correlation operator is differentiable in time, which can be translated in practice by requiring that the maximum neglected eigenvalue μ_{S+1}' of the correlation operator, would never cross the trajectories $\mu_1, ..., \mu_S$ at any time. Under this assumption the best rank-S approximation z_S coincides to (3.73) and the approximation error is given by

$$\epsilon_S(t) = \sum_{i=S+1}^{\infty} \mu_i'(t) = \sum_{i=S+1}^{\infty} \mu_i(t). \tag{3.75}$$

We see that, for a deterministic linear operator $\mathcal{L}(\cdot)$, the continuous time differentiability of z_S is a sufficient condition for the DO solution to coincide to the best rank S approximation.

3.4.3 Case III: effect of truncation - z_S is not continuously time differentiable

We remove any hypothesis of regularity on the evolution of the eigenvalues of the correlation operator. This implies that the trajectories of μ_i, μ_j may cross each other at any time instant, for any i, $j \in \mathbb{N}$. In particular, if the Sth eigenvalue of the correlation operator is not continuously time differentiable, which means that $\mu_k'(t)$ would cross $\mu_i(t)$ at some $t \in \mathcal{T}$, for some $i = S+1, ..., \infty$ and $k = 1, ..., S$, then the best approximation z_S will not be continuously differentiable in time. In this case the DO approximate solution and the best rank-S approximation

do not coincide:

$$z_S(x, t, \omega) = \sum_{i=1}^{S} \gamma_i'(\omega, t)\sqrt{\mu_i'(t)}\, Z_i'(x, t)$$
$$u_S(x, t, \omega) = \sum_{i=1}^{S} \gamma_i(\omega, t)\sqrt{\mu_i(t)}\, Z_i(x, t)$$

The DO approximation error is then strictly larger then the best approximation error:

$$\epsilon_S^{DO}(t) = \sum_{i=1}^{S} \mu_i(t) > \sum_{i=1}^{S} \mu_i'(t) = \epsilon_S(t). \tag{3.76}$$

and we do not have any control on it in terms of best approximation error. However observe that, for the specific problem considered in this section, the DO approximation error is always bounded by the initial truncation error:

$$\epsilon_S^{DO}(t) \leq \epsilon_S(0), \qquad \forall\, t \in \mathscr{T}. \tag{3.77}$$

3.4.4 An Illustrative Example

Consider the following problem:

$$
\begin{cases}
\dfrac{\partial u(x, t, \omega)}{\partial t} - \dfrac{\partial^2 u(x, t, \omega)}{\partial x^2} = 0 & x \in (0, 2\pi),\ t \in [0, \mathrm{T}],\ \omega \in \Omega \\
u(0, t; \omega) = u(2\pi, t; \omega) = 0 & t \in [0, \mathrm{T}],\ \omega \in \Omega \\
u(x, 0; \omega) = \alpha_1(\omega)\dfrac{1}{\sqrt{\pi}}\sin(x) + \alpha_2(\omega)\dfrac{1}{\sqrt{\pi}}\sin(2x) & x \in (0, 2\pi),\ \omega \in \Omega
\end{cases} \tag{3.78}
$$

where α_1, α_2 are independent uniform random variables with zero mean and variance $\mathbb{E}[\alpha_1^2] = 1$, $\mathbb{E}[\alpha_2^2] = 2$. As one can easily verify, the exact solution as well the total variance can be calculate analytically, i.e.:

$$
\begin{aligned}
u(x, t, \omega) &= \alpha_1(\omega)e^{-t}\frac{1}{\sqrt{\pi}}\sin(x) + \alpha_2(\omega)e^{-4t}\frac{1}{\sqrt{\pi}}\sin(2x), \\
\mathbb{V}ar_T[u](t) &= \mathbb{E}[\alpha_1^2]e^{-2t} + \mathbb{E}[\alpha_2^2]e^{-8t}.
\end{aligned} \tag{3.79}
$$

Observe that $u(x, 0, \omega)$ is a 2-rank function in the span of the first two eigenfunctions of the Laplace operator. Consequently the exact solution evolves in the manifold $\mathcal{M}_2$ at any time instant, the DO method degenerates to the POD method and, with $S = 2$, both the DO and the Karhunen-Loève solutions coincide with the exact solution.

Think now that we want to approximate the solution in a manifold of dimension 1. The initial datum is approximated according to the Karhunen-Loève decomposition by the principal component with largest variance, i.e. $z_1(x, 0, \omega) = u_1(x, 0, \omega) = \alpha_2(\omega)\dfrac{1}{\sqrt{\pi}}\sin(2x)$, and the DO method develops the following approximate solution:

$$u_1(x, t, \omega) = \alpha_2(\omega)e^{-4t}\frac{1}{\sqrt{\pi}}\sin(2x) \qquad x \in [0, 2\pi],\ t \in [0, \mathrm{T}],\ \omega \in \Omega \tag{3.80}$$
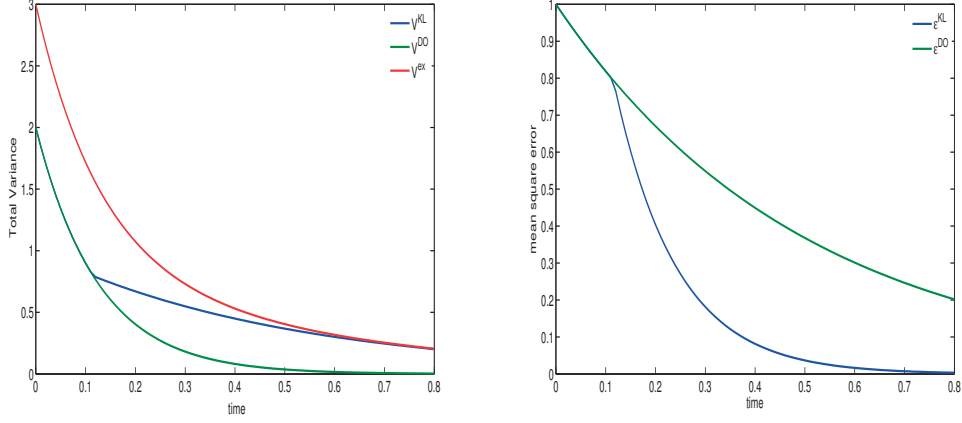
Figure 3.1 – On the left: Evolution of the total variance $\mathbb{V}ar_T(t)$ of the exact solution as well as the KL and DO approximate solution with $S = 1$. On the right: Time evolution of the mean square error $\epsilon(t)$ of the DO method with $S = 1$, compared to the best approximation error.

On the contrary the Karhunen-Loève approximate solution is given by:

$$
z_1(x, t, \omega) = \begin{cases} \alpha_2(\omega)e^{-4t}\dfrac{1}{\sqrt{\pi}}\sin(2x) & \text{for } t \in [0, \mathrm{T}] : \mathbb{E}[\alpha_1^2]e^{-2t} \le \mathbb{E}[\alpha_2^2]e^{-8t} \\[4mm] \alpha_1(\omega)e^{-t}\dfrac{1}{\sqrt{\pi}}\sin(x) & \text{for } t \in [0, \mathrm{T}] : \mathbb{E}[\alpha_1^2]e^{-2t} > \mathbb{E}[\alpha_2^2]e^{-8t} \end{cases}
\tag{3.81}
$$

That is not continuously time differentiable at $t^* = \frac{1}{6}\log\left(\frac{\mathbb{E}[\alpha_1^2]}{\mathbb{E}[\alpha_2^2]}\right)$. Figure 3.1 shows the evolution of the exact and approximate total variance (left) and the mean square error of the DO method compared to the best 1-rank approximation.

One can see that the error of the DO method is bounded by the initial truncation error and goes asymptotically to zero as t goes to infinity, but it is strictly larger than the best approximation error as soon as the eigenvalues cross each other. Indeed, while the best approximation error asymptotically goes to zero with exponential rate given by the second eigenvalue of the Laplace operator, i.e.:

$$
\epsilon(t)_{KL} = \min\left(\mathbb{E}[\alpha_1^2]e^{-2t}, \mathbb{E}[\alpha_2^2]e^{-8t}\right) \text{ and } \epsilon(t)_{KL} = \mathbb{E}[\alpha_2^2]e^{-8t} \quad \text{for } t > t^*,
$$

the exponential rate of DO approximation error is given by the smallest eigenvalue of the Laplace operator:

$$
\epsilon(t)_{DO} = \mathbb{E}[\alpha_1^2]e^{-2t} = \begin{cases} \epsilon_{KL}(t) \text{ for } t \in [0, \mathrm{T}] : t \le t^* \\[4mm] \frac{\mathbb{E}[\alpha_1^2]}{\mathbb{E}[\alpha_2^2]}e^{6t}\epsilon_{KL}(t) \text{ for } t \in [0, \mathrm{T}] : t > t^* \end{cases}
\tag{3.82}
$$

which shows that the DO error can not be bounded uniformly by the Karhunen-Loève error. This result does not contradict Theorem 4.1. Indeed at time t^* the truncated Karhunen-Loève expansion with rank $S = 1$ is not differentiable in time, so one important assumption in the Theorem 4.1 is not fulfilled.

3.5 Numerical examples

In this section, we will give some numerical examples to verify the performance of the DO approximation. Thus, we need to numerically solve the following DO system:

$$
\begin{cases}
\dfrac{\partial \bar{u}_S(x,t)}{\partial t} = \mathbb{E}\left[\mathscr{L}(u_S(x,t,\cdot))\right] & (3.83) \\[2ex]
\displaystyle\sum_{i=1}^{S} \mathbf{C}_{ij}(t)\dfrac{\partial U_i(x,t)}{\partial t} = \Pi^{\perp}_{\mathscr{U}}\mathbb{E}\left[\mathscr{L}(u_S(x,t,\cdot))\,Y_j(t,\cdot)\right], & j=1,\cdots,S, \quad (3.84) \\[2ex]
\dfrac{\partial Y_i(t,\omega)}{\partial t} = \left\langle \mathscr{L}^*(u_S(\cdot,t,\omega),\omega), U_i(\cdot,t)\right\rangle, & i=1,\cdots,S \quad (3.85)
\end{cases}
$$

3.5.1 Numerical discretization

For what concerns the numerical discretization of the system (3.83)-(3.85) we use the Finite Element method in the physical space, for equations (3.83)-(3.84), and Stochastic Collocation method (see e.g.[5, 11]) for equation (3.85). We assume that the input data are functions of a uniformly distributed random vector $\boldsymbol{\eta} = (\eta_1,...,\eta_N)$ so that the stochastic space $(\Omega,\mathscr{A},\mathscr{P})$ is parametrized by $\boldsymbol{\eta}$ and replaced by $(\Lambda,\mathscr{B}(\Lambda),f(\boldsymbol{\eta})d\boldsymbol{\eta})$ where Λ, $\mathscr{B}(\Lambda)$ and f denote respectively the domain, the Borel σ-algebra and the density function of $\boldsymbol{\eta}$. For the discretization in time we use a backward Euler scheme in which however eventual non linear terms are computed explicitly. Both the covariance matrix and the projection operator in (3.83) are treated explicitly, this allow us to linearize and completely decouple equations (3.83)-(3.84) from equations (3.85). In particular, the projection onto the orthogonal space in (3.84) is done on a basis freezed at the previous time step whereas the update of the random variables $\{Y_i\}$ in (3.84) is done on the newly computed basis. The splitting scheme is therefore of "Gauss-Seidel" type. Let $\mathscr{U}_h$ denote the finite element space of continuous piecewise linear functions on a regular triangulation of the spatial domain D with mesh size h, $\{\xi_{\boldsymbol{i}} \in \Lambda\}$ the set of N_y tensorized Gauss-Legendre collocation points and Δt the time step. Then the DO approximate solution at time $t^n = n\Delta t$ is discretized as follow:

$$
u^n_{S,h,N_y}(x,\boldsymbol{\eta}) = \sum_{j=1}^{N_h} \bar{U}^n_j \rho_j(x) + \sum_{i=1}^{S}\left(\sum_{j=1}^{N_h} U^n_{j,i}\rho_j(x)\sum_{k=1}^{N_y} Y^n_{k,i}\mathscr{L}_k(\boldsymbol{\eta})\right)
$$

where $\{\rho_i\}_{i=1}^{N_h}$ and $\{\mathscr{L}_k\}_{k=1}^{N_y}$ are respectively the finite elements basis functions in D and the multivariate tensorized Lagrange polynomials on the grid $\{\xi_{\boldsymbol{k}}\}$ in Λ. Observe that the first moment of the DO approximate solution at $t = t^n$ corresponds to the function $\bar{U}^n_h(x) =$

$\sum_{j=1}^{N_h} \bar{U}_j^n \rho_j(x)$ and the total variance can be easily computed as the sum of the variances of the stochastic coefficients $\{Y_i\}$, i.e.:

$$\mathrm{Var}_T[u_{S,h,N_y}^n] = \sum_{i=1}^{S} \mathrm{Var}[Y_i^n] = \sum_{i=1}^{S} \sum_{k=1}^{N_y} Y_i^n(\boldsymbol{\xi}_k)^2 w_k$$

where $\{w_k\}_{k=1}^{N_y}$ are the weights of the Gaussian quadrature formula associated to the collocation points of the stochastic grid. (For further details concerning other possible types of stochastic grids we refer to [11, 101]). Moreover the computation of the covariance of the stochastic coefficients is explicitly required in the equations (3.84). Indeed the equations for the deterministic basis functions $\{U_i\}$ are coupled together by the covariance matrix. A "natural" option to decouple the equations consists in multiplying both sides in (3.84) by the inverse of the covariance matrix. Unfortunately this is often not possible, since the covariance matrix $\mathbf{C}(t^n) = \mathbf{C}^n$ may be singular or very ill conditioned at some time instant t^n. A straightforward example is provided by any system of PDEs with stochastic coefficients and deterministic initial data: the DO approximate solution will require $S \geq 1$ number of modes, in general, even if the initial covariance matrix is identically equal to zero and then singular at least for the very first iteration. Furthermore the rank of the covariance matrix typically evolves in time whatever the initial condition is. For instance in the very simple case of linear diffusion equations with no forcing terms, the rank tends asymptotically to zero as t goes to infinity whereas for non linear problem it may drastically increases or decreases during the time evolution. This makes unsuitable also the direct use of the pseudo-inverse of $\mathbf{C}^n$, since such approach automatically sets to zero the "non active" deterministic basis functions and then prevents the rank from increasing. Instead of multiplying both sides of (3.84) by the pseudo-inverse of $\mathbf{C}$, denoted by $\mathbf{C}^\dagger$, we reformulate directly the problem in this form:

$$\frac{\partial \mathbf{U}}{\partial t} = \mathbf{C}^\dagger \Pi_{\mathcal{U}}^\perp \mathbb{E}\left[\mathbf{Y} \mathcal{L}(u_S)\right],$$

that is equivalent to solve (3.84) when the covariance matrix has full rank. From a numerical point of view the strategy that we have adopted in this work, is based on diagonalizing the covariance matrix at each time step in order to completely decouple the system of equation (3.84). Indeed, even if the covariance matrix at $t = 0$ is diagonal, the DO method does not preserve in general the un-correlation of the stochastic coefficients for $t > 0$. For a better understanding let us write equation (3.84) in algebraic form with notations:

$$\mathbf{U}^n \in \mathbb{R}^{N_h \times S}: U_{i,j}^n = \langle U_j(\cdot, t^n), \rho_i \rangle, \qquad \mathbf{Y}^n \in \mathbb{R}^{N_y \times S}: Y_{i,j}^n = Y_j(\xi_i, t^n)$$

$$\mathbf{F}^n \in \mathbb{R}^{N_h \times N_y}: \mathbf{F}_{jk}^n = \langle \mathcal{L}(u_S(\cdot, \xi_k, t^n)), \rho_j \rangle$$

Furthermore we denote with $\mathbb{M}$ the Finite Element mass matrix and with $\mathbb{E}_{N_y}[.]$ the discretized expected value, computed by the quadrature formula on the collocation points $\{\xi_k\}_{k=1}^{N_y}$. Then

the algebraic formulation of system (3.84) is the following:

$$\mathbb{M}\mathbf{U}^{n+1}\mathbf{C}^n = \mathbb{M}\mathbf{U}^n\mathbf{C}^n + \Delta t(\mathbb{I} - \mathbb{M}\mathbf{U}^n\mathbf{U}^{nT})\mathbb{E}_{N_y}[\mathbf{F}^*\mathbf{Y}^n] \tag{3.86}$$

where $\mathbf{F}^*$ denotes that the diffusion term is always treated implicitly, while eventual non linear terms are computed at $t = t^n$. Let $\mathbf{V}^n = (v_1^n, ..., v_S^n) \in \mathbb{R}^{S \times S}$ be the matrix of eigenvectors of $\mathbf{C}^n$ and $\Sigma^n \in \mathbb{R}^{S \times S} : \Sigma_{ij}^n = \delta_{ij}\mu_i^n$ the matrix of eigenvalues, such that $\mathbf{C}^n\mathbf{V}^n = \mathbf{V}^n\Sigma^n$. Then multiplying both sides in (3.86) by $\mathbf{V}^n$ we get:

$$\mathbb{M}\mathbf{U}^{n+1}\mathbf{V}^n\Sigma^n = \mathbb{M}\mathbf{U}^n\mathbf{V}^n\Sigma^n + \Delta t(\mathbb{I} - \mathbb{M}\mathbf{U}^n\mathbf{U}^{nT})\mathbb{E}_{N_y}[\mathbf{F}^*\mathbf{Y}^n\mathbf{V}^n]$$

Observe that $\mathbf{Z}^n = \mathbf{Y}^n\mathbf{V}^n$ is a vector of uncorrelated random variables with variance $\mathbb{E}_{N_y}[(Z_i^n)^2] = \mu_i^n$. Finally we solve:

$$\mathbb{M}\mathbf{U}^{n+1}\mathbf{V}^n = \mathbb{M}\mathbf{U}^n\mathbf{V}^n + \Delta t(\mathbb{I} - \mathbb{M}\mathbf{U}^n\mathbf{U}^{nT})\mathbb{E}_{N_y}[\mathbf{F}^*\mathbf{Y}^n\mathbf{V}^n]\Sigma^{\dagger n} \tag{3.87}$$

where $\Sigma^{\dagger n}$ is the pseudoinverse of Σ^n with tolerance ϵ, that is:

$$\Sigma_{ij}^{\dagger n} = \begin{cases} \delta_{ij}\mu_i^n & \mu_i^n > \epsilon\mu_{max}^n \\ 0 & \text{otherwise} \end{cases}$$

Roughly speaking we impose that only the "directions" associated to the eigenvalues $\mu_i^n > \epsilon\mu_{max}^n$ evolve, while the others remain constant. An alternative integrator for the low rank approximation of time dependent matrix is proposed in [82]. This is based on a suitable splitting of the orthogonal projector onto the tangent space.

As already mentioned, the DO method explicitly requires the deterministic basis functions to be orthonormal in $L^2(D)$. At the continuous level the orthonormality is preserved at any time instant thanks to condition (3.11). On the other hand, many numerical schemes, including the one discussed here, will not preserve the orthonormality of the discrete basis (see e.g. [56] for a discussion on orthogonality preserving numerical schemes). We therefore re-orthogonalize at each time step the spatial basis $\{U_i\}$ by means of a QR factorization (where the matrix Q is orthogonal with respect to the continuous $L^2(D)$ inner product, i.e. $Q^T\mathbb{M}Q = \mathbb{I}$).

3.5.2 Linear parabolic problem with random initial conditions

We start by considering the following simple problem already discussed in Section 5:

$$\begin{cases} \dfrac{\partial u(x,t,\omega)}{\partial t} - \dfrac{\partial^2 u(x,t,\omega)}{\partial^2 x} = 0, & x \in [0,8], \, t \in \mathcal{T}, \omega \in \Omega \tag{3.88} \\ u(0,t;\omega) = u(8,t,\omega) = 0, & t \in \mathcal{T}, \omega \in \Omega \tag{3.89} \end{cases}$$

where the initial condition is a random field. Here we take

$$u_0(x,\omega) = \alpha(\omega)u_{01}(x) = \alpha(\omega)\frac{1}{4}| - x + 4| + 1$$
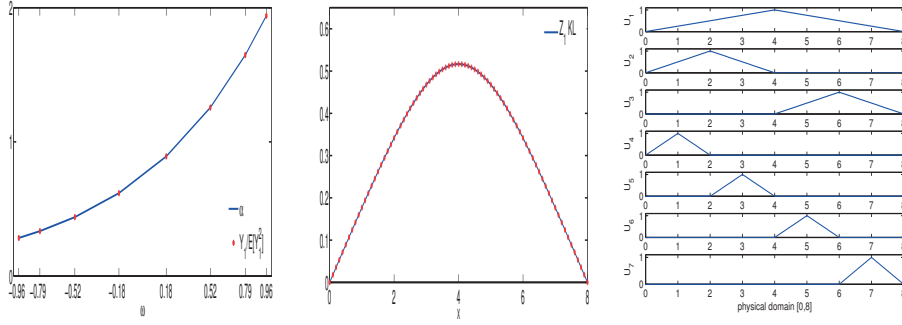
Figure 3.2 – Left: The random variable $\alpha(\eta)$, blue, and $Y_1(t)$ (scaled by the variance $\mathbb{E}[Y_1(t)^2]$), red markers, at the collocation points. Middle: The first mode with $S = 1$ (red markers) and the principal component Z_1 (blue) at T=1. Right: Hierarchical basis function.

where $\alpha(\omega) = e^{\eta(\omega)} - \mathbb{E}[e^{\eta}]$ and $\eta(\omega)$ is a uniformly distributed random variable in $[-1,1]$. It is easy to see that the exact solution is a stochastic field with the same distribution of $\alpha(\omega)$. Analytically it can be calculated as $u_{ex}(x,\omega,t) = \alpha(\omega)\psi(x,t)$, being ψ the solution of the deterministic diffusion PDE with initial condition $\psi(x,0) = u_{01}(x)$. By normalizing the field $\psi(x,t)$ the solution can be rewritten in accordance with the Karhunen-Loève decomposition, having only one principal component $Z_1(x,t) = \psi(x,t)/\|\psi(.,t)\|_0$ and one stochastic coefficient $\gamma_1(\omega,t) = \alpha(\omega)\|\psi(.,t)\|_0$. We can easily verify that the couple (Z_1, γ_1) satisfy the DO system (3.13)-(3.14) with $S = 1$:

$$
\begin{cases}
\dot{Z}_1 = \dfrac{\dot{\psi}}{\|\psi\|_0} - < \dfrac{\dot{\psi}}{\|\psi\|_0}, \dfrac{\psi}{\|\psi\|_0} > \dfrac{\psi}{\|\psi\|_0} = \Delta Z_1 - < \Delta Z_1, Z_1 > Z_1 = \Pi_{\mathcal{Z}}^{\perp}\{\Delta Z_1\} \\
\dot{\gamma}_1 = < \dfrac{\dot{\psi}}{\|\psi\|_0}, \dfrac{\psi}{\|\psi\|_0} > \gamma_1 = < \Delta Z_1, Z_1 > \gamma_1
\end{cases}
\tag{3.90}
$$

and initial condition $Z_1(x,0) = u_{01}(x)/\|u_{01}\|_0$, $\gamma_1(\omega,0) = \alpha(\omega)$. This confirms again the exactness of the DO method in case of deterministic operator and initial condition that belongs to a finite dimensional manifold. Now we want to show that the exactness of the DO method is preserved at the discrete level as well: let $u_{h,N_y,\Delta t}$ denote the discrete solution of (3.88), obtained by using piecewise linear continuous Finite Elements in space with mesh size h, Stochastic Collocation method on N_y Gauss-Legendre points in $\eta(\omega)$ and backward Euler discretization in time with step Δt. We show that the corresponding DO approximate solution coincides with $u_{h,N_y,\Delta t}$. Let $\mathbb{M}$ and $\mathbb{K}$ be respectively the mass and stiffness matrix of the Finite Element discretization and let μ_1^n denotes the variance of the random variable Y_1^n. The algebraic system for the DO solution $u_{1,h,N_y,\Delta t} = U_{1,h,\Delta t} Y_{1,Ny,\Delta t}^T$ with rank one is given by:

$$
\begin{cases}
MU^{n+1}\mu^n + \Delta t \mathbb{K} U^{n+1}\mu^n = \mathbb{M}U^n\mu^n + \Delta t \mathbb{M}U^n U^{nT}\mathbb{K}U^n\mu^n \tag{3.91} \\
Y^{n+1T} + \Delta t U^n \mathbb{K} U^n Y^{n+1T} = Y^{nT} \tag{3.92}
\end{cases}
$$

where for simplicity of notation, we have omitted the subscripts. By multiplying the first equation by Y^{n+1} and using (3.92) we get:

$$\mathbb{M}U^{n+1}Y^{n+1T} + \Delta t\mathbb{K}U^{n+1}Y^{n+1T} = \mathbb{M}U^n Y^{n+1T} - \mathbb{M}U^n Y^{n+1T} + \mathbb{M}U^n Y^{nT}$$

or equivalently:

$$\mathbb{M}u_1^{n+1} + \Delta t\mathbb{K}u_1^{n+1} = \mathbb{M}u_1^n$$

which exactly corresponds to the algebraic system of the discretized problem for $u_{h,N_y,\Delta t}$. Figure 3.2 (middle) shows that the deterministic basis function U_1 evolves in time and coincides to the principal component Z_1 of the discrete solution $u_{h,N_y,\Delta t}$ at each time step. The stochastic coefficient Y_1 is as well proportional to the initial random parameter, with variance that decreases in time and coincides with the total variance of the solution. Figure 3.2 (left) shows that Y_1^n, normalized with respect to the variance at time t^n ($\mathbf{E}[(Y_1^n)^2]$), is equal to $\alpha(\omega)$ at each time step.

Finally we aim at analyzing the efficacy of the DO method in case of over-approximation, that occurs when the DO approximate solution is defined in a manifold of dimension larger then the rank of the exact solution. To this purpose we again apply the DO method to problem (3.88) with $S > 1$ whereas we have seen that only one mode is really needed. We initialize the deterministic basis functions $\{U_i\}_{i=1}^S$ to a sequence of hierarchical functions as in Figure 3.2 (right). To preserve the consistency with the initial datum, the first stochastic coefficient is initialized to $\alpha(\omega)$ and all the other coefficients are initialized to zero. Since the DO method requires the deterministic basis functions to be orthonormal in $L^2(D)$, the first step consists in the re-orthonormalizing the initial hierarchical basis functions. From a computational point of view, this is achieved by using the QR decomposition, with respect to the continuous $L^2(D)$ inner product. Let $(\hat{U}_1,...,\hat{U}_S)$ denotes the set of orthonormalized basis functions. Then the initial datum is expanded as

$$u_0(x,\omega) = \sum_{i=1}^S \hat{U}_i(x,0)\hat{Y}_i(\omega,0) \tag{3.93}$$

with $\hat{Y}_i(\omega,0) = 0$ for $i = 2,...,S$. As the system evolves in time, all the spatial basis functions evolve and all the random variables become in general different from zero (Figure 3.3). However the stochastic coefficients $\{\hat{Y}_i\}$ are all linearly dependent and the rank of the covariance matrix $\mathbf{C}_{ij}^n = \mathbb{E}[\hat{Y}_i^{nT}\hat{Y}_j^n]$ remains constantly equal to one at each time step, as long as the total variance of the solution is larger than zero. This confirms that the DO method in the version proposed here, effectively deals with singular covariance matrices in case of over-approximation and is able to identify the effective dimension of the manifold of the solution. Moreover we remark that at each time step only one deterministic PDE is actually solved, thanks to the diagonalization technique discussed in Section 6.1. Also in case of over-approximation we have verified numerically that the DO solution corresponds to the discrete solution $u_{h,N_y,\Delta t}$ at each time step.
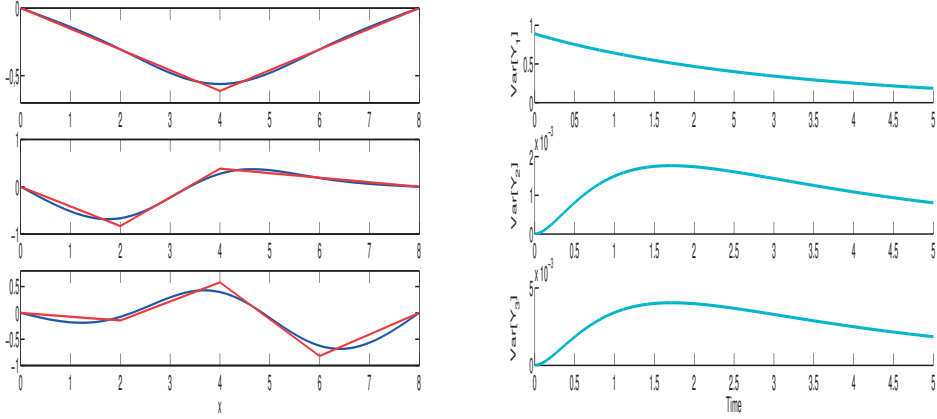
Figure 3.3 – Left: The re-orthogonalized modes at time $T = 0$ (red) and $T = 5$ (blue), with $S = 3$. Collocation points $N_y = 11$, time step $\Delta t = 0.001$, spatial discretization $h = 0.1$ and threshold $\epsilon = 1.e - 16$. Right: Evolution in time of the variance of the stochastic coefficients.
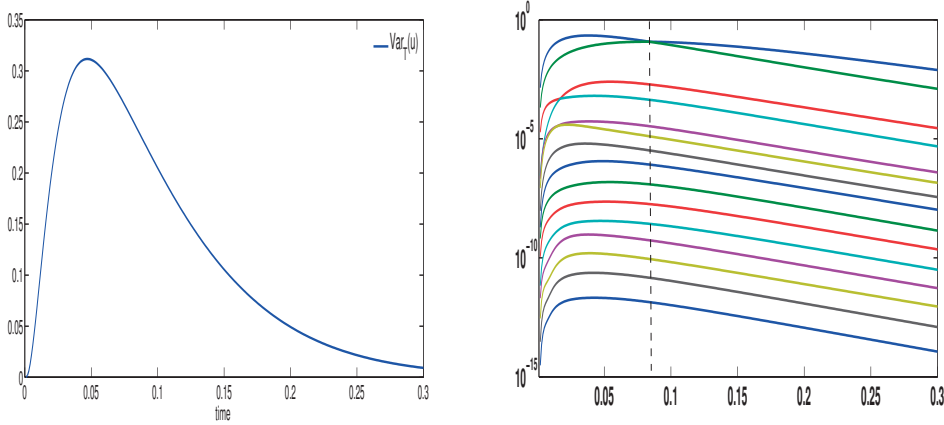
Figure 3.4 – Left: The evolution of the total variance of the solution. Right: Evolution of the first 15 eigenvalues of the correlation operator in log scale

3.5.3 Linear parabolic problem with random diffusion coefficient

We consider now the following linear parabolic equation:

$$\begin{cases} \dfrac{\partial u(x,t,\omega)}{\partial t} - \operatorname{div}\left(a(x,\omega)\dfrac{\partial u(x,t,\omega)}{\partial x} \right) = 0, & x \in [0,1], \, t \in \mathcal{T}, \, \omega \in \Omega & (3.94) \\ u(0,t,\omega) = u(1,t,\omega) = 0, & t \in \mathcal{T}, \, \omega \in \Omega & (3.95) \end{cases}$$

where now $\mathscr{L}(\cdot)$ is actually a stochastic differential operator, being the diffusion coefficient a a random field on $(\Omega, \mathscr{A}, \mathscr{P})$ taking values in $L^\infty(D)$. This means that the eigenvalues and the eigenfunctions of $\mathscr{L}(\cdot)$ are random fields in $(\Omega, \mathscr{A}, \mathscr{P})$ as well. For reasons of existence and uniqueness of the solution we assume $a(\cdot, \omega)$ to be a strictly positive and bounded function
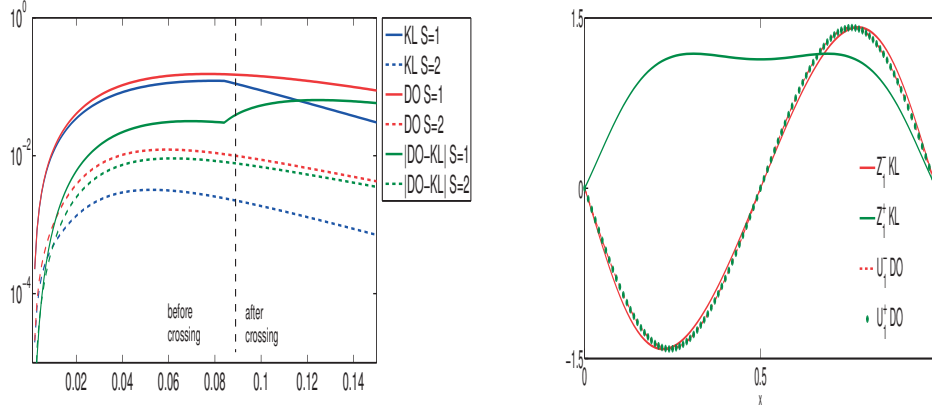
Figure 3.5 – Left: The best approximation error (blue) and the error of the DO approximate solution (red) in the $L^2(D) \times L^2(\Omega)$ norm with $S = 1$ (solid line) and $S = 2$ (dotted line), in log scale. In green the $L^2(D) \times L^2(\Omega)$ norm of the difference between the DO and the Karhunen-Loève solutions with $S = 1$ (solid line) and $S = 2$ (dotted line). Right: The first mode of the Karhunen-Loève expansion and DO solution at the time step just before (red) and after the crossing of eigenvalues (green solid line-Karhunen-Loève , green markers-DO). Discretization parameters: Gauss-Legendre collocation points in tensor grid with $N_y = 11$ in each direction, spatial discretization $h = 0.02$, time step $\Delta t = 0.001$

over D for each random event $\omega \in \Omega$. Here we consider a coefficient having the following form:

$$a(x, \omega) = \bar{a}(x) + \sum_{i=1}^{2} \left(\eta_{2i-1}(\omega) \cos(i\pi x) + \eta_{2i}(\omega) \sin(i\pi x) \right) \tag{3.96}$$

where $\bar{a} = 1.45$ and $\eta_1, ..., \eta_4$ are zero mean uniform independent random variables with variance $\mathbb{E}[\eta_i^2] = (1/3) \cdot 10^{-i+1}$.

3.5.4   Deterministic initial condition

The initial condition is taken to be a deterministic function and is given by:

$$u_0(x) = 10 \sin(\pi x)$$

By this choice we can also verify the stability of the DO method in case of an initial zero rank covariance matrix and emphasize the differences with respect to the type of problems discussed in Section 5.1 and 6.2. Here the solution $u(x, t, \omega)$ is actually a function of the random variables $\boldsymbol{\eta} = (\eta_1, ..., \eta_4)$. Figure 3.4 (left) shows the evolution of the total variance of the solution.

At each time step, we can introduce the parameter-to-solution map $u(\cdot, t, \boldsymbol{\eta}) : [-1, 1]^4 \to H_0^1(D)$. Defining now the set $\mathcal{V}(t) = \{u(\cdot, t, \boldsymbol{\eta}), \boldsymbol{\eta} \in [-1, 1]^4\} \subset H_0^1(D)$, at each time step the solution $u(\cdot, t, \boldsymbol{\eta}(\omega))$ is in $\mathcal{V}(t)$ for all $\omega \in \Omega$. The manifold $\mathcal{V}(t)$ is a multidimensional manifold which, in this example, evolves in time.   First of all we compute and analyze the Karhunen-Loève
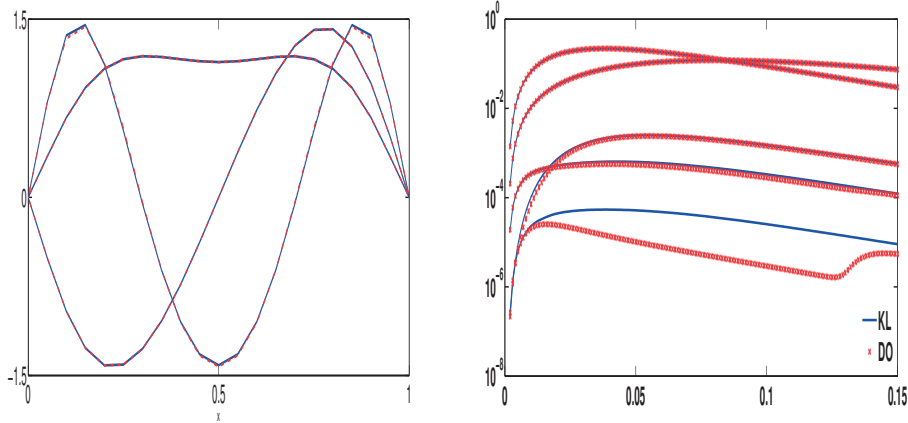
Figure 3.6 – Left: The first three modes of the DO approximate solution (red markers) with $S = 5$ and of the Karhunen-Loève expansion (blue, solid line) just after the crossing t^*. Left: The eigenvalues of the covariance matrix of the stochastic coefficient of the DO solution (red markers) with $S = 5$ and the first 5 eigenvalues of the covariance operator of the reference solution (blue, solid line), in log scale. Note that the first DO and KL modes and eigenvalues are almost indistinguishable. Discretization parameters: Gauss-Legendre collocation points in tensor grid with $N_y = 11$ in each direction, spatial discretization $h = 0.02$, time step $\Delta t = 0.001$

decomposition of the numerical reference solution computed with a very accurate (and costly) Stochastic Collocation method. The analysis of the Karhunen-Loève decomposition is very useful to understand what we can expect from the DO method. Moreover, it allows us to directly compare the DO solution with the best approximation. In Figure 3.4 (right) we see the evolution of the eigenvalues of the covariance operator. Observe that only few of them reach remarkable values. This immediately shows that the solution u can be well approximated in a low rank format. On the other hand, notice that the first two eigenvalues cross each other at time $t^* \approx 0.08$. This implies that the 1-truncated Karhunen-Loève expansion is not continuously differentiable in time at the crossing. Hence Theorem 4.1 only applies for $t < t^*$, for $S = 1$. On the other hand, the case $S = 2$ seems to be smooth (at least up to the final computational time $T = 0.3$). Similar considerations apply to successive modes, for $S = 3, 4$, $S = 5, 6$, etc.

The numerical results confirm the theoretical ones given in Section 4 and are consistent with the observations above. The errors have been calculated with respect to a reference solution, numerically computed with the Stochastic Collocation method in a fine tensor grid and by using the same discretization parameters, in time and space, chosen for solving the DO system. Figure 3.5 (left) shows the approximation error in the $L^2(D) \times L^2(\Omega)$ norm of the truncated Karhunen-Loève expansion as well as the DO solution, with rank $S = 1$ (solid line) and $S = 2$ (dotted line). We see that the approximation error of the DO solution with rank equal to 1 is proportional to the best approximation error only until the first two eigenvalues of the Karhunen-Loève expansion cross each other. Before the crossing , the difference in the $L^2(D) \times L^2(\Omega)$ norm between the DO solution and the truncated Karhunen-Loève expansion
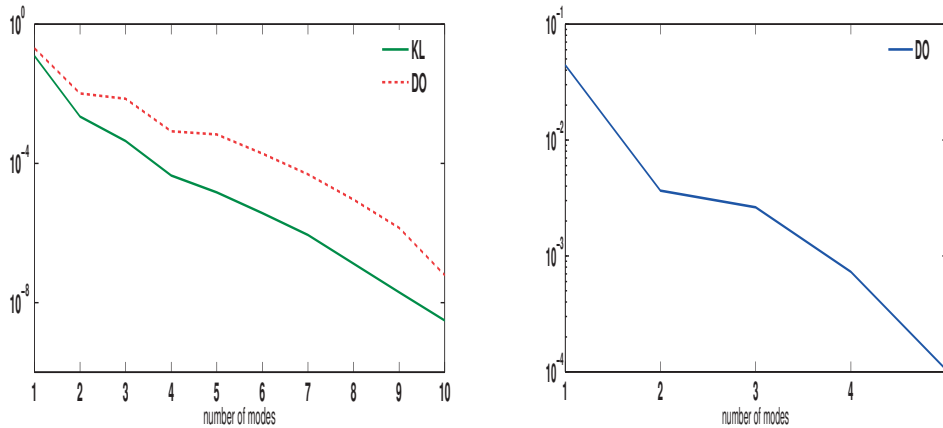
Figure 3.7 – Left: The best approximation error (green) and the error of the DO approximate solution (red, dotted line) in $L^2(D) \times L^2(\Omega)$ norm w.r.t. the number of modes at time $T = 0.1$. Right: The $L^2(D)$-error on the mean of the DO approximate solution w.r.t. the number of modes, at time $T = 0.1$

with rank 1 is bounded by the best approximation error so the error of the DO method is well controlled by the error of the Karhunen-Loève expansion. After the crossing, the bound clearly degenerates. The problem is due to the evolution of the first mode of the Karhunen-Loève expansion which is no longer continuous in time. In Figure 3.5 (right) we see the first mode of the Karhunen-Loève expansion in two consecutive time step, just before and after the crossing. By using only one mode, the DO method is not able to follow accurately the evolution of the first mode of the Karhunen-Loève expansion after the crossing. Indeed the hypothesis of continuous differentiability of the rank-1 Karhunen-Loève approximation is not fulfilled and Theorem 4.1 can not be applied after the crossing. However this problem can be overcome by using an approximation with rank larger than 1, as we can see in Figure 3.6 (left). The plot shows the first three modes of the DO approximate solution which are very close to those of the Karhunen-Loève expansion even after the crossing t^*. Moreover the evolution of the eigenvalues of the covariance matrix related to the DO solution, is comparable to the evolution of the eigenvalues of the covariance operator of the exact solution (Figure 3.6, right). For what concerns the $L^2(D) \times L^2(\Omega)$ error (Figure 3.7, left) and the $L^2(D)$ error on the mean (Figure 3.7, right), good levels of accuracy can be achieved by using only few modes. Moreover, the plot clearly shows an exponential convergence rate with respect to the number of modes.

Stochastic initial condition

We analyze again problem (3.94), by adding this time some sources of stochasticity in the initial condition as well. While in the first case the goal was to verify the effectiveness of the proposed numerical method in case of singular covariance matrix, and study the accuracy and the convergence of the DO method with respect to the regularity of the Karhunen-Loève decomposition and the number of modes; now our aim is to analyze the effect of the truncation
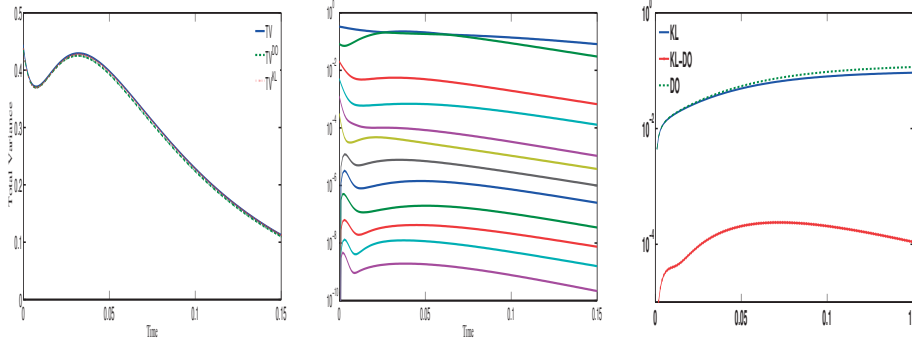
Figure 3.8 – Left: time evolution of the total variance of the reference solution (green), the truncated KL expansion (blue) and the DO approximate solution (red), with $S = 3$. Middle: time evolution of the first 12 eigenvalues of the covariance operator. Right: Plot of $\| u^{DO} - u^{KL} \|$ with the norm in (3.52) (red, dotted line), error in norm $L^2([0, t], H^1(D) \times L^2(\Omega))$ of the KL (blue, solid line) and DO (green, marked line) approximate solution. Discretization parameters: Gauss-Legendre collocation points with an isotropic sparse-grid of Smolyak type of level 4, spatial discretization $h = 0.02$, time step $\Delta t = 0.001$

on the initial datum and especially verify the bound for the DO approximation error, obtained in (3.52). We assume:

$$u(x, 0, \omega) = 10 sin(\pi x) + \sum_{i=1}^{6} \sqrt{\frac{2}{4^{i-1}}} y_i(\omega) sin(i\pi x) \tag{3.97}$$

where $y_1, ..., y_6$ are zero mean, uniform and independent random variables in $[-1, 1]$. Furthermore we assume that the random vector $(y_1, ..., y_6)$ is independent to $(\eta_1, ..., \eta_4)$, which implies that for this example we deal with a stochastic space of dimension 10. Figure 3.8 (left) shows the evolution of the total variance of the reference solution, computed with the Stochastic Collocation method by using an isotropic sparse-grid of Smolyak type with level 4 (see [101] for details) and the same discretization parameters, in time and space, used for solving the DO system. Figure 3.8 (middle) shows the evolution in time of the first 12 eigenvalues of the covariance operator where we can see that only the first two cross in the time interval considered. In Figure 3.8 (right) we show the results for the DO approximate solution as well as the truncated Karhunen-Loève expansion with $S = 3$ number of modes. Observe that the initial condition for the DO system is simply obtained by truncating the sum in (3.97) to $i = 3$. Furthermore we remark that, even if the first two eigenvalues cross, the best 3-rank approximate solution is continuously differentiable in time, at least up to the final computational time $T = 0.15$, and all the hypothesis of Theorem 4.1 hold. In the figure we verify the bound for the DO approximation error, and specifically inequality (3.52) at each time step. In particular we show in red (dotted line) the difference between the DO and the KL approximate solution (with the norm specified in (3.52)), in blue (solid line) and in green (marked line) respectively the error of the KL and DO approximate solution in norm $L^2([0, t], H^1(D) \times L^2(\Omega))$. We see once more that the DO error is well controlled by the KL error (best $S-$rank approximation).
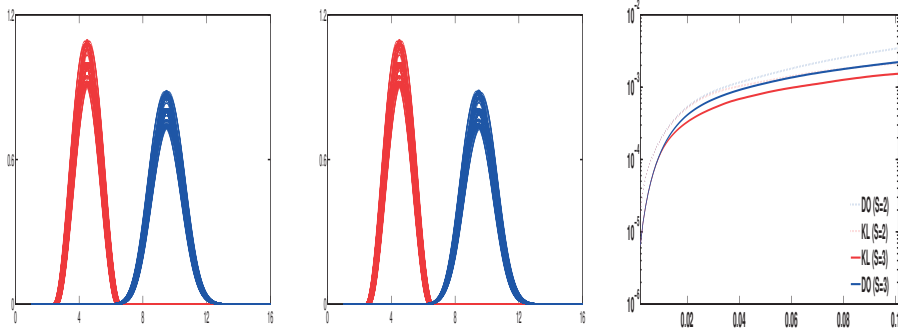
Figure 3.9 – Left: The reference solution in the collocation points at $t = 0$ and $t = 0.1$. Middle: The DO approximate solution in the collocation points at $t = 0$ and $t = 0.1$ with $S = 3$. Left: Approximation error of the Karhunen-Loève (red) and DO (blue) approximate solution in norm $L^2([0, t], H^1(D) \otimes L^2(\Omega))$ (solid line $S = 3$, dotted line $S = 2$). Discretization parameters: time step $dt = 1.e - 3$, $h = 0.008$, Gauss-Legendre collocation points with an isotropic sparse-grid of Smolyak type of level 3.

Linear convection-dominated problem

We consider now a linear convection-dominated problem with a constant transport coefficient and stochastic diffusion and initial datum:

$$\frac{\partial u(x, t, \omega)}{\partial t} - \mu \mathrm{div}\left(a(x, \omega)\frac{\partial u(x, t, \omega)}{\partial x}\right) + b\frac{\partial u(x, t, \omega)}{\partial x} = 0, \qquad x \in [1, 16], t \in \mathcal{T}, \omega \in \Omega$$
$$u(1, t, \omega) = u(16, t, \omega) = 0, \qquad t \in \mathcal{T}, \omega \in \Omega$$
$$(3.98)$$

with the diffusion coefficient a defined in (3.96), μ and b constant and equal to 1 and 100, respectively. The initial condition is a stochastic field defined as:

$$u_0(x, \omega) = \begin{cases} \delta(\omega)(\sin(0.5\pi(x - 3.5\gamma(\omega))) + 1) & (3.5\gamma(\omega) - 1) < x < (3.5\gamma(\omega) + 3) \\ 0 & x \leq (3.5\gamma(\omega) - 1) \\ 0 & x \geq (3.5\gamma(\omega) + 3) \end{cases} \qquad (3.99)$$

where δ and γ are two independent uniformly distributed random variables with mean 1 and variance $1/3 \cdot 10^{-2}$ and $4/3 \cdot 10^{-2}$, respectively. The mesh size in the physical space is chosen small enough such that the numerical scheme used for computing both the DO and the reference solution does not require any stabilization strategy. The results achieved show that the DO method performs quite well for the example under consideration. Figure 3.9 (left) and Figure 3.9 (middle) display the reference solution and the DO approximate solution with $S = 3$, respectively, in the collocations points, at $t = 0$ and $t = 0.1$. The plots show that the DO approximation effectively describes the dynamics and the variance of the traveling wave. The analysis of the approximation error is reported in Figure 3.9 (right) where both the DO and the Karhunen-Loève approximate solutions are compared with the reference solution in
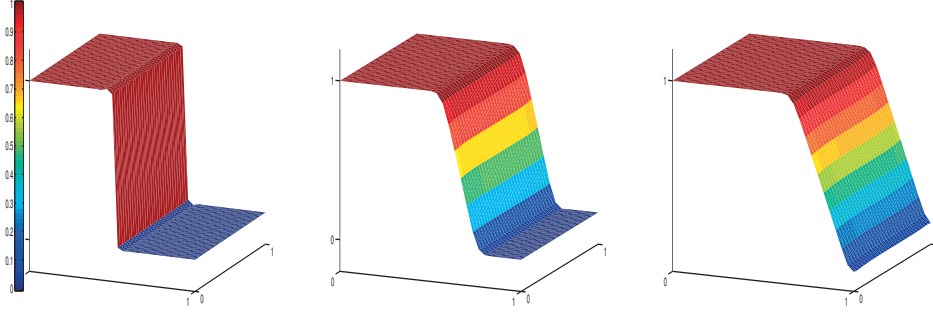
Figure 3.10 – Mean solution at time $t = 0$ (left), $t = 0.35$ (middle), $t = 1$ (right). Discretization parameters: Gauss-Legendre collocation points $N_y = 41$, spatial discretization $h = 0.03$, time step $\Delta t = 0.001$

norm $L^2([0, t], L^2(\Omega) \otimes H^1(D))$. Also in this case the DO approximation error behaves similarly as the KL error. We remark, however, that in this case of a convection-dominated problem, the numerical scheme proposed here may require a smaller time discretization step than for the pure diffusion case to guarantee a good accuracy level. This is probably due to the strategy adopted for the discretization of the projection operator which is indeed treated explicitly. A possible better strategy would consist in adopting a splitting scheme when solving the equations for the deterministic modes in order to first transfer the modes according to the advection field before performing the projection in the second step.

3.5.5 Parabolic equation with non linear reaction term

To conclude we consider a reaction-driven non liner parabolic operator with stochastic coefficient. The problem is defined as follows:

$$
\begin{aligned}
&\frac{\partial u(\mathbf{x}, t, \omega)}{\partial t} - \mu \Delta u(\mathbf{x}, t, \omega) = f(u(\mathbf{x}, t, \omega)), && \mathbf{x} \in D = [0,1]^2, t \in \mathcal{T}, \omega \in \Omega, \\
&\frac{\partial u}{\partial \mathbf{n}}(\boldsymbol{\sigma}, t, \omega) = 0, && \boldsymbol{\sigma} \in \partial D, t \in \mathcal{T}, \omega \in \Omega \\
&u(\mathbf{x}, 0, \omega) = \begin{cases} 1 & \text{if } x_1 \leq 0.5, \\ 0 & \text{if } x_1 > 0.5. \end{cases} && \mathbf{x} \in D, \omega \in \Omega,
\end{aligned}
\tag{3.100}
$$

The reaction term is a cubic polynomial in u, i.e. $f(u) = \beta u(u - 1)(\alpha(\omega) - u)$, with constant excitation rate β and stochastic threshold potential $\alpha(\omega)$. We assume $\alpha(\omega)$ to be a uniformly distributed random variable. The initial condition is instead deterministic and it is represented by a step function equal to one for values of coordinate x_1 smaller than 0.5 and zero otherwise. The solution is a traveling wave, whose direction and speed, proportional to β, is determined by the value of the random variable α. After a while, the wave exits the computational domain and the solution tends to the constant function $u = 1$ if $\alpha(\omega) < 0.5$, and $u = 0$ if $\alpha(\omega) > 0.5$, irrespectively of the initial datum. Therefore for $t \gg 0$ the solution is at most a rank-1 function
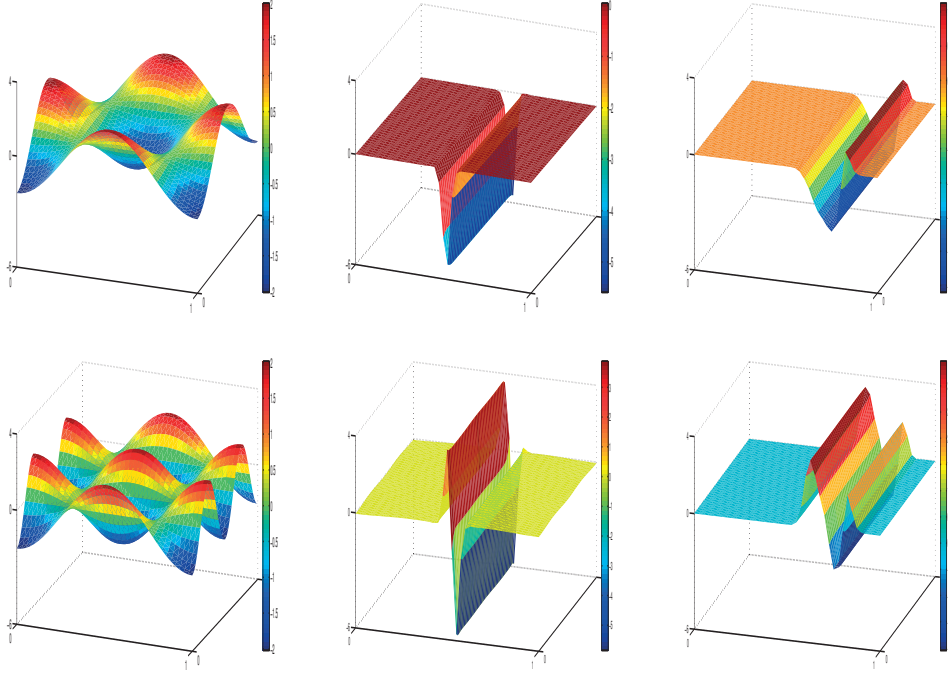
Figure 3.11 – First (top) and second (bottom) mode of the DO approximate solution at time $t = 0$ (left), $t = 0.05$ (middle) and $t = 0.5$ (right) with a number of modes $S = 6$, excitation rate $\beta = 100$ and threshold potential $\alpha(\omega)$ uniform r.v. in $[0, 0.4]$. Discretization parameters: Gauss-Legendre collocation points $N_y = 41$, spatial discretization $h = 0.05$, time step $\Delta t = 0.001$

and is deterministic if either $\alpha(\omega) < 0.5$ or $\alpha(\omega) > 0.5 \; \forall \omega \in \Omega$. Figure 3.10 shows the evolution of the expected value of the solution, by assuming $\alpha(\omega) < 0.5$. Observe that the solution is a function of the random variable α. By defining as before the parameter-to-solution map $u(\cdot, t, \alpha) : [-1, 1] \to H^1(D)$ and the set $\mathcal{V}(t) = \{u(\cdot, t, \alpha), \alpha \in [-1, 1]\} \subset H^1(D)$, at each $t \in \mathcal{T}$ the solution $u(\cdot, t, \alpha(\omega))$ is in the one dimensional manifold $\mathcal{V}(t)$ for all $\omega \in \Omega$. However, the manifold $\mathcal{V}(t)$ evolves in time, driven by the non linear reaction term, and may feature a complex structure for large times. In the DO approach, we try to approximate such manifold by a linear combination of S modes. The number of basis functions that the DO approximate solution needs to well describe the solution depends obviously on the complexity of the manifold $\mathcal{V}(t)$.

We analyze here the performance of the DO approach. First of all, since the initial condition is deterministic, we have arbitrarily initialized the modes to a set of orthonormal functions. Due to the zero Neumann boundary conditions, we have chosen S orthonormal cosine functions of increasing frequency, i.e. $u_i(x_1, x_2) = k_i \cos(i_1 \pi x_1) cos(i_2 \pi x_2)$ (where k_i is the normalization constant). Figure 3.11 shows the first two deterministic basis functions at different time iterations. Observe that the modes adapt very fast to the structure of the problem and assume values different than zero only around the front of the traveling wave. On the other, by analyzing the evolution of the eigenvalues of the covariance matrix, we see that good levels

87
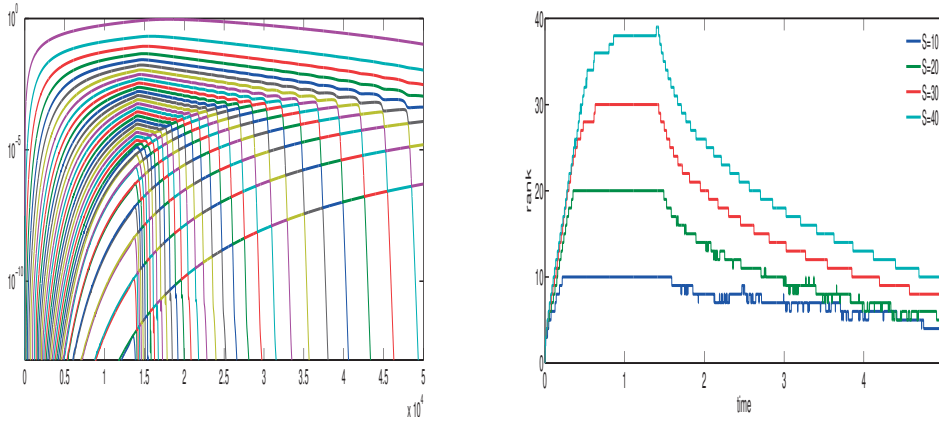
Figure 3.12 – On the left: Evolution of the eigenvalues of the covariance matrix of the DO approximate solution (in logarithmic scale) with $S = 40$. Right: Evolution of the rank of the covariance matrix with number of modes $S = 10/20/30/40$ and excitation rate $\beta = 100$. Discretization parameters: Gauss-Legendre collocation points $N_y = S$, spatial discretization step $h = 0.05$, time step $dt = 0.001$.

of accuracy can be achieved only if a relatively large number of modes is used. Figure 3.12 (left) shows that several eigenvalues reach remarkable values and many of them cross each other. This poses a serious limitation in the use of low-rank formats for this type of problems, which is intrinsically due to the nature of the problem and the structure of the solution. This is confirmed also by the analysis of the effective rank of the DO approximate solution. Figure 3.12 (right) shows the evolution of the rank of the DO approximate solution, for different values of S. Being the rank of the covariance matrix bounded by the number of collocation points that are used in the stochastic discretization, in the plot we consider a number of collocation points at least equal to S. We see that for $S < 40$ the rank tends to reach the saturation level
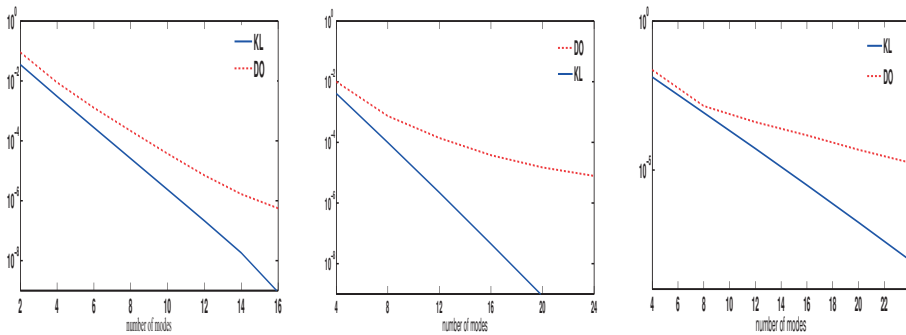
Figure 3.13 – Error in norm $L^2(D) \times L^2(\Omega)$ of the DO (red, dotted line) and the truncated Karhunen-Loève (blue, solid line) approximate solution w.r.t. the number of modes. (log scale). Left: $\beta = 10$, T=3. Middle: $\beta = 100$, $T = 0.5$. Right: $\beta = 100$, $T = 0.7$.Discretization parameters: Gauss-Legendre collocation points $N_y = 41$, spatial discretization $h = 0.05$, time step $\Delta t = 0.001$

S, whereas it does not exceed 39 by using $S > 40$. However that bound is influenced by the space discretization and tends to slightly increase by refining the deterministic discretization parameter. Furthermore, the bound is related to the value of the excitation rate β, which affects the "sharpness" of the front. When the excitation rate is small, e.g $\beta \approx 10$, the maximum rank achieved is relatively smaller (≈ 22). In the latter case indeed the reaction has weaker predominance on the diffusion term and the solution has fewer step gradients. Looking at the solution for larger times, when the front is about to exit the computational domain, the total variation of the solution decreases, and the rank decreases as well. If, on the one hand, these results show the ability of the DO method to capture the effective dimension of the solution at each time step, on the other hand, they also show the need of using a large number of modes. Figure 3.13 shows the $L^2(D) \times L^2(\Omega)$ error of the DO approximate solution (red, dotted line) and of the truncated Karhunen-Loève expansion (blue, solid line) with respect to the number of modes (the errors have been calculated with respect to a reference solution, numerically computed with the Stochastic Collocation method in a fine tensor grid and by using the same discretization parameters, in time and space, chosen for solving the DO system). We observe that in both cases, high levels of accuracy may be achieved only for large values of S. Furthermore, the rate of convergence of the DO method is slower than the one of the truncated Karhunen-Loève expansion. This is due to the fast increasing/decreasing of the eigenvalues of the covariance matrix, that makes them frequently cross each other. As discussed in Section 4, this fact may negatively affect the performance of the DO method. On the other hand, numerical evidence reveals that better performances for these types of problems can be achieved when the stochastic input concerns the initial condition instead of the coefficients of the reaction-operator.

3.6 Conclusion

In this work we established and formalized a link between the DO approximation of PDEs with random initial datum and the MCTDH method proposed for the approximation of deterministic Schr*ö*dinger equations, or the discrete analogue Dynamical Low Rank approximation of evolution matrix or tensor equations. We have reinterpreted the DO approximation as a Galerkin projection onto the tangent space to the manifold $\mathcal{M}_S$ of all rank S functions, at any time instant and in light of the theoretical results developed in [39, 66] for the MCTDH method and the Dynamical Low Rank approximation, we investigated the properties of the manifold $\mathcal{M}_S$ for a linear parabolic equation with random parameters. Specifically we exploited the curvature bounds of $\mathcal{M}_S$ to show that the DO approximation error can be bounded in terms of the best rank S approximation of the solution, at each time instant, under the assumption that the latter is differentiable in time. On the other hand, we have seen that the regularity assumption on the Karhunen-Loève decomposition is actually a necessary condition to maintain an effective control on the DO approximation error. As confirmed by the numerical results, the DO approximation error is properly bounded in terms of best approximation error as long as the eigenvalues of Karhunen-Loève expansion included in the S rank approximation, do not

cross the ones which have been initially omitted. In conclusion our work sets the bases for a theoretical analysis of the DO approximation for random PDEs and provides indication of the effectiveness of the DO method for different types of problems.

4 Dual DO approximation of Navier Stokes equations with random boundary conditions

This Chapter is mainly based on the paper [95] with only minor changes (in particular, for consistency with the rest of the thesis, the stochastic coefficients **Y** are defined here as column vectors instead of row vectors). In this Chapter we propose a method for the strong imposition of random Dirichlet boundary conditions in the Dynamical Low Rank (DLR) approximation of parabolic PDEs, focusing on incompressible Navier-Stokes equations. In particular, in Section 4.3 we start by defining the Dual DO formulation in which the stochastic modes are kept orthonormal, then we show that the collection of all S rank random fields which satisfy a suitable low-rank approximation of the exact boundary condition, admits a structure of differential manifold. We formulate the DLR variational principle in the constrained manifold and we characterize its tangent space by means of the Dual DO formulation which allows us to identify the proper boundary conditions for each (time dependent) deterministic mode and guarantees that the boundary constraint is fulfilled at each time. In Section 4.4 we apply the Dual DO formulation to Navier-Stokes equations with random parameters, including Dirichlet boundary conditions. We conclude with two numerical test cases: the classical benchmark of a laminar flow around a cylinder with random inflow velocity, and a biomedical application simulating blood flow in a realistic carotid artery reconstructed from MRI data with random inflow conditions coming from Doppler measurements.

Introduction

Uncertainty quantification received a lot of attention in the last decades and is nowadays an active research field [134, 120, 79, 123, 75]. Mathematical models and numerical methods for efficient propagation of uncertainties are more and more needed in many application areas, from aerospace and mechanical engineering to life and geosciences. Numerical techniques for uncertainty propagation typically require a lot of problem solves for many values of the uncertain/random parameters and this may result in an unaffordable computational cost for complex applications, mostly if the phenomenon under study is time dependent. In this context, the use of reduced order models is very appealing as they reduce dramatically the computational cost of each solve, provided they guarantee a certain accuracy level. Many

techniques have been developed, especially in the deterministic parametric framework, starting from the "classical" Proper Orthogonal Decomposition (POD) [23, 33] and the Reduced Basis (RB) method [58, 110]. All these techniques are based on the assumption that in many situations the solution manifold can be well approximated by a small number of dominant modes extracted from the covariance matrix of several snapshots precomputed in the offline stage, for different values of the parameters and different time instants. By performing a Galerkin projection into the subspace spanned by the dominant modes the size of the original problem is drastically reduced and the online stage simply consists in low-cost reduced-order simulations for new instances of the input. The drawback of this approach is that the solution manifold at time t, i.e. the collection $\mathcal{U}(t) = \{u(t, \omega), \omega \in \Omega\}$ of all solutions at time t for all parameters $\omega \in \Omega$, may change significantly during the time evolution, which implies that the (fixed in time) reduced basis has to be sufficiently rich to be able to approximate $\mathcal{U}(t)$ at all time. This may lead to a fairly large reduced model thus compromising its efficiency.

An alternative approach that has been proposed in the literature consists in expanding the solution on a fixed basis in the probability space by assuming that the randomness can be accurately parametrized in terms of a finite dimensional vector. For instance, in the Polynomial Chaos (PC) expansion, polynomial basis functions are chosen, which are orthonormal with respect to the underlying probability measure of the input random vector used to parametrize the stochastic space [75]. However, it has been reported in literature [128] that, for certain classes of problems, long time integration might need an increasing number of terms in the expansion to keep an acceptable accuracy level.

To overcome the limitations related to expansions of the solution on a fixed basis, either deterministic or stochastic, here we propose a dynamical low-rank approximation. In the UQ context, this method has been introduced in [116, 117] and is known under the name of Dynamically Orthogonal Field equations (DO). Equivalent formulations of the same approach can be found in [30, 31] (Dynamically Bi-Orthogonal method or DyBO) and [35, 36] (Bi-Orthogonal method or BO). The DO is a reduced order method in which both the spatial and random modes are computed on the fly and are free to evolve in time, thus adjusting at each time to the current structure of the solution. The approximate solution is sought on the manifold $\mathcal{M}_S$ of S-rank functions $u_S(x, t, \omega) = \sum_{i=1}^{S} U_i(x, t) Y_i(t, \omega)$ with both $\{U_i\}$ and $\{Y_i\}$ linearly independent and is obtained by Galerkin projection of the governing equations onto the tangent space of $\mathcal{M}_S$ along the solution trajectory. As the manifold is parametrized in terms of dynamic constraints, one can derive evolution equations for both the deterministic modes U_i and the random modes Y_i, suitable for numerical computation. The same approach was independently proposed in the literature in different fields. In the context of deterministic time-dependent Schrödinger equations, its abstract formulation is known as Dirac-Frenkel time-dependent variational principle [80] and leads to the derivation of the so-called multi-configuration time-dependent Hartree (MCTDH) method [92, 12, 65, 138]. In a finite dimensional setting, the same is known as Dynamical Low Rank approximation [66, 68]. Extensions to tensor formats can be found in [39, 67, 83]. In [96] the link between the DLR, or MCTDH, and the DO method has been exploited to derive a quasi-optimal error bound for the approximation of linear parabolic equations. More precisely the approximation

error is bounded in terms of the best approximation error of the exact solution in $\mathcal{M}_S$, and holds in the largest time interval in which the best rank S approximation remains full rank and continuously differentiable in time.

In this chapter, we focus on the approximation of parabolic PDEs and, in particular, incompressible Navier-Stokes equations, with random Dirichlet boundary conditions. Our interest is motivated by the observation that, in fluid dynamics problems, small variations on inflow boundary conditions can have a strong impact on the dynamics of the flow. Applications can be found both in engineering and biomedical problems. A judicious approximation of the problem by low-rank techniques necessarily has to address the issue concerning which boundary conditions should be satisfied by the approximate solution and if and how the randomness coming from the boundary should be compressed. This problem has not been investigated in depth in the literature so far, at least in the context of dynamically low-rank approximation, and the answer is far from being straightforward. In fact, it is not possible to say "a priori" which parameters have the strongest impact on the dynamics of the solution and at which time. Moreover, no results can be derived by the comparison with the truncated Karhunen-Loève expansion which does not necessary approximate well the solution on the boundary. It is clear that the truncated Karhunen-Loève expansion, being the result of a (volumetric) L^2-projection at a fixed time, is not able to quantify the discrepancy on the boundary and, least of all, evaluate its impact on the dynamics. We mention that in the first formulation of the DO method for random time-dependent PDEs, as introduced in [116], the source of randomness includes boundary terms. The strategy proposed there consists in projecting the Dirichlet boundary conditions onto the subspace spanned by the stochastic modes at each time. However, we observe that this subspace evolves in time and is part of the solution of the approximate problem and not known "a priori". It is then not clear which boundary conditions are actually satisfied by the approximate solution as times evolves and how the randomness arising from the boundary data is taken into consideration. An alternative approach, common in the reduced basis community [61, 110], consists in computing explicitly a lift of the random boundary function, which needs to be written in a separable form, and then solving for the homogeneous part of the solution (zero on the boundary). In such case, the deterministic modes always vanish on the boundary. However, on the one hand, the explicit construction of the lifting may be difficult and time-consuming for time-dependent random boundary conditions, and the quality of the approximation may be influenced by the choice of the lift. Theses issues are reflected in a similar way in the DO approximation and in particular the latter concerns the difficulty in evaluating the loss of information in deriving the reduced order system when the lift is projected in the tangent space.

In this work, we investigate the possibility of strong imposition of the random boundary conditions in the dynamical low-rank approximation. We require that the approximate solution satisfies the same boundary conditions as the exact solution, or a well-controlled approximation of them. To do so we assume that the datum on the boundary is "almost low rank", which is not a too restrictive assumption in our context: since we are looking for an approximate solution u_S of rank S such that $u_S \approx u$, it is reasonable to ask that the boundary value $u_{|\partial D} = g$ is properly approximated in separable form by $g_M = \sum_{i=1}^{M} Z_i(\omega) v_i(t, x)$ with $M \leq S$. In the

context of dynamical low rank approximation, an approximation g_M of the boundary datum g in separable form with $M < S$ terms, will allow us to identify the proper boundary conditions to impose on each deterministic mode at each time for the solution $u_S = \sum_{i=1}^{S} U_i(x, t) Y_i(t, \omega)$. Indeed the reduced system for the evolution of the deterministic modes consists of S coupled PDEs of the same type as the original problem, which have to be completed with suitable boundary conditions. Our strategy consists in seeking for a dynamically low-rank solution in the manifold $\mathcal{M}_S$ of rank S functions constrained to take the approximate value g_M on the boundary. We show, in particular, that, as long as the datum g_M has rank M, the constrained manifold is indeed a manifold and we provide a characterization of its tangent space. To derive a proper set of equations for the deterministic and stochastic modes, we propose a Dual-DO formulation, in which the stochastic modes are kept orthonormal, instead of the deterministic modes as in the original DO formulation of [116]. It turns out that such a formulation is very convenient for the "strong" imposition of random Dirichlet boundary conditions and results in a set of S coupled PDEs for the evolution of the deterministic modes (M of which with non homogeneous boundary conditions) coupled with $S - M$ ODEs for the evolution of the stochastic modes. Also when dealing with the incompressible Navier-Stokes equations, the Dual-DO is also very convenient to include the incompressibility constraint.

The Dual DO method has been tested on two fluid dynamics problems. In the first one, our goal is to test the performance of the Dual DO approximation in the challenging case in which the rank of the solution continues to increase in time. We consider the classical benchmark 2D problem of an incompressible viscous fluid flowing around a cylindrical obstacle in a channel at moderate Reynold numbers Re$\in$ [80, 120]. The challenge of this test is due to the inflow velocity that depends on some random parameters. The patterns of the solutions correspond to flows with random vortex shedding frequency. Intuitively one can imagine the solution manifolds $\mathcal{U}(t)$ as the collection of infinitely many flow patterns which become more and more out of phase, one with respect to the others, as time evolves. The obtained numerical results show good performance of the method, at least in the initial phase, in approximating the whole solution manifold at each time instant with a relatively small number of time evolving modes. However, as one might expect, the performance deteriorates for larger times due to the "phase" issue. To alleviate the problem, we introduce a simple time rescaling based on an empirical linear relation between Reynolds number and shedding frequency considerably improves the performance of the method as it allows to "rephase" all solutions. In this setting, we were able to simulate the transition phase and few shedding periods in the whole range Re$\in$ [80, 120] with good accuracy with $S = 4$ modes.

The second numerical problem addressed in this work aims at testing the possibility of applying the Dynamical Low-Rank method for biomedical applications. Indeed in this field, numerical simulations of parameter dependent PDEs can be used as a virtual platform for the prediction of input/output response of biological values, and the speed up of the computational time is a crucial issue. We consider the problem of simulating blood flow in a realistic carotid artery reconstructed from MRI data, where the inflow boundary conditions are taken as random due to the uncertainty and large errors in Doppler measurements of the inflow velocity profile [50, 111]. The results highlight the remarkable potential of the Dual DO

method for this type of problems.

The paper is organized as follows: in Section 1 we introduce the problem setting and the notations used throughout, in Section 4.2 we recall the DO approach for a general parabolic problem with deterministic boundary conditions, in Section 3 we describe the Dual DO formulation for a second order elliptic operator with random Dirichlet boundary conditions and in Section 4 we apply the Dual DO to the Navier-Stokes equations. Section 5 presents the two numerical tests mentioned above.

4.1   Problem setting and Notation

Let $D \subset \mathrm{R}^d$, $1 \le d \le 3$, be an open bounded domain and $(\Omega, \mathscr{A}, \mathscr{P})$ a complete probability space, where Ω is the set of outcomes, $\mathscr{A}$ a σ-algebra and $P : \mathscr{A} \to [0,1]$ a probability measure. We consider a general time dependent PDE of the type:

$$\dot{u}(x,t,\omega) = \mathscr{L}(u(x,t,\omega),x,t,\omega), \quad x \in D, \ t \in (0,T], \ \omega \in \Omega, \tag{4.1}$$

where $\mathscr{L}$ is a linear or non-linear differential operator, $x \in D$ is the spatial coordinate and t is the time variable in $[0,T]$. For the ease of notation in what follows we omit to write the explicit dependence of $\mathscr{L}$ on (x,t,ω) and use the shorthand notation $\mathscr{L}(u(x,t,\omega))$. The initial state of the system is described by

$$u(x,t=0,\omega) = u_0(x,\omega), \quad x \in D, \ \omega \in \Omega, \tag{4.2}$$

and equation (4.1) is complemented with suitable boundary conditions

$$\mathscr{B}(u(x,t,\omega),\omega) = g(x,t,\omega), \quad x \in \partial D, \ \omega \in \Omega, \ t \in (0,T].$$

Here $\omega \in \Omega$ represents a random elementary event which may affect the operator $\mathscr{L}$ (as e.g. a coefficient or a forcing term), the boundary conditions or the initial conditions. Specifically, in Section 4, we consider a second order deterministic elliptic operator $\mathscr{L}$ completed with random Dirichlet boundary conditions and in Section 4.4 we consider the Navier Stokes equations, with random viscosity and Dirichlet boundary conditions.

We introduce here some notation that will be used throughout. Let $v : \Omega \to \mathbb{R}$ be an integrable random variable; we define the mean of v as:

$$\bar{v} = \mathbb{E}[v] = \int_\Omega v(\omega) d\mathscr{P}(\omega),$$

and the variance as:

$$\mathrm{Var}[v] = \mathbb{E}[(v - \bar{v})^2] = \int_\Omega (v(\omega) - \bar{v})^2 d\mathscr{P}(\omega).$$

We will use the shorthand notation: $v^* = v - \mathbb{E}[v]$, and $L_0^2(\Omega)$ will denote the set of all zero mean, square integrable random variables. Let now $v, u : D \times \Omega \to \mathbb{R}$ be x-indexed random

95

fields. We denote the L^2 inner product in the physical space by:

$$\langle u(\cdot,\omega), v(\cdot,\omega)\rangle = \int_D u(x,\omega)v(x,\omega)dx.$$

We also recall that $L^2(D \times \Omega)$ denotes the space of all square integrable random fields, i.e.:

$$L^2(D \times \Omega) := \left\{ u : D \times \Omega \to \mathbb{R} \text{ s.t. } \int_D \mathbb{E}\left[(u(x,\cdot) - \bar{u}(x))^2\right]dx < \infty \right\}$$

Observe that $L^2(D \times \Omega)$ is isometrically isomorphic to $L^2(D) \otimes L^2(\Omega)$.

A vector-valued random field will be denoted by small bold letters $\mathbf{u} := (u_1,..,u_N)^T$ and is conventionally a column vector. The $L^2(D)$ and the $L^2(D \times \Omega)$ norms are respectively defined as:

$$\|\mathbf{u}(\cdot,\omega)\|^2_{[L^2(D)]^N} := \sum_{i=1}^N \|u_i(\cdot,\omega)\|^2_{L^2(D)} \text{ and } \|\mathbf{u}\|^2_{[L^2(D)]^N \otimes L^2(\Omega)} := \sum_{i=1}^N \|u_i\|^2_{L^2(D) \otimes L^2(\Omega)}.$$

In the following we will denote by $\|\cdot\|$ both the scalar and vector norm in $L^2(D \times \Omega)$. Capital bold letters will be instead used for denoting a vector of deterministic scalar (or vector-valued) functions $\mathbf{U} = (U_1,...,U_S)$ (or $\mathbf{U} = (\mathbf{U}_1,...,\mathbf{U}_S)$ in the case of vector valued functions) which will be written as row vector, and the notion $\ll \mathbf{U},\mathbf{V} \gg$ denotes the $S \times S$ matrix with entries:

$$\ll \mathbf{U},\mathbf{V} \gg_{ij} = \int_D V_i(x)U_j(x)dx$$

(or $\ll \mathbf{U},\mathbf{V} \gg_{ij} = \int_D \mathbf{V}_i(x)^T \mathbf{U}_j(x)dx$ if $\mathbf{U}_i,\mathbf{V}_j$ are vector functions).

Lastly, we recall the well known Karhunen-Loève expansion. Let $u \in L^2(D \times \Omega)$ be a square integrable random field, the covariance function $\text{Cov}_u : \overline{D \times D} \to \mathbb{R}$ is defined as:

$$\text{Cov}_u(x,y) = \mathbb{E}\left[u^*(x,\cdot)u^*(y,\cdot)\right], \quad x,y \in D.$$

and defines a trace class operator $T_u : L^2(D) \to L^2(D)$ as

$$T_u v(\cdot) = \int_D \text{Cov}_u(x,\cdot)v(x)dx, \quad \forall v \in L^2(D); \tag{4.3}$$

Then, u can be written as:

$$u(x,\omega) = \bar{u}(x) + \sum_{i=1}^\infty \sqrt{\lambda_i} Z_i^{KL}(\omega) V_i^{KL}(x)$$

where:

- (λ_i, V_i^{KL}) are respectively the eigenvalues and the ($L^2(D)$-orthonormal) eigenfunctions of the covariance operator T_u,

- Z_i are mutually uncorrelated real-valued random variables given by:

$$Z_i^{KL}(\omega) := \frac{1}{\sqrt{\lambda_i}} \int_D u^*(x,\omega) V_i^{KL}(x) dx \qquad \forall i \in \mathbb{N}^+, \tag{4.4}$$

with zero mean, $\mathbb{E}[Z_i^{KL}] = 0$, and unit variance, $\mathbb{E}[Z_i^{KL} Z_j^{KL}] = \delta_{ij}$.

Assuming that the eigenvalues are sorted in decreasing order, it is well known (see e.g. [78, 60, 49]) that the best L^2-approximation of u^* with S terms written in separable form is given by the truncated Karhunen-Loève expansion:

$$u(x,\omega) \approx u_S^{KL}(x,\omega) := \bar{u}(x) + \sum_{i=1}^{S} \sqrt{\lambda_i} Z_i^{KL}(\omega) V_i^{KL}(x), \tag{4.5}$$

Assuming $\text{Cov}_u \in C^0(\overline{D \times D})$ and D compact, by Mercer's theorem [112], it follows that

$$\lim_{S \to \infty} \sup_{x \in D} \mathbb{E}[(u(x,\cdot) - u_S^{KL}(x,\cdot))^2] = \lim_{S \to \infty} \sup_{x \in D} \sum_{i=S+1}^{\infty} \lambda_i (V_i^{KL}(x))^2 = 0.$$

All the previous definitions can be generalized to a time-varying random field $u(x,t,\omega)$ and in particular the Karhunen-Loève expansion can either be defined at each fixed $t \in [0,T]$:

$$u_S^{KL}(x,t,\omega) = \bar{u}(x,t) + \sum_{i=1}^{S} \sqrt{\lambda_i(t)} Z_i^{KL}(t,\omega) V_i^{KL}(x,t), \ \ \forall S \in \mathbb{N}^+ \tag{4.6}$$

with $\langle V_i^{KL}(\cdot,t), V_j^{KL}(\cdot,t) \rangle = \delta_{ij}$ for all $t \in [0,T]$, or as a global space-time approximation

$$\tilde{u}_S^{KL}(x,t,\omega) = \bar{u}(x,t) + \sum_{i=1}^{S} \sqrt{\lambda_i(t)} \tilde{Z}_i^{KL}(\omega) \tilde{V}_i^{KL}(x,t), \ \ \forall S \in \mathbb{N}^+,$$

provided $u \in L^2(D \times [0,T] \times \Omega)$. In what follows we refer always to (4.6) as the best S-terms approximation of a space-time random field.

Let us also define the Stiefel manifold $St(S, \mathcal{H})$, for a general Hilbert space $\mathcal{H}$, as the set of orthonormal frames of S vectors in $\mathcal{H}$, i.e.:

$$St(S, \mathcal{H}) = \left\{ \mathbf{V} = (V_1, ..., V_S) : V_i \in \mathcal{H} \text{ and } < V_i, V_j >_{\mathcal{H}} = \delta_{ij} \ \forall i,j = 1, ..., S \right\}$$

where $< \cdot, \cdot >_{\mathcal{H}}$ is the inner product in $\mathcal{H}$. We denote by $\mathcal{G}(S, \mathcal{H})$ the Grassmann manifold of dimension S that consists of all the $S-$dimensional linear subspaces of $\mathcal{H}$. Observe that the truncated Karhunen-Loève expansion can be characterized as:

$$u_S^{KL}(x,t,\omega) = \Pi_{\mathcal{V}_S^{KL}(t) \otimes \mathcal{Z}_S^{KL}(t)} \left[u_S(x,t,\omega) \right] \tag{4.7}$$

where Π is the $L^2(D \times \Omega)$ projector and $\mathcal{V}_S^{KL}(t) \in \mathcal{G}(S, L^2(D))$, $\mathcal{Z}_S^{KL}(t) \in \mathcal{G}(S, L^2(\Omega))$ coincide respectively to the span of the first S deterministic and stochastic modes: $(V_1^{KL}, ..., V_S^{KL}) \in St(S, L^2(D))$ and $(Z_1^{KL}, ..., Z_S^{KL}) \in St(S, L^2(\Omega))$, in the Karhunen-Loève expansion (4.6).

However we would like to emphasize that, in our context, the Karhunen-Loève decomposition (4.6) of the solution to problem (4.1), as well as the L^2-orthogonal projector $\Pi_{\mathcal{V}^{KL} \otimes \mathcal{Z}^{KL}}$ in (4.7), are not available in practice. In other words the optimality of the Karhunen-Loève approximation is suitable only for the purpose of analysis, since it provides a lower bound for the approximation error of low rank methods.

4.2   Dynamical Low rank methods

The Dynamical Low-rank approach [66, 80] is a reduced order method according to which the solution of the governing equation is approximated in a low dimensional manifold of functions with fixed rank, written in separable form. The peculiarity of this reduced basis approach relies on the fact that both the deterministic modes and the stochastic coefficients can evolve in time and are thus able to dynamically adapt to the features of the solution. The approximate solution is obtained by performing a Galerkin projection of the governing equations onto the (time-varying) tangent space to the approximation manifold along the solution trajectory. Let us assume that the solution $u(\cdot, t, \omega)$ to problem (4.1) is in a certain Hilbert space $\mathcal{H} \subset L^2(D)$ for (almost) all $t \in [0, T]$ and $\omega \in \Omega$ and that $\mathcal{L}(u) \in \mathcal{H}'$ for all $u \in \mathcal{H}$ and almost everywhere in $[0, T] \times \Omega$. Moreover let us define S *rank random field* any function $u_S \in \mathcal{H} \otimes L^2(\Omega)$ which can be expressed as a sum of S (and not less than S) linearly independent deterministic modes combined with S linearly independent stochastic modes.

Definition 4.2.1. *We define $\mathcal{M}_S \subset \mathcal{H} \otimes L^2(\Omega)$ the manifold of all the S rank random fields, i.e.:*

$$\mathcal{M}_S = \left\{ u_S \in \mathcal{H} \otimes L^2(\Omega) : u_S = \sum_{i=1}^{S} U_i Y_i \mid \begin{array}{l} span(U_1, ..., U_S) \in \mathcal{G}(S, \mathcal{H}), \\ span(Y_1, ..., Y_S) \in \mathcal{G}(S, L^2(\Omega)) \end{array} \right\} \tag{4.8}$$

Observe that the definition of S rank random field can be characterized in several different ways. We recall in the following box few of the many possible representations that have been proposed and used in literature. For simplicity we describe the different options for time independent random fields.

Representations of S rank random field:

- *Double-Orthogonal decomposition* (used e.g. in [66, 68]), thereafter named DDO:

$$u_S(x,\omega) = \sum_{i=1}^{S} \sum_{j=1}^{S} \mathbf{A}_{ij} Z_i(\omega) V_j(x) = \mathbf{VAZ} \tag{4.9}$$

where:

  - $\mathbf{A}_{ij} \in \mathbb{R}^{S \times S}$ is a full rank matrix,
  - $\mathbf{V}$ is a row vector of S $L^2(D)$−orthonormal deterministic functions,
  - $\mathbf{Z}$ is a column vector of S $L^2(\Omega)$−orthonormal random variables.

- *Decomposition with orthonormal deterministic modes* (used e.g. in [116, 117]), thereafter named DO:

$$u_S(x,\omega) = \sum_{i=1}^{S} \tilde{Y}_i(\omega) \tilde{U}_i(x) = \tilde{\mathbf{U}}\tilde{\mathbf{Y}} \tag{4.10}$$

where:

  - $\tilde{\mathbf{U}} = \mathbf{V}$ is a row vector of $L^2(D)$−orthonormal deterministic functions,
  - $\tilde{\mathbf{Y}} = \mathbf{ZA}$ is a column vector of S linearly independent random variables, hence with full rank covariance matrix $\mathbf{C} = \mathbb{E}[\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}]$.

- *Decomposition with orthonormal stochastic modes* (see Section 4.3), thereafter named Dual DO:

$$u_S(x,\omega) = \sum_{i=1}^{S} Y_i(\omega) U_i(x) = \mathbf{UY} \tag{4.11}$$

where:

  - $\mathbf{U} = \mathbf{VA}$ is a row vector of S linearly independent deterministic functions. Namely, $\mathbf{M} \in \mathbb{R}^{S \times S}$, defined as $\mathbf{M}_{ij} = \langle U_j, U_i \rangle$, is a full rank matrix.
  - $\mathbf{Y} = \mathbf{Z}$ is a column vector of S $L^2(\Omega)$−orthonormal random variables.

In this chapter we adopt the *decomposition with orthonormal stochastic modes* that turns out to be more suitable to approximate the incompressible Navier Stokes equations with random Dirichlet boundary conditions.

Observe, however, that none of the previous formats leads to a unique representation of u_S. Namely, it is always possible to rewrite u_S in the same format but with a different set of bases. This implies that the Dynamical Low-Rank solution (DLR solution), or generally any arbitrary continuously differentiable path $t \to u_S(t)$ from $[0, T]$ to $\mathcal{M}_S$, is not uniquely described in terms of time-dependent bases, whatever the format in which it is represented is. However, the uniqueness of the representation is recovered by imposing dynamic constraints in the

evolution of the bases. These constraints can be formally derived by exploiting the geometrical differential structure of the approximation manifold, see Section 4.3 and [7, 90, 80].

4.2.1 DLR Variational Principle

We introduced here the Dynamical Low Rank (DLR) approach for a general problem (4.1), all details concerning the boundary conditions have been postponed to Section 4.3.2.

Consider problem (4.1): the solution u describes a path $t \to u(t)$ from $[0, T]$ in $\mathcal{H} \otimes L^2(\Omega)$. The idea behind the DLR approach is to approximate this curve $t \to u(t) \approx u_S(t)$ by dynamically constraining the time derivative $\dot{u}_S$ to be in the tangent space to the manifold $\mathcal{M}_S \subset \mathcal{H} \otimes L^2(\Omega)$ at $u_S(t)$ by Galerkin projection of the governing equation (4.1). Precisely, the DLR variational principle for problem (4.1) reads as follows:

DLR Variational Principle. *At each $t \in (0, T]$, find $u_S(t) \in \mathcal{M}_S$ such that:*

$$\mathbb{E}\left[\langle \dot{u}_S(\cdot, t, \cdot) - \mathcal{L}(u_S(\cdot, t, \cdot)), v \rangle\right] = 0, \quad \forall v \in \mathcal{T}_{u_S(t)}\mathcal{M}_S \tag{4.12}$$

where $\mathcal{T}_{u_S(t)}\mathcal{M}_S$ is the tangent space to $\mathcal{M}_S$ at $u_S(t)$.

If $\mathcal{L}(u_S(\cdot, t, \cdot))$ is in the tangent space itself at $u_S(t)$ for any $u_S \in \mathcal{M}_S$, and at any time, and u_0 is a S rank function, then the DLR approximation recovers the exact solution. If u_0 is not S rank and the DLR method is initialized with its best S rank approximation, u_{0S}^{KL}, then the DLR solution coincides with the truncated Karhunen-Loève expansion (i.e. the best S rank approximation), under the assumption that the eigenvalues considered in the approximation of u_0 do not cross the ones that have been omitted at initial time [96].

Observe that the variational principle in (4.12) does not depend on the parametrization of the manifold $\mathcal{M}_S$, as long as the solution is full rank. Specifically, the tangent space $\mathcal{T}_{u_S(t)}\mathcal{M}_S$ is time dependent and depends only on $u_S(t)$ and not on its representation. The variational principle in (4.12) provides indeed a unified formulation for the DO method, as proposed by Sapsis [116], and similar approaches proposed in the literature, including the DyBO method [30] and the DDO method [39]. We refer to [96] for further details.

In order to numerically compute the approximate solution, one needs to uniquely characterize u_S in terms of deterministic and stochastic bases (modes). This is achieved by locally characterizing the manifold by means of a parametrization of the tangent space. This is detailed in the next section for the Dual DO formulation with orthonormal stochastic modes.

4.3 Dual DO formulation

We have seen that from the variational point of view, the DLR approximate solution $u_S \in \mathcal{M}_S$ is defined as a solution of the variational principle (4.12) at each time. However to numerically compute u_S, we need to parametrize the tangent space, hence the manifold, in terms of local charts, corresponding in our context to the deterministic and stochastic modes. Once a

parametrization has been fixed, one can easily derive a set of equations that uniquely describe the dynamics of both the deterministic and stochastic modes. We emphasize that problem (4.12) leads to different sets of equations depending on the parametrization of the tangent space and any of these parametrizations leads to a different reduced order system.

Here we adopt the two fields formulation (4.11) in which we assume the stochastic modes to be orthonormal. We will refer to it as *Dual DO Formulation* (as opposed to (4.10) which keeps the deterministic modes orthonormal and was originally proposed in [116]). This representation turns out to be computationally more efficient and more suitable for dealing with random Dirichlet boundary conditions and solenoidal constraints. We define:

$$B(S, \mathcal{H}) = \{\mathbf{U} = (U_1, ..., U_S) : U_i \in \mathcal{H} \text{ with } \mathbf{M}_{ij} = <U_i, U_j> \text{ s.t. } \text{rank}(\mathbf{M}) = S\}$$

and the map:

$$
\begin{aligned}
\tilde{\pi}: \quad & \left(B(S, \mathcal{H}), St(S, L^2(\Omega)) \right) && \to \mathcal{M}_S \subset \mathcal{H} \otimes L^2(\Omega) \\
& (\mathbf{U}, \mathbf{Y}) && \mapsto \tilde{\pi}(\mathbf{U}, \mathbf{Y}) = \sum_{i=1}^S U_i Y_i =: u_S
\end{aligned}
\tag{4.13}
$$

The image of $\tilde{\pi}$ is the manifold of S rank random fields $\mathcal{M}_S$ defined in (4.8). Observe that:

- the DLR variational principle (4.12) is defined in $\mathcal{M}_S$ while we want to write the DLR approximate solution in terms of $(\mathbf{U}, \mathbf{Y}) \in B(S, \mathcal{H}) \times St(S, L^2(\Omega))$;

- the map $\tilde{\pi}$ is not injective, indeed for any orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{S \times S}$, $\tilde{\pi}(\mathbf{UQ}, \mathbf{Q}^T\mathbf{Y}) = \tilde{\pi}(\mathbf{U}, \mathbf{Y})$.

The uniqueness of the representation (4.13) can be recovered in terms of unique decomposition in tangent space by imposing the following Gauge constraint [40, 90]:

$$\mathbb{E}[\delta Y_i Y_j] = 0 \qquad \forall i, j = 1, ..., S,
\tag{4.14}$$

which leads to the following parametrization of the tangent space at $u_S = \sum_{i=1}^S U_i Y_i$ as [93, 66]:

$$
\begin{aligned}
\mathcal{T}_{u_S}\mathcal{M}_S \quad = \Big\{ & \dot{v} = \sum_{i=1}^S \left(\delta U_i Y_i + U_i \delta Y_i \right) && \in \mathcal{H} \otimes L^2(\Omega), \text{ with } \delta U_i \in \mathcal{H}, \\
& \delta Y_i \in L^2(\Omega) \text{ s.t. } \mathbb{E}[\delta Y_i Y_j] = 0 \; \forall i, j = 1, ..., S \Big\}.
\end{aligned}
\tag{4.15}
$$

Finally the variational problem (4.12) can be rewritten in terms of evolution equations for $(\mathbf{U}, \mathbf{Y})$.

Proposition 4.3.1. *Let $(\mathbf{U}(t), \mathbf{Y}(t)) \in B(S, \mathcal{H}) \times St(S, L^2(\Omega))$ be a solution of the following system:*

$$
\begin{cases}
\dfrac{\partial U_i(x, t)}{\partial t} = \mathbb{E}\left[\mathcal{L}(u_S(x, t, \cdot)) Y_i(t, \cdot) \right] & i = 1, \cdots, S \tag{4.16} \\[4mm]
\displaystyle\sum_{i=1}^S \mathbf{M}_{ji}(t) \dfrac{\partial Y_i(t, \omega)}{\partial t} = \Pi_{\mathcal{Y}}^{\perp} \langle \mathcal{L}(u_S(\cdot, t, \omega)), U_i(\cdot, t) \rangle & j = 1, \cdots, S \tag{4.17}
\end{cases}
$$

then $u_S(t) = \tilde{\pi}(\mathbf{U}(t), \mathbf{Y}(t)) \in \mathcal{M}_S$ satisfies the DLR variational principle (4.12) at any $t \in [0, T]$.

Note the symmetry with the DO system proposed in [116].

4.3.1 Isolating the mean

In our context of partial differential equations with random parameters, since we are usually interested in computing the statistics of the solution, it may be worth approximating separately the mean of the solution. This is achieved by adopting a slightly different definition of S-rank random field and leads to an approximation closer to the Karhunen-Loève expansion of the solution (4.5) where the mean is treated separately in the expansion. The idea of isolating the mean and the corresponding DO formulation was introduced in [116] and adopted in [117],[96],[30]. We detail here only the Dual DO formulation and re-define S rank random field as follows.

Definition 4.3.1. *We call S rank random field (with the isolating mean format) any function that can be exactly expressed as:*

$$
\begin{aligned}
u_S \;&=\; \bar{u}_S &&+\; \textstyle\sum_{i=1}^{S} U_i Y_i \\
&=\; U_0 Y_0 &&+\; \textstyle\sum_{i=1}^{S} U_i Y_i &&=\; \mathbf{U}\mathbf{Y}^{\mathrm{T}}
\end{aligned}
\tag{4.18}
$$

where:

- $\mathbf{Y}$ *is a column vector of $S + 1$ $L^2(\Omega)$-orthonormal random variables such that $Y_0 = 1$ and $\mathbb{E}[Y_i] = 0$ for all $i = 1, ..., S$,*

- $U_1, ..., U_S$ *are linearly independent deterministic functions.*

One can think that the difference with respect to definition (4.11) consists in fixing the first random variable to be constant ($Y_0 = 1$), with the zero mean condition of the remaining random variables coming simply from the orthonormality of the random modes. However, observe that (4.18) is not necessarily a $S + 1$ rank function since $\bar{u}_S$ is not assumed to be linearly independent of $U_1, ..., U_S$, or more precisely the subspace spanned by $\mathbf{U}$ does not have necessarily dimension $S + 1$ (*at most* dimension $S + 1$ and at least S). According to the new definition of S rank random field given in (4.18) the Dual DO system derived in (4.16)-(4.17) becomes:

$$
\left\{
\begin{aligned}
&\dot{U}_0(x, t) = \mathbb{E}[\mathscr{L}(\tilde{u}_S(x, t, \cdot))] && \text{(4.19)} \\
&\dot{U}_i(x, t) = \mathbb{E}[\mathscr{L}(\tilde{u}_S(x, t, \cdot)) Y_i(t, \cdot)] && i = 1, \cdots, S && \text{(4.20)} \\
&\sum_{j=1}^{S} \mathbf{M}_{ji}(t) \dot{Y}_j(x, t) = \Pi_{\tilde{\mathcal{Y}}}^{\perp} \langle \mathscr{L}^*(\tilde{u}_S(\cdot, t, \omega), \omega), U_i(\cdot, t) \rangle && i = 1, \cdots, S && \text{(4.21)} \\
&\quad\quad = \Pi_{\tilde{\mathcal{Y}}}^{\perp} \langle \mathscr{L}(\tilde{u}_S(\cdot, t, \omega), \omega), U_i(\cdot, t) \rangle && i = 1, \cdots, S && \text{(4.22)}
\end{aligned}
\right.
$$

where $\mathcal{Y} = span(Y_1,...,Y_S)$, $\mathcal{L}^*(\bar{u}_S(x,t,\omega),\omega) := \mathcal{L}(\bar{u}_S(x,t,\omega),\omega) - \mathbb{E}[\mathcal{L}(\bar{u}_S(x,t,\cdot))]$, $\tilde{\mathcal{Y}} = span(Y_0,...,Y_S)$ and $\mathbf{M}_{ij} = \langle U_j, U_i \rangle$, $i,j = 1,...,S$.

Remark 4.3.1. *Observe that definition (4.18) does not guarantee the optimality in the manifold of function with rank $S+1$. This is due to the fact that we do not assume U_0 linearly independent of $U_1,...,U_S$ and so (4.18) may have a deficient rank. To clarify this point, just consider the function $u(x,\omega) = x + \alpha(\omega)x$, with $\mathbb{E}[\alpha] \neq -1$. This is clearly a 1-rank function, but representation (4.18) would require 2 modes.*

The previous remark may appear quite obvious when we only want to isolate the mean, but it turns out to be a crucial point when the number of constraints is larger. In the following, we investigate the latter situation. For this purpose, the main questions that we need to address are: how to define the low-rank manifold with constraints and which is the best approximation in this manifold.

4.3.2 Dual DO under boundary constraints

We now explicitly assume that $\mathcal{L}$ in (4.1) is a second order elliptic operator of the form $\mathcal{L}(u) = -div(A(x,\omega)\nabla u) - b(x,\omega) \cdot \nabla u + c(x,\omega)u - f(x,t,\omega)$ where $A_{ij}(x,\omega), b_i(x,\omega), c(x,\omega)$, $i,j = 1,...,d$, are bounded random variables in the open bounded Lipschitz domain $D \subset \mathbb{R}^d$ and under the assumptions that $A(x,\omega)$ is uniformly coercive almost surely and $f \in L^2([0,T], L^2(D \times \Omega))$. The problem is set in $H^1(D) \otimes L^2(\Omega)$ and completed with Dirichlet boundary conditions $u_{|\partial D} = g$. If the boundary condition is deterministic it is reasonable to adopt formulation (4.18) in which the first deterministic mode, that approximates the mean, is required to fulfill the constraint on the boundary, while all other modes satisfy homogeneous conditions:

- $U_0(x,t) = g(x,t)$ for $x \in \partial D$,

- $U_i(x,t) = 0$ for $x \in \partial D$ and for all $i = 1,...,S$.

This is consistent with the Karhunen-Loève decomposition given in (4.5), for which we have:

$$
\begin{aligned}
\lambda_i V_i^{KL}(y)_{|\partial D} &= \left[\int_D \text{Cov}_u(x,y) V_i^{KL}(x)dx \right]_{|y \in \partial D} \\
&= \int_D \mathbb{E}\left[u^*(x,\cdot) u^*(y,\cdot)_{|y \in \partial D} \right] V_i^{KL}(x)dx = 0 \qquad \forall i \in \mathbb{N}^+
\end{aligned}
\tag{4.23}
$$

since $\bar{u}_{|\partial D} = \mathbb{E}[u]_{|\partial D} = g$ and $u^*_{|\partial D} = u_{|\partial D} - \bar{u}_{|\partial D} = 0$. The case in which the boundary data are random is more cumbersome. The first question to be addressed is which boundary conditions should be satisfied by a general low rank approximate solution. One can easily verify that the truncated Karhunen-Loève expansion does not necessarily satisfy the same boundary conditions satisfied by the exact solution. Consider for example the following toy problem in

$D \equiv (0, 2\pi)$:

$$\begin{cases} \dot{u}(x,t,\omega) - \Delta u(x,t,\omega) = 0 \\ u(0,t,\omega) = u(2\pi, t, \omega) = \alpha(\omega)e^{-t} \\ u(x,0,\omega) = \alpha(\omega)\cos(x) + \beta(\omega)\sin(x) \end{cases} \qquad (4.24)$$

where α, β are two uncorrelated zero mean random variables such that $\mathbb{E}[\alpha^2] < \mathbb{E}[\beta^2]$. The exact solution is $u(x,t,\omega) = e^{-t}(\alpha(\omega)cos(x) + \beta(\omega)sin(x))$. It is clear that the Karhunen-Loève approximate solution of rank 1 is $u_1^{KL}(x,t,\omega) = \beta(\omega)e^{-t}\sin(x)$ and $u_{1|\partial D}^{KL} \neq u_{|\partial D}$. Generally the values of the truncated Karhunen-Loève expansion on the boundary are unknown. Secondly, in the context of Dynamical Low Rank approximation, we need to specify the boundary conditions to impose on each deterministic mode $U_1, ..., U_S$. We remind that the Dual DO reduced system consists of dynamic differential equations for all the factors in (4.11). In particular, in the equations (4.16)-(4.17), boundary conditions for each deterministic mode U_i are needed to have a well posed problem.

Dual DO under random boundary constraints

Our strategy consists in enforcing that the low rank approximation satisfies the same boundary conditions as those of the exact solution. This is motivated by the fact that we can not say "a priori" which parameters have the strongest impact on the dynamics and at which time the dynamic of the solution is influenced by the uncertain parameters in the boundary data. It may therefore be important to impose these constraints as accurately as possible. To do so we assume that the datum on the boundary is "almost low rank", which is not a too restrictive assumption in our context: since we are looking for an approximate solution $u_S \approx u$ is reasonable to ask that $u_{|\partial D}$ is properly approximated by a function of rank at most S. We start considering Dirichlet boundary conditions that do not depend on time.

Assumption 3. *The boundary function g can be properly approximated on the manifold of M-rank functions for some $M \leq S$:*

$$u(x,t,\omega) = g(x,\omega) \approx g_M(x,\omega) = \sum_{i=1}^{M} v_i(x) Z_i(\omega) \quad \forall x \in \partial D, \ a.s. \qquad (4.25)$$

with:

- $\mathbb{E}[Z_i Z_j] = \delta_{ij}$,

- $v_1, ..., v_M$ *linearly independent.*

We denote by R the difference $S - M$ and by $\mathcal{Z}$ the subspace spanned by $\{Z_1, ..., Z_M\}$. Then in the DO formulation (4.12), we impose strongly condition (4.25). Precisely we ask:

$$u_S(x,t,\omega) = g_M(x,\omega) \quad \forall x \in \partial D, \ a.s. \qquad (4.26)$$

For the sake of clarity, we start considering a general Dual DO representation, as defined in (4.11). The similar formulation with the isolation of the mean is discussed in Remark 4.3.2.

Definition 4.3.2. *A S rank random field under constraint (4.26) is a S rank function that satisfies the boundary condition in (4.25) and can be written as:*

$$u_S^{g_M}(x,\omega) = \sum_{i=1}^{S} U_i(x) Y_i(\omega) = \mathbf{U}\mathbf{Y} \tag{4.27}$$

with:

- *$u_{S|\partial D} = g_M$ a. s.,*

- *$U_1, ..., U_S$ linearly independent deterministic functions.*

- *$Y_1, ..., Y_S$ uncorrelated random variables.*

We denote by $\mathcal{M}_S^{g_M}$ the set of all the S rank random fields under constraint (4.26).

For the sake of notation, we omit the superscript in $u_S^{g_M}$ in the following. However observe that definition (4.27) strongly depends on the boundary conditions. Our first aim is to show that $\mathcal{M}_S^{g_M}$ is indeed a manifold. and precisely we aim to show that $\mathcal{M}_S^{g_M}$ is the *manifold* of all random fields of rank S that satisfy the same boundary condition as the solution, up to the approximation in (4.25). We now claim that any function in $\mathcal{M}_S^{g_M}$ can be written in terms of the random modes $Z_1, ..., Z_M$ in (4.25) and $R = S - M$ "free" random variables, in the orthogonal complement of $\mathcal{Z}$.

Lemma 4.3.1. *Let $\mathcal{M}_{R,M}$ denote the manifold of all the functions $u_{R,M}$ written as:*

$$u_{R,M}(x,\omega) = \sum_{i=1}^{R} U_i(x) Y_i(\omega) + \sum_{i=1}^{M} V_i(x) Z_i(\omega) \tag{4.28}$$

where we assume:

- *$R + M = S$,*

- *$u_{R,M}(x,\omega) = g_M(x,\omega) = \sum_{i=1}^{M} v_i(x) Z_i(\omega)$ for $x \in \partial D$ a.s.,*

- *all the random variables are mutually $L^2(\Omega)$-orthonormal:*

    - *$\mathbb{E}[Z_i Z_j] = \delta_{ij}$ for all $i, j = 1, ..., M$;*

    - *$\mathbb{E}[Y_i Y_j] = \delta_{ij}$ for all $i, j = 1, ..., R$;*

    - *$\mathbb{E}[Z_i Y_j] = 0$ for all $i = 1, ..., M$ and for all $j = 1, ..., R$.*

- *$U_1, ..., U_R$ are linearly independent.*

Then, the set $\mathcal{M}_S^{g_M}$ coincides with $\mathcal{M}_{R,M}$, hence it is a manifold.

To prove this lemma we need a preliminary result:

Lemma 4.3.2. *Given a random field $u_{R,M} \in \mathcal{M}_{R,M}$ defined as in (4.28), it holds:*

- $V_{i|\partial D} = v_i$ *a.s. for all $i = 1, ..., M$;*

- $U_{i|\partial D} = 0$ *a.s. for all $i = 1, ..., R$;*

and in particular $\{U_1, ..., U_R, V_1, ..., V_M\}$ are linearly independent.

Proof. To verify the values of U_i and V_i on the boundary is enough to observe that:

$$\begin{aligned}
V_{i|\partial D} &= \mathbb{E}[u_{R,M|\partial D} Z_i] = \mathbb{E}[g_M Z_i] = \sum_{j=1}^{M} v_j \mathbb{E}[Z_j Z_i] = v_i \\
U_{i|\partial D} &= \mathbb{E}[u_{R,M|\partial D} Y_i] = \mathbb{E}[g_M Y_i] = \sum_{j=1}^{M} v_j \mathbb{E}[Z_j Y_i] = 0
\end{aligned}$$

where we have used the fact that $u_{R,M|\partial D} = g_M$ a.s. and the random modes are mutually orthogonal. Then the fact that $v_1, ..., v_M$ are linearly independent implies that $\{U_1, ..., U_R, V_1, ..., V_M\}$ are linearly independent. $\qquad\square$

Proof (Lemma 4.3.1). The fact that $\mathcal{M}_{R,M} \subseteq \mathcal{M}_S^{g_M}$ follows directly from Lemma 4.3.2. Now we need to show that $\mathcal{M}_S^{g_M} \subseteq \mathcal{M}_{R,M}$. Let $u_S \in \mathcal{M}_S^{g_M}$, we have that:

$$\begin{aligned}
u_S &= \sum_{i=1}^{S} U_i Y_i \\
&= \sum_{i=1}^{S} U_i \Pi_{\mathcal{Z}} Y_i + \sum_{i=1}^{S} U_i \Pi_{\mathcal{Z}}^{\perp} Y_i \\
&= \sum_{j=1}^{M} \left(\sum_{i=1}^{S} U_i \mathbb{E}[Y_i Z_j] \right) Z_j + \sum_{i=1}^{S} U_i \Pi_{\mathcal{Z}}^{\perp} Y_i \\
&= \sum_{j=1}^{M} V_j Z_j + \sum_{i=1}^{S} U_i \Pi_{\mathcal{Z}}^{\perp} Y_i
\end{aligned}$$

Since $u_{S|\partial D} = g_M$ and Z_i are orthogonal, we necessary have that $V_{j|\partial D} = v_j$. Moreover, the fact that v_i are linearly independent implies that $V_1, ..., V_M$ are linearly independent. We can write:

$$\langle v_j, u_{S|\partial D} \rangle_{L^2(\partial D)} = \langle v_j, \sum_{i=1}^{M} v_i \rangle_{L^2(\partial D)} Z_i, \tag{4.29}$$

that implies

$$\sum_{l=1}^{S} \langle v_j, U_{l|\partial D} \rangle_{L^2(\partial D)} Y_l = \langle v_j, \sum_{i=1}^{M} v_i \rangle_{L^2(\partial D)} Z_i. \tag{4.30}$$

Let $\mathbf{B} \in \mathbb{R}^{M \times M}$, $\mathbf{C} \in \mathbb{R}^{M \times S}$ denote respectively:

$$\mathbf{B}_{ij} = \langle v_j, v_i \rangle_{L^2(\partial D)} \text{ and } \mathbf{C}_{ij} = \langle v_i, U_{j|\partial D} \rangle_{L^2(\partial D)}.$$

Then, equation (4.30) can be rewritten as:

$$\mathbf{CY} = \mathbf{MZ} \quad \Rightarrow \mathbf{Z} = \mathbf{M}^{-1}\mathbf{CY} \tag{4.31}$$

where we use the fact that $v_1, ..., v_M$ are linearly independent. This shows that:

$$\text{span}(Z_1, ..., Z_M) \subset \text{span}(Y_1, ..., Y_S).$$

In particular there exist $(\tilde{Y}_1, ..., \tilde{Y}_R)$ orthonormal random variables, orthogonal to the subspace spanned by $(Z_1, ..., Z_M)$ such that:

$$\text{span}(Y_1, ..., Y_S) = \text{span}(Z_1, ..., Z_M) \oplus \text{span}(\tilde{Y}_1, ..., \tilde{Y}_R).$$

Hence u_S can be written according to (4.28), as:

$$u_S = \sum_{j=1}^{M} V_j Z_j + \sum_{i=1}^{R} \tilde{U}_i \tilde{Y}_i$$

where the linear independence of $\tilde{U}_1, ..., \tilde{U}_R$ follows form the fact that u_S is a S-rank random field.

$\square$

In view of Lemma 4.3.1 we can exploit representation (4.28) that enables us to derive the boundary conditions for each mode in the DO reduced system. In particular any $u_S \in \mathcal{M}_S^{g_M}$ is written as:

$$u_S = \sum_{i=1}^{S} U_i Y_i$$

where

- $\{Y_1, .., Y_S\}$ are $L^2(\Omega)$-orthonormal random variables;

- $\{U_1, ..., U_R\}$ are linearly independent;

- $U_{i|\partial D} = 0$ for all $i = 1, ..., R$ ($R = S - M$) and $U_{i|\partial D} = v_i$ for all $i = R+1, ..., S$;

- $Y_i = Z_i$ for all $i = R+1, ..., S$;

Observe that:

$$\mathcal{M}_S^{g_M} \cong \mathcal{M}_R \oplus \mathcal{M}_M^{g_M} \tag{4.32}$$

where we recall that:

$$\mathcal{M}_M^{g_M} = \left\{ u = \sum_{i=1}^{M} U_i Z_i \,\Big|\, u_{|\partial D} = g_M, \, U_i \in H^1(D) \text{ linearly independent} \right\} \subset H^1(D) \otimes L^2(\Omega)$$

and $\mathcal{M}_R$ is the manifold embedded in $H_0^1(D) \otimes \mathcal{Z}^\perp$ of all the random fields of rank R. Observe that if $M = 0$ we recover the standard formulation without constraints according to which $\mathcal{M}_S$ is the manifold of all the random fields of rank S that vanish on the boundary. On the other hand, if $M = S$, $\mathcal{M}_S^{g_S}$ reduces to the S dimensional affine subspace, spanned by $Z_1, ..., Z_M$ and the Dynamical Low Rank approximation reduces to a standard Galerkin projection.

We are now ready to define the Dynamical Low Rank variational principle in $\mathcal{M}_S^{g_M}$, i.e. the manifold of all the S rank random fields that satisfy the (approximate) boundary conditions (4.25). Observe that, in light of (4.32), for any $u_S \in \mathcal{M}_S^{g_M}$, we have that:

$$\mathcal{T}_{u_S} \mathcal{M}_S^{g_M} \cong \mathcal{T}_{u_R} \mathcal{M}_R \oplus \left(\mathcal{Z} \otimes H_0^1(D) \right) \tag{4.33}$$

where $u_R = \Pi_{\mathcal{Z}}^\perp [u_S]$. Assuming that we adopt a parametrization of the manifold such that $u_S \in \mathcal{M}_S^{g_M}$ is represented as $\sum_{i=1}^S U_i Y_i$ where the last M random variables coincide with $Z_1, .., Z_M$ in (4.25), then the tangent space can be parametrized as:

$$\mathcal{T}_{u_S} \mathcal{M}_S^{g_M} \cong \left\{ \dot{u} = \sum_{i=1}^S \left(\delta U_i Y_i + U_i \delta Y_i \right) \in H_0^1(D) \otimes L^2(\Omega) \text{ s.t.} \quad \begin{aligned} \mathbb{E}[\delta Y_i Y_j] = 0 & \quad \forall i, j = 1, ..., S \\ \delta Y_i = 0 \text{ a.s.} & \quad \forall i = R+1, ..., S \end{aligned} \right\}$$
$$\tag{4.34}$$

This construction of constrained approximation manifolds can be generalized to Dirichlet boundary conditions which depend on time. In this case the decomposition (4.25) is time dependent and the approximation manifold changes in time: $\mathcal{M}_S^{g_M(t)} \cong \mathcal{M}_R \oplus \mathcal{M}_M^{g_M(t)}$. The tangent space is defined at each fixed time, according to (4.34). Formally the DLR variational principle reads the same as in (4.12). What changes is the definition of the manifold. We project the governing equation into the tangent space to $\mathcal{M}_S^{g_M(t)}$ at $u_S(t)$ at each time where now $\mathcal{M}_S^{g_M(t)}$ is the manifold constrained to $g_M(t)$ which may change in time, hence the approximate solution $u_S(t)$ automatically satisfies the Dirichlet boundary conditions of the original problem.

According to the parametrization of the tangent space in (4.34) the reduced order system for the Dual DO formulation under random boundary constraints becomes:

$$
\begin{cases}
\dot{U}_i(x,t) = \mathbb{E}\left[\mathscr{L}(u_S(x,t,\cdot))\,Y_i(t,\cdot)\right] & x \in D,\ t \in (0,T], \\
& i = 1,\cdots,S \\[4pt]
U_i(x,t) = 0 & (x,t) \in \partial D \times (0,T], \\
& i = 1,\cdots,R \\[4pt]
U_i(x,t) = v_i(x,t) & (x,t) \in \partial D \times (0,T], \\
& i = R+1,\cdots,S
\end{cases}
\tag{4.35}
$$

$$
\sum_{j=1}^{R} \mathbf{M}_{ji}(t)\dot{Y}_j(t,\omega) = \Pi_{\mathscr{Y}}^{\perp}\langle \mathscr{L}(u_S(\cdot,t,\omega),\omega), U_i(\cdot,t)\rangle
\qquad
\begin{aligned}
&(t,\omega) \in (0,T] \times \Omega, \\
&i = 1,\cdots,R
\end{aligned}
\tag{4.36}
$$

$$
\dot{Y}_i(t,\omega) = 0
\qquad
\begin{aligned}
&(t,\omega) \in (0,T] \times \Omega, \\
&i = R+1,\cdots,S
\end{aligned}
\tag{4.37}
$$

where $\mathscr{Y} = \mathrm{span}(Y_1,...,Y_S)$ and $\mathbf{M} \in \mathbb{R}^{R \times R}$ is the correlation matrix of the first R deterministic modes $\mathbf{M}_{i,j} = \langle U_i, U_j \rangle$ for $i,j = 1,...,R$. Observe that the system (4.36) consists of only $R = S - M$ equations since the last M random variables remain constant.

Remark 4.3.2. *Again, since in our context of partial differential equations with random parameters we are usually interested in computing the statistics of the solution it may be worth approximating separately the mean of the solution as in (4.18). Observe that we can distinguish two cases:*

- *(non homogeneous) deterministic boundary conditions. In this case isolating the mean only reduces to a special case of the Dual DO formulation under boundary constraints with $S+1$ modes and $M = 1$ constraint: $\bar{u}_{S+1}$ satisfies the constraints and all other deterministic modes are homogeneous on the boundary. The approximation manifold can be defined including the mean, by taking $U_{S+1} = \bar{u}_{S+1}$ and $Y_{S+1} = 1$. Observe that the non homogeneous boundary conditions guarantee that U_{S+1} and U_i are linearly independent for any $i = 1,...,S$. In practice, we work in a manifold of rank $S+1$ under one constraint, given by $Y_{S+1} = 1$ at each time.*

- *random boundary conditions. Consider a boundary datum $g = \bar{g} + \sum_{i=1}^{M} v_i Z_i$, with $\mathbb{E}[Z_i] = 0$ and $\bar{g} = \mathbb{E}[g]$, if $\bar{g}$ is linearly independent from $\{v_1,...,v_M\}$ we fall back to the first case, namely isolating the mean coincides with defining the constrained approximation manifold by including the mean: we have $S+1$ linearly independent modes and $M+1$ constraints. On the other hand in the general case, isolating the mean does not guarantee any kind of orthogonality for $\bar{u}_S$ with respect to the other deterministic modes, thus the constrained set which includes the mean is not necessarily a manifold. If $\bar{g}$ is not linearly independent from $v_1,...,v_M$, we can either isolate the mean and work in a manifold of dimension S embedded in $H^1(D) \otimes L_0^2(\Omega)$, or write g_M as in (4.25). Observe that in the first case the approximate solution is in $H^1(D) \oplus \mathcal{M}_S^{g_M}$ and has rank at least equal to S.*

An alternative strategy for dealing with random boundary conditions was proposed in [116],[117] and consists in projecting the boundary conditions $g(x, t, \omega)$ onto $\mathcal{Y}(t) = \text{span} < Y_1, ..., Y_S >$ at each t. Combining this approach to the Dual DO framework would imply enforcing:

$$U_i(x, t)_{|\partial D} = \mathbb{E}[g(x, t, \cdot) Y_i(t, \cdot)] \quad \forall x \in \partial D, \forall t \in (0, T], \forall i = 1, ..., S$$

which further implies:

$$u_S(x, t, \cdot)_{|\partial D} = \sum_{i=1}^{S} \mathbb{E}[g(x, t, \cdot) Y_i(t, \cdot)] Y_i(t, \omega) \quad \forall x \in \partial D, \forall t \in (0, T], \forall i = 1, ..., S.$$

However, note that the subspace spanned by $Y_1, ..., Y_S$ evolves in time and it is implicitly determined by the approximate solution itself. It is not clear then at time $t \gg 0$ which boundary conditions are actually satisfied by u_S and how the randomness arising from the boundary data is taken into consideration. The two different strategies are numerically compared in Section 4.5. The results for the problems under analysis show that strong imposition of boundary constrains leads to better performances in terms of accuracy versus number of modes, especially for long time intervals.

4.3.3   Best S rank approximation

We now look at the problem of finding the best S rank approximation in $\mathcal{M}_S^{g_M(t)}$ at any fixed time $t \in [0, T]$. This can be seen as an optimization problem under constraints.

Definition 4.3.3. *Fix $t \in [0, T]$ and let $u(t) \in H^1(D) \otimes L^2(\Omega)$ be a square integrable random field with rank greater or equal to S and such that $u(x, t, \omega) = g_M(x, t, \omega)$ for all $x \in \partial D$ a.s.. We define best rank S approximation a solution of the following problem:* find $u_S^{KL}(t) \in \mathcal{M}_S^{g_M(t)}$ *such that*

$$u_S^{KL}(t) = \underset{v_S \in \mathcal{M}_S^{g_M(t)}}{argmin} \|u(t) - v_S\|_{L^2(D) \otimes L^2(\Omega)} \tag{4.38}$$

Lemma 4.3.3. *The solution to problem (4.38) is given by:*

$$\begin{aligned} u_S^{KL}(x, t, \omega) &= \sum_{i=1}^{R} \sqrt{\lambda_i(t)} V_i^{KL}(x, t) Z_i^{KL}(t, \omega) + \sum_{i=1}^{M} V_i(x, t) Z_i(\omega, t) \\ &= u_R^{KL}(t) + \underbrace{\sum_{i=1}^{M} V_i(x, t) Z_i(t, \omega)}_{u_M^*(t)} \end{aligned} \tag{4.39}$$

where

- *$u_M^*(t) \in \mathcal{M}_M^{g_M(t)}$ is the Galerkin projection of $u(t)$ in $H^1(D) \otimes \mathcal{Z}(t)$. Specifically $V_i(x, t) = \mathbb{E}[u(x, t, \cdot) Z_i(t, \cdot)]$;*

- *$u_R^{KL}(t)$ is the best approximation with R terms of $\Pi_{\mathcal{Z}(t)}^{\perp}[u(t)] = u(t) - u_M^*(t): (\lambda_i(t), V_i^{KL}(t), Z_i^{KL}(t))_{i=1}^{R}$*

are the first R terms of the Karhunen-Loève expansion of $\Pi^{\perp}_{\mathcal{Z}(t)}[u(t)]$.

Proof. Observe that $\mathcal{M}^{g_M(t)}_S \subset H^1(D) \otimes L^2(\Omega)$ while the minimization in (4.38) is defined in $L^2(D) \otimes L^2(\Omega)$. In order to formally recover the constraint (4.26), i.e. $u^{KL}_S(x,t,\omega) = g_M(x,t,\omega)$, $\forall x \in \partial D$ a.s., we set problem (4.38) in the larger space $\mathcal{V} \subset L^2(D) \otimes L^2(\Omega)$ defined as:

$$\mathcal{V} = L^2(D) \otimes \mathcal{Z}(t) \oplus \mathcal{M}_R(L^2(D) \otimes \mathcal{Z}^{\perp}(t))$$

where $\mathcal{M}_R(L^2(D) \otimes \mathcal{Z}^{\perp}(t))$ is the manifold of R rank random fields embedded in $L^2(D) \otimes \mathcal{Z}^{\perp}(t)$. Now let us define the following problem: *Find* $\tilde{u}^{KL}_S(t) \in \mathcal{V}$ *such that*

$$\tilde{u}^{KL}_S(t) = \underset{v_S \in \mathcal{V}}{\mathrm{argmin}} \| u(t) - v_S \|_{L^2(D) \otimes L^2(\Omega)} \tag{4.40}$$

The problem (4.40) reduces to two well known problems: a Galerkin projection in $\mathcal{Z}(t)$ plus an optimization problem without constraints in $\mathcal{M}_R(L^2(D) \otimes \mathcal{Z}^{\perp}(t))$. This implies that problem (4.40) is well posed and admits a solution $\tilde{u}^{KL}_S(t)$ that can be written as in (4.39). Moreover observe that $\tilde{u}^{KL}_S(t) \in \mathcal{M}^{g_M(t)}_S \subset \mathcal{V}$, which implies that $\tilde{u}^{KL}_S(t) = u^{KL}_S(t)$ is a solution of problem (4.40). $\qquad\square$

In the following we call best S rank approximation $u^{KL}_S(t)$ the solution to problem (4.38).

Remark 4.3.3. *The error analysis derived in [96] for linear parabolic equations with random parameters applies as well to the Dual DO approximation under constraints. In this case, the DLR approximation error is bounded in term of the best approximation (4.39), i.e. the solution of the optimization problem under constraints (4.38). The proof follows very closely the one derived in [96].*

4.4   Application to Navier Stokes equations

In this Section we focus on fluid flow dynamics governed by the non-stationary Navier Stokes equations for incompressible, constant-density fluids. In this setting the uncertainty may arise from the parameters of the equations such as the fluid viscosity, or from the forcing term or initial or boundary conditions. The general problem, in a open, bounded Lipschitz domain $D \subset \mathbb{R}^d$, with $d = 2, 3$, reads a.s. in Ω as:

$$\begin{cases} \dot{\mathbf{u}}(\mathbf{x},t,\omega) - \nu(\mathbf{x},t,\omega)\Delta\mathbf{u}(\mathbf{x},t,\omega) + \mathbf{u}(\mathbf{x},t,\omega) \cdot \nabla\mathbf{u}(\mathbf{x},t,\omega) + \nabla p(\mathbf{x},t,\omega) = \mathbf{f}(\mathbf{x},t,\omega) & (\mathbf{x},t) \in D \times (0,T] \\ \nabla \cdot \mathbf{u}(\mathbf{x},t,\omega) = 0 & \\ \mathbf{u}(\mathbf{x},0,\omega) = \mathbf{u}_0(\mathbf{x},\omega) & \mathbf{x} \in D \\ \mathbf{u}(\mathbf{x},t,\omega) = \mathbf{g}(\mathbf{x},t,\omega) & \mathbf{x} \in \Gamma_D, t \in (0,T] \\ \nu\partial_{\mathbf{n}}\mathbf{u}(\mathbf{x},t,\omega) - p(\mathbf{x},t,\omega) \cdot \mathbf{n} = \mathbf{h}(\mathbf{x},t,\omega) & \mathbf{x} \in \Gamma_N, t \in (0,T] \end{cases}$$

$$\tag{4.41}$$

where $\mathbf{u}$ is the velocity (column) vector field, p is the scalar pressure and ν is the kinematic viscosity that may eventually be modeled as a random variable or random field. Γ_D and Γ_N are disjointed parts of the boundary ∂D, such that $\overline{\Gamma}_D \cup \overline{\Gamma}_N = \partial D$, on which we impose Dirichlet and Neumann boundary conditions respectively.

Our goal is to find a low rank approximation of the velocity field. We apply the Dual DO method described in Section 4.3 and we derive evolution equations for all the factors $(\mathbf{U}, \mathbf{Y})$ of the approximate velocity vector field. We start by recalling the definition of Karhunen-Loève expansion for a square integrable random *vector* field.

Definition 4.4.1. *Let $\mathbf{u} \in L^2\big(\Omega, [L^2(D)]^d\big)$ be a square integrable random field with covariance function $\mathrm{Cov}_{\mathbf{u}} : D \times D \to \mathbb{R}^{d \times d}$, defined as:*

$$\mathbf{Cov_u}(\mathbf{x}, \mathbf{y}) = \mathbb{E}[\mathbf{u}^*(\mathbf{x}, \cdot)\mathbf{u}^{*\mathrm{T}}(\mathbf{y}, \cdot)]$$

with $\mathbf{u}^* = \mathbf{u} - \mathbb{E}[\mathbf{u}]$.
Then $\mathbf{u}$ can be written as:

$$\mathbf{u}(\mathbf{x}, \omega) = \bar{\mathbf{u}}(\mathbf{x}, \omega) + \underbrace{\sum_{i=1}^{\infty} \sqrt{\lambda_i} \mathbf{V}_i^{KL}(\mathbf{x}) Z_i^{KL}(\omega)}_{\mathbf{u}^*} \tag{4.42}$$

where:

- *$\{\lambda_i, \mathbf{V}_i^{KL}\}$ are respectively the (non-zero) eigenvalues and eigenfunctions (column vectors of scalar functions) of the covariance operator $T_{\mathbf{u}} : [L^2(D)]^d \to [L^2(D)]^d$ defined as*

$$\begin{aligned} T_{\mathbf{u}}\mathbf{V}(\mathbf{x}) &= \int_D \mathbf{Cov_u}(\mathbf{y}, \mathbf{x})\mathbf{V}(\mathbf{y})\mathrm{d}\mathbf{y}, \quad \mathbf{V} \in [L^2(D)]^d \\ T_{\mathbf{u}}\mathbf{V}_i^{KL} &= \lambda_i \mathbf{V}_i^{KL} \end{aligned}$$

- *Z_i^{KL} are mutually uncorrelated scalar random variables given by:*

$$Z_i^{KL}(\omega) := \frac{1}{\sqrt{\lambda_i}} \int_D (\mathbf{u}^*(x, \omega))^T \mathbf{V}_i^{KL}(x)\, dx \qquad \forall i \in \mathbb{N}^+,$$

  with zero mean and unit variance.

Observe that the deterministic modes are vector valued functions while the stochastic modes are scalar functions. We denote by $H_{div}^1(D)$ and $H_{\Gamma_D}^1(D)$ the following spaces:

$$H_{div}^1(D) := \big\{\mathbf{v} \in [H^1(D)]^d : \nabla \cdot \mathbf{v} = 0\big\},$$

$$H_{\mathbf{g}}^1(D) := \big\{\mathbf{v} \in [H^1(D)]^d : \mathbf{v}_{|\Gamma_D} = \mathbf{g}\big\}, \quad H_{\Gamma_D}^1(D) := \big\{\mathbf{v} \in [H^1(D)]^d : \mathbf{v}_{|\Gamma_D} = \mathbf{0}\big\}.$$

Remark 4.4.1. *Let $\mathbf{u}$ be in $H_{div}^1(D) \otimes L^2(\Omega)$, then the mean and all the deterministic eigen-modes in (4.42) are divergence free. Indeed*

$$\nabla \cdot \bar{\mathbf{u}} = \nabla \cdot \mathbb{E}[\mathbf{u}] = \mathbb{E}[\nabla \cdot \mathbf{u}] = 0, \qquad \lambda_i \nabla \cdot \mathbf{V}_i^{KL} = \nabla \cdot \mathbb{E}[\mathbf{u} Z_i^{KL}] = \mathbb{E}[(\nabla \cdot \mathbf{u}) Z_i^{KL}] = 0$$

In light of Remark 4.4.1, we look for a Dynamical Low Rank approximation written as a linear combination of divergence free modes.

We consider the general case of problem (4.41) with random Dirichlet boundary conditions, and detail the Dual DO formulation introduced in Section 4.3.1 in which we also isolate the mean. Following the discussion in Section 4.3.2, we assume that the datum on the Dirichlet boundary can be properly approximated by a M rank random field, with $M \leq S$. In particular, for consistency with the approximate solution, the boundary constraint is decomposed according to Definition (4.18), by isolating the mean:

$$\mathbf{u}(\mathbf{x}, t, \omega) = \mathbf{g}(\mathbf{x}, t, \omega) \approx \mathbf{g}_M(\mathbf{x}, t, \omega) = \bar{\mathbf{g}}(\mathbf{x}, t) + \mathbf{g}_M^*(\mathbf{x}, t, \omega), \quad \mathbf{x} \in \Gamma_D, \ t \in [0, T], \text{ a.s.} \qquad (4.43)$$

Hence $\bar{\mathbf{g}}(\mathbf{x}, t)$ is the deterministic Dirichlet boundary condition for the mean, while $\mathbf{g}_M^*(\mathbf{x}, t, \omega)$, written as a linear combination of $M \leq S$ zero mean random variables, is the constraint of the approximation manifold. To be precise:

$$\mathbf{g}_M(\mathbf{x}, t, \omega) = \bar{\mathbf{g}}(\mathbf{x}, t) + \sum_{i=1}^{M} \boldsymbol{v}_i(\mathbf{x}, t) Z_i(t, \omega), \quad \forall \mathbf{x} \in \Gamma_D, \ t \in [0, T], \text{ a.s.} \qquad (4.44)$$

with:

- $Z_1, ..., Z_M$ zero mean $L^2(\Omega)$-orthonormal random variables: $\mathbb{E}[Z_i(\cdot, t)] = 0$, $\mathbb{E}[Z_i(\cdot, t) Z_j(\cdot, t)] = \delta_{ij}$ for all $i, j = 1, ..., M$

- $\boldsymbol{v}_1, ..., \boldsymbol{v}_M$ linearly independent vector valued deterministic functions.

and the approximation manifold of *zero mean S rank random fields* constrained to $\mathbf{g}_M^*(t)$ is parametrized as follows:

$$\mathcal{M}_{S,div}^{\mathbf{g}_M^*(t)} = \left\{ \mathbf{u}_S^* = \sum_{i=1}^{S} \mathbf{U}_i Y_i \text{ s.t.} \quad \mathbf{u}_{S|\Gamma_D} = \mathbf{g}_M^*(t), \text{ and } \mathbf{U}_i \in H_{div}^1(D), \\ \mathbb{E}[Y_i] = 0, \mathbb{E}[Y_i Y_j] = \delta_{ij}, \text{ rank}(\mathbf{M}) = R \right\} \qquad (4.45)$$

where $R = S - M$ and $\mathbf{M} = \ll \mathbf{U}, \mathbf{U} \gg \in \mathbb{R}^{R \times R}$ is again the full rank correlation matrix of the first R deterministic modes: $\mathbf{M}_{ij} = \langle \mathbf{U}_i, \mathbf{U}_j \rangle = \sum_{k=1}^{d} \langle U_{i,k}, U_{j,k} \rangle$. Thus, the DLR approximate solution is written at each time as:

$$\mathbf{u}_S(t) = \bar{\mathbf{u}}_S(t) + \mathbf{u}_S^*(t)$$

with:

- $\bar{\mathbf{u}}_S(t) \in H_{div}^1(D) \cap H_{\bar{\mathbf{g}}(t)}^1(D)$,

- $\mathbf{u}_S^*(t) \in \mathcal{M}_{S,div}^{\mathbf{g}_M^*(t)}$.

Finally, the DLR variational principle (4.12) applied to the Navier Stokes problem in (4.41) becomes:

DLR Variational Principle. *At each time t, find $(\bar{\mathbf{u}}_S(t), \mathbf{u}_S^*(t)) \in (H_{div}^1(D) \cap H_{\bar{\mathbf{g}}(t)}^1(D)) \times \mathcal{M}_{S,div}^{\mathbf{g}_M^*(t)}$ such that $\mathbf{u}_S(t) = \bar{\mathbf{u}}_S(t) + \mathbf{u}_S^*(t) = \bar{\mathbf{u}}_S(t) + \sum_{i=1}^S \mathbf{U}_i(t) Y_i(t)$ satisfies:*

$$\mathbb{E}\big[<\dot{\mathbf{u}}_S + \mathbf{u}_S \cdot \nabla \mathbf{u}_S - \mathbf{f}, \mathbf{v}> + \quad <\nu \nabla \mathbf{u}_S, \nabla \mathbf{v}> - <\mathbf{h}, \mathbf{v}>_{\Gamma_N} \big] = 0$$
$$\forall \mathbf{v} \in (H_{div}^1(D) \cap H_{\Gamma_D}^1(D)) \times \mathcal{T}_{\mathbf{u}_S(t)} \mathcal{M}_{S,div}^{\mathbf{g}_M^*(t)}. \quad (4.46)$$

with initial condition given by $\bar{\mathbf{u}}(0) + \mathbf{u}_S^*(0)$, where $\bar{\mathbf{u}}(0) = \mathbb{E}[\mathbf{u}_0]$ and $\mathbf{u}_S^*(0)$ is the best S rank approximation of $\mathbf{u}_0 - \bar{\mathbf{u}}(0)$ in $\mathcal{M}_{S,div}^{\mathbf{g}_M^*(0)}$, provided $\mathbf{u}_0 \in H_{div}^1 \otimes L^2(\Omega)$.

Observe that the term $<\mathbf{h}, \mathbf{v}>_{\Gamma_N}$ derives from the integration by part of $-\nu <\Delta \mathbf{u}_S, \mathbf{v}> + <\nabla p, \mathbf{v}>$ combined with the Neumann boundary conditions in Γ_N. Again, by imposing condition (4.14), we can equip $\mathcal{M}_{S,div}^{\mathbf{g}_M^*(t)}$ with a differential manifold structure and derive the Dual DO reduced order system for Navier Stokes equations with random parameters (including boundary conditions).

Proposition 4.4.1. *Let $(\bar{\mathbf{u}}_S, \mathbf{U}, \mathbf{Y})$ be a smooth solution of*

$$\begin{cases} <\dot{\bar{\mathbf{u}}}_S + \mathbb{E}[\mathbf{u}_S \cdot \nabla \mathbf{u}_S - \mathbf{f}], \delta \mathbf{U}> + <\mathbb{E}[\nu \nabla \mathbf{u}_S], \nabla \delta \mathbf{U}> - <\mathbb{E}[\mathbf{h}], \delta \mathbf{U}>_{\Gamma_N} = 0 \\ \bar{\mathbf{u}}_{S_{|\Gamma_D}} = \bar{\mathbf{g}} \\ <\dot{\mathbf{U}}_i + \mathbb{E}[(\mathbf{u}_S \cdot \nabla \mathbf{u}_S - \mathbf{f}) Y_i], \delta \mathbf{U}> + <\mathbb{E}[\nu \nabla \mathbf{u}_S Y_i], \nabla \delta \mathbf{U}> - <\mathbb{E}[\mathbf{h} Y_i], \delta \mathbf{U}>_{\Gamma_N} = 0 \\ \mathbf{U}_{i|\Gamma_D} = 0 & \forall i = 1, ..., R \\ \mathbf{U}_{i|\Gamma_D} = \boldsymbol{\nu}_i & \forall i = R+1, ..., S \\ \\ \sum_{k=1}^R \mathbf{M}_{jk} \mathbb{E}[\dot{Y}_k \delta Y] = <\mathbf{U}_j, \mathbb{E}[(\mathbf{f} - \mathbf{u}_S \cdot \nabla \mathbf{u}_S) \delta Y]> - <\nabla \mathbf{U}_j, \mathbb{E}[\nu \nabla \mathbf{u}_S \delta Y]> + <\mathbf{U}_j, \mathbb{E}[\mathbf{h} \delta Y]>_{\Gamma_N} \\ \hfill \forall j = 1, ..., R \end{cases}$$
$$(4.47)$$

$\forall \delta \mathbf{U} \in H_{div}^1(D) \cap H_{\Gamma_D}^1(D)$ and $\forall \delta Y \in \mathcal{Y}_0^\perp$ (the orthogonal complement of $\mathcal{Y}$ in $L_0^2(\Omega)$), with initial conditions given by $\bar{\mathbf{u}}(0)$ and the best S rank approximation of $\mathbf{u}_0 - \bar{\mathbf{u}}(0)$ in $\mathcal{M}_{S,div}^{\mathbf{g}_M^*(0)}$. Then $\mathbf{u}_S = \bar{\mathbf{u}}_S + \sum_{i=1}^S \mathbf{U}_i Y_i$ is solution of (4.46), at each time.

We treat the divergence free constraint, that is imposed on each deterministic mode, by introducing $S + 1$ Lagrange multipliers $\bar{p}, p_1, ..., p_S$. Then, by reintegrating by part, we finally get:

Proposition 4.4.2. *Let* $(\bar{\mathbf{u}}_S, \mathbf{U}, \mathbf{Y})$ *be a smooth solution of*

$$\dot{\bar{\mathbf{u}}}_S + \nabla \bar{p} = \mathbb{E}\left[\nu \Delta \mathbf{u}_S - \mathbf{u}_S \cdot \nabla \mathbf{u}_S + \mathbf{f}\right]$$
$$\nabla \cdot \bar{\mathbf{u}}_S = 0$$
$$\dot{\mathbf{U}}_i + \nabla p_i = \mathbb{E}\left[\left(\nu \Delta \mathbf{u}_S - \mathbf{u}_S \cdot \nabla \mathbf{u}_S + \mathbf{f}\right) Y_i\right] \qquad (4.48)$$
$$\nabla \cdot \mathbf{U}_i = 0 \qquad\qquad\qquad \forall i = 1, ..., S$$
$$\sum_{k=1}^{R} \mathbf{M}_{ik} \dot{Y}_k = <\mathbf{U}_i, \Pi_{1 \cup \mathcal{Y}}^{\perp}\left[\nu \Delta \mathbf{u}_S - \mathbf{u}_S \cdot \nabla \mathbf{u}_S + \mathbf{f}\right]> \qquad \forall i = 1, ..., R$$

then $\mathbf{u}_S = \bar{\mathbf{u}}_S + \mathbf{U}\mathbf{Y}$ *is solution of (4.46), at each time.*

The initial conditions are given by $\bar{\mathbf{u}}(0)$ *and the best S rank approximation of* $\mathbf{u}_0 - \bar{\mathbf{u}}(0)$ *in* $\mathcal{M}_{S,div}^{\mathbf{g}_M^*(0)}$, *while the boundary conditions are the following:*

$$\bar{\mathbf{u}}_S(x, t) = \bar{\mathbf{g}}(x, t) \qquad\qquad (\mathbf{x}, t) \in \Gamma_D \times [0, T],$$
$$\mathbf{U}_i(x, t) = \boldsymbol{v}_i(x, t) \qquad\qquad (\mathbf{x}, t) \in \Gamma_D \times [0, T], \forall i = 1, ..., R$$
$$\mathbf{U}_i(x, t) = \mathbf{0} \qquad\qquad (\mathbf{x}, t) \in \Gamma_D \times [0, T], \forall i = R+1, ..., S$$
$$\nu \partial_\mathbf{n} \mathbf{U}_i(\mathbf{x}, t) - p_i(\mathbf{x}, t) \cdot \mathbf{n} = \mathbb{E}[\mathbf{h}(\mathbf{x}, t, \cdot) Y_i] \quad (\mathbf{x}, t) \in \Gamma_N \times [0, T], \forall i = 1, ..., S.$$

In conclusion, the Navier Stokes equations with random parameters in (4.41) is reduced to S deterministic problems of Navier Stokes type, coupled to a system of R stochastic ODEs.

4.5 Numerical Test

4.5.1 Flow around a cylinder: stochastic boundary condition

In this section we consider a two-dimensional incompressible flow over a circular cylinder at moderate Reynold's Numbers for which a periodic vortex shedding phenomenon is observed around the obstacle. The geometry and the mesh used for the simulations are shown in Figure 4.1. The height and length of the channel are respectively $H = 0.41$ and $l = 2.2$. The cylinder hole has radius $r = 0.05$ and is slightly uncentered, the coordinate of the center being (0.2, 0.2) with respect to the origin located on the lower-left corner of the channel. We consider homogeneous initial conditions and we assume a parabolic inflow profile with random peak
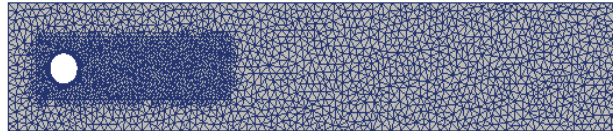
Figure 4.1 – Left: mesh used for the simulation, 2592 number of vertices, h_{max}= 0.055, h_{min}=0.006.

velocity U_{max} that varies in the range [1.2, 1.8]. More precisely we have:

$$
\begin{aligned}
\mathbf{u}(\mathbf{x}, \omega, t) \quad &= (4\underbrace{(\bar{U} + \sigma Z(\omega))}_{U_{max}} x_2(H - x_2)/H^2, 0) \qquad\qquad \mathbf{x} = (x_1, x_2) \in \Gamma_{in} \\
&= (4\bar{U}x_2(H - x_2)/H^2, 0) + (4\sigma Z(\omega)x_2(H - x_2)/H^2, 0) \qquad\qquad (4.49) \\
&= \bar{\mathbf{g}}(\mathbf{x}) + \underbrace{Z(\omega)\boldsymbol{v}_1(\mathbf{x})}_{\mathbf{g_1}}
\end{aligned}
$$

where $\bar{U} = 1.5$, $\sigma = 0.1$ and Z is a uniform random variable with zero mean and unit variance. An initial ramp is applied on the boundary data to guarantee consistency with the homogeneous initial conditions. We use a cubic polynomial smoothing function that reaches 1 at time t=1. No slip conditions are applied on the top, bottom and cylinder side-walls, Neumann homogeneous conditions at the outlet. We recall that for this problem the Reynold's number (that can be computed as $Re = \frac{U_m D}{\nu}$ where here $\nu = 10^{-3}$ is the kinematic viscosity and $U_m = \frac{2}{3}U_{max}$ is the mean inflow velocity) determines the frequency of the vortex shedding and the length of the recirculation region. Observe that the random boundary condition at the inflow directly influences the Reynold's Numbers, that here varies in the range $Re \in [80, 120]$. It follows that the pattern of the solutions corresponds to flows with random vortex shedding frequency. Before presenting the numerical results we remark that even if we have only one random variable as input, the problem is not straightforward since the solution depends non-linearly on it. Indeed the test case under consideration is challenging for model order reduction techniques, which are unable to approximate the solution manifold with a relatively small number of modes. This is due to the fact that each value of the input parameter leads to a vortex shedding with different frequency and characteristic length. Somehow we can imagine the manifold of solutions to be constituted of infinitely many flow patterns which become more and more out of phase one with respect to the others as time evolves. Consequently, even if we start with a low-rank initial condition (or even deterministic) the rank of the solution will significantly increase in time. This can be verified by looking at the evolution of the eigenvalues of the covariance operator in Figure 4.8 (left). Consider for instance the POD method [13][23], in which the approximate solution is sought as a linear combination of (deterministic and fixed in time) precomputed modes. For a fixed value of Reynold's number the solution is periodic and the vortex shedding can be well reproduced as a linear combination of few pairs of modes with alternating symmetry properties: the dynamics of the solution is approximately low rank (or well approximated in a low dimensional manifold), at least once the vortex shedding is fully developed. On the other hand, it has been shown (see for instance [99][85] for details), that the span of the eigen-modes changes significantly with the Reynold's number, making the POD approach very sensitive to the choice of the parameters used to compute the snapshots. Indeed POD techniques may fail to capture the dynamics for values of the Reynold's number different from those used to pre-compute the modes. We are interested in understanding if the dynamical approach of the DLR method can, at least partially, overcome this problem. In particular, we analyze the performance of the Dual DO method in describing both the transient period and the long-term periodic dynamics.

Dual DO system and numerical discretization

We apply the Dual DO formulation to the Navier-Stokes (NS) equations derived in Section 4.4. By isolating the mean, the Dual DO approximate solution is written as:

$$\mathbf{u}_S(\mathbf{x}, t, \omega) = \bar{\mathbf{u}}_S(\mathbf{x}, t) + \sum_{i=1}^{S-1} \mathbf{U}_i(\mathbf{x}, t) Y_i(t, \omega) + \mathbf{U}_S(\mathbf{x}, t) Z(\omega)$$

where Z is the random variable in (4.49), and $\mathbf{u}_S^* = \mathbf{u}_S - \bar{\mathbf{u}}_S$ belongs to the low dimensional manifold:

$$\mathcal{M}_{S,div}^{\mathbf{g}_1(t)} = \left\{ \mathbf{u}_S = \sum_{i=1}^{S} \mathbf{U}_i Y_i \text{ s.t. } \mathbf{u}_{S|\Gamma_{in}} = \mathbf{g}_1(t), \mathbf{U}_i \in H_{div}^1, Y_i \in L_0^2(\Omega), \mathbb{E}[Y_i Y_j] = \delta_{ij}, \text{ rank}(\mathbf{M}) = S-1 \right\}$$

Let $\mathbb{B}$ denote the third order tensor defined as $\mathbb{B}_{ijk} := \mathbb{E}[Y_i Y_j Y_k]$, the Dual DO system for this problem reads:

$$
\begin{cases}
\dot{\bar{\mathbf{u}}}_S - \nu \Delta \bar{\mathbf{u}}_S + \bar{\mathbf{u}}_S \cdot \nabla \bar{\mathbf{u}}_S + \sum\limits_{i=1}^{S} \mathbf{U}_i \cdot \nabla \mathbf{U}_i + \nabla \bar{p}_S = 0 \\
\nabla \cdot \bar{\mathbf{u}}_S = 0 \\
\bar{\mathbf{u}}_{S|\Gamma_{in}} = \bar{\mathbf{g}} \\
\dot{\mathbf{U}}_k - \nu \Delta \mathbf{U}_k + \sum\limits_{i=1}^{S} \sum\limits_{j=1}^{S} \mathbb{B}_{ijk} \mathbf{U}_i \cdot \nabla \mathbf{U}_j + \mathbf{U}_k \cdot \nabla \bar{\mathbf{u}}_S + \bar{\mathbf{u}}_S \cdot \nabla \mathbf{U}_k + \nabla p_k = 0 \qquad k = 1, ..., S \\
\nabla \cdot \mathbf{U}_k = 0 \\
\mathbf{U}_{1|\Gamma_{in}} = \mathbf{v}_1 \qquad \mathbf{U}_{k|\Gamma_{in}} = \mathbf{0} \qquad \text{for } k \neq 1
\end{cases}
$$
$$\sum_{i=1}^{S-1} \mathbf{M}_{ki} \dot{Y}_i + \Pi_{1 \cup \mathcal{Y}}^{\perp} < \mathbf{U}_k, \sum_{i=1}^{S} \sum_{j=1}^{S} \mathbf{U}_i \cdot \nabla \mathbf{U}_j > Y_i Y_j = 0 \qquad k = 1, ..., S-1$$
$$Y_S = Z$$

No slip conditions on the top, bottom and cylinder side-walls and homogeneous Neumann boundary conditions on the outflow are applied for $\bar{\mathbf{u}}_S, \mathbf{U}_1, ..., \mathbf{U}_S$. Observe that the Dual DO system (4.5.1)-(4.5.1) reduces to $S+1$ (coupled) Navier-Stokes equations, plus a system of $S-1$ ODEs. However playing with the time discretization it is possible to decouple the system to save computational time and effectively compute the approximate solution without losing the stability. In particular, we used a splitting scheme of "Gauss-Seidel" type to linearize and completely decouple the system of ODEs from the system of PDEs. Specifically both the third order tensor $\mathbb{B}$ and the projection operator $\Pi_{1 \cup \mathcal{Y}}^{\perp}(\cdot)$ are treated explicitly whereas the update of the random variables $\{Y_i\}$ is done on the newly computed basis $\{U_i\}$. In particular, denoting by $\mathbf{u}_S^n = \bar{\mathbf{u}}_S^n + \mathbf{U}^n(\mathbf{Y}^n)$ the approximate solution at time $t^n = n\Delta t$, equations (4.5.1) and (4.5.1) are discretized in time as follows:

$$\frac{1}{\Delta t} \bar{\mathbf{u}}_S^{n+1} - \nu \Delta \bar{\mathbf{u}}_S^{n+1} + \bar{\mathbf{u}}_S^n \cdot \nabla \bar{\mathbf{u}}_S^{n+1} + \nabla \bar{p}_S^{n+1} = \frac{1}{\Delta t} \bar{\mathbf{u}}_S^n - \sum_{i=1}^{S} \mathbf{U}_i^n \cdot \nabla \mathbf{U}_i^n$$

$$\frac{1}{\Delta t} \mathbf{U}_k^{n+1} - \nu \Delta \mathbf{U}_k^{n+1} + \left(\bar{\mathbf{u}}_S^n + \sum_{i=1}^{S} \mathbb{B}_{ikk}^n \mathbf{U}_i^n \right) \cdot \nabla \mathbf{U}_k^{n+1} + \nabla p_k^{n+1} = \frac{1}{\Delta t} \mathbf{U}_k^n - \sum_{\substack{j=1 \\ j \neq k}}^{S} \sum_{i=1}^{S} \mathbb{B}_{ijk}^n \mathbf{U}_i^n \cdot \nabla \mathbf{U}_j^n - \mathbf{U}_k^n \cdot \nabla \bar{\mathbf{u}}_S^n$$

In conclusion at each time step we first solve in parallel $S+1$ decoupled deterministic NS equations and then a system of $S-1$ ODEs. All implementations has been developed within the open source Finite Element library FEniCs. We chose to discretize functions defined in the physical space by a Finite Element method, with P_2 elements for the velocity and P_1 for the pressure, to ensure that the inf-sup condition is satisfied at the discrete level. Then each of the $S+1$ deterministic NS equations system is solved by using the Chorin-Teman projection scheme with rotational incremental pressure correction [125]. The system of ODEs is instead discretized by using the Stochastic Collocation method with Gauss-Legendre points (the stochastic space has been parametrized by a uniform random variable, according with the input data).

The first difficulty in applying the Dual DO method concerns the initialization of the modes. Indeed, the initial condition is deterministic, i.e. a zero rank random field, but the rank is expected to increase in time, due to the randomness in the boundary data and the non linearity of the problem. It follows that we may expect $S>1$ modes to be needed to effectively describe the dynamics of the solutions. In practice we look for the Dual DO approximate solution $\mathbf{u}_S$ even if the initial condition has clearly defective rank. We initialize the last random mode to Z in (4.49) and the first $S-1$ to an orthonormal polynomial basis in $Z^\perp$, whereas the deterministic modes are set to zero. Observe that $\mathbf{u}_S$ does not belong to the approximation manifold $\mathscr{M}_{S,div}^{\mathbf{g}_1(t)}$ at least at the initial time steps. This is a common problem of DLR methods and may arise even if the initial solution is full rank because nothing prevent the DLR solution to become rank deficient at some point in time. To treat the case of rank deficiency here we used the same strategy proposed in [96] that consists in diagonalizing the correlation matrix $\mathbf{M}$ at each time step and solving the equations only for the eigen-modes with eigenvalues larger than a prescribed (small) tolerance, whereas the other modes are kept constant in time. By this way, the stochastic coefficients associated to deterministic modes with L^2 norm below the threshold, have a negligible influence on the approximation of the solution. However, they are kept in the approximation and may become active again at a later time when the rank of the solution increases. See [96] for more details. We mention that an alternative strategy to treat rank deficiency is proposed in [82], in the context of time-dependent matrices, which makes use of a projector splitting integrator.

Numerical results

First of all, we assess the accuracy of the Dual DO formulation with constraints according to which the random boundary conditions are imposed strongly. This technique is compared to the one proposed in [116] that consists in projecting the boundary data in the subspace spanned by the random modes at each time. The approximation error is calculated with respect to the reference solution computed by using the Stochastic Collocation method with Gauss-Legendre points and with the same discretization parameters in time and space. In Figure 4.2 (left) we compare the approximation error in norm $H^1(D) \otimes L^2(\Omega)$ for the two strategies with S (number of modes) equal to 7 and 11. We observe that for the problem under consideration the strategy proposed here exhibits a smaller error for long time integration. We
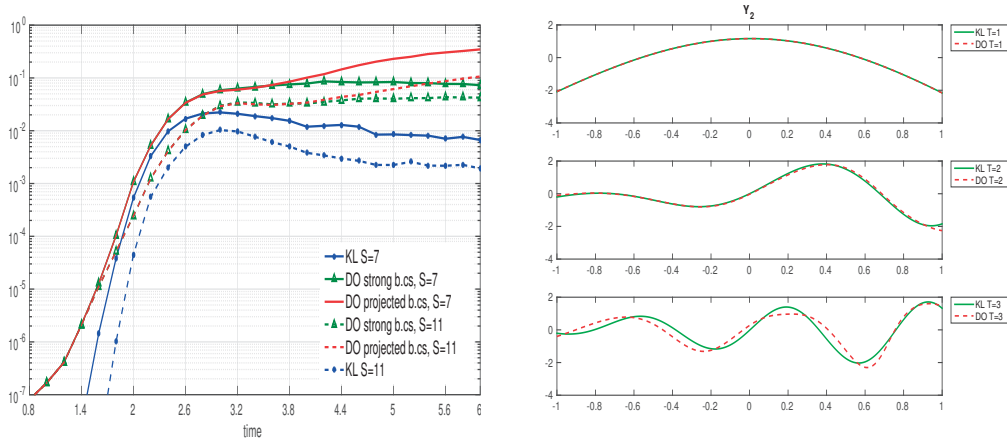
Figure 4.2 – Left: time evolution of the approximation error in norm $H^1(D) \otimes L^2(\Omega)$ with $S = 7$ modes (and $S = 11$, dashed lines). In blue the best approximation error, in red the approximation error of DO method with projected boundary conditions ([116][117]), in green approximation error of the Dual DO with strong imposition of boundary constraints. Right: The second stochastic mode of the KL decomposition of the reference solution (green) and the DO approximate solution (red dashed line) with $S = 5$ at time $T = 1$, $T = 2$, $T = 3$
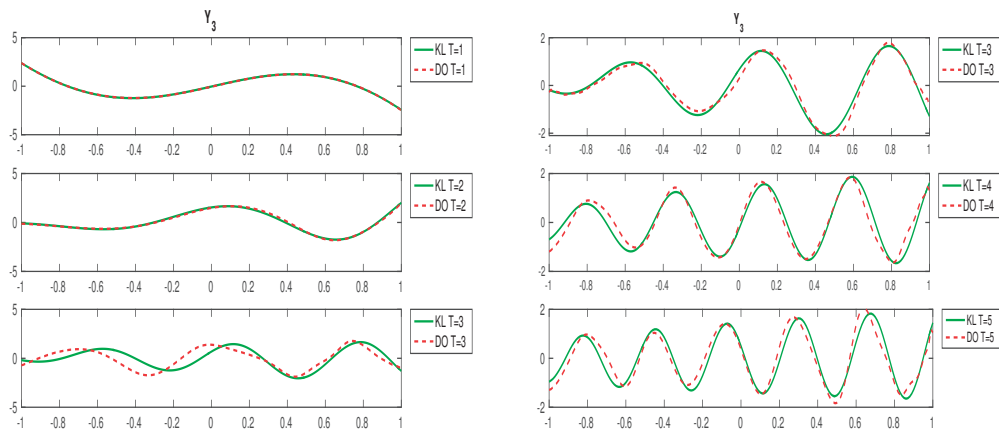
Figure 4.3 – Left: The third stochastic mode of the (constrained) KL decomposition (green) of the reference solution and of the Dual DO approximate solution (red dashed line) with $S = 5$ at time $T = 1$, $T = 2$, $T = 3$ seconds (left) and $S = 11$ at time $T = 3$, $T = 4$, $T = 5$ (right). seconds
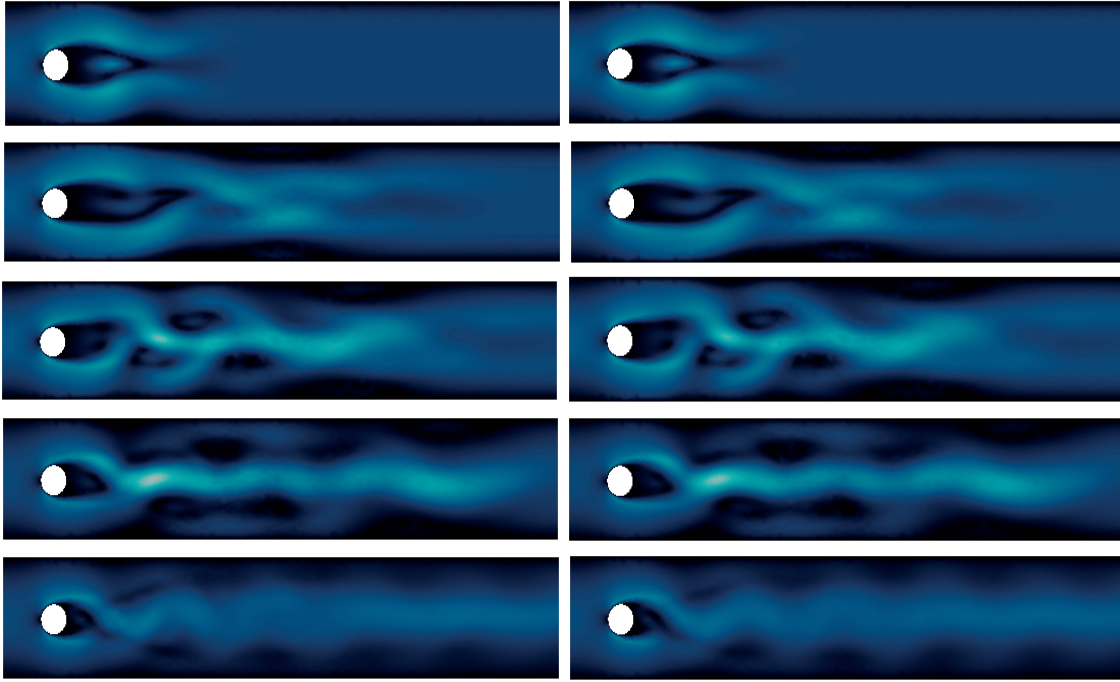
Figure 4.4 – The first deterministic mode of the Dual DO approximate solution with $S = 11$ (on the left) and the first KL eigen-mode of the reference solution (on the right) at $t = 0.6$, $t = 1.6$, $t = 2$, $t = 2.4$ and $t = 5.2$.

stress that for this problem the first random mode is fixed and distributed as Z in (4.49) while the other $S - 1$ random modes, initialized to an arbitrary orthonormal basis in the orthogonal complement to Z, "automatically" adapt to the structure of the solution. In Figure 4.2 (right) and Figure 4.3 we compare the random modes of the Dual DO approximate solution to the random modes of the best rank S approximation at different times. We recall that we denote by $\mathbf{u}_S^{KL}$ the solution to problem (4.38), namely the best S rank approximation in $\mathscr{M}_{S,div}^{\mathbf{g}_1(t)}$, the approximation manifold with constraints. As expected the accuracy in the evolution of the random basis depends on the number of modes used to compute the Dual DO approximate solution: the modes stay closer and closer to the optimal ones (and for a longer time interval) as S increases. In the first part of the transition phase, the stochastic modes properly adapt in time also when very few modes are used, whereas the effectiveness of the method tends to decrease for long time intervals, see Figure 4.2 (right) and Figure 4.3 (left). Better agreement for longer times is achieved by increasing the number of modes, Figure 4.3 (right). Similar conclusions can be drawn by analyzing the deterministic modes, see Figure 4.4 and Figure 4.5. We conclude this section by analyzing the rate of convergence of the Dual DO method with respect to the number of modes. The Dual DO approximation error is again computed in norm $H^1(D) \otimes L^2(\Omega)$ with respect to the reference solution computed by using the Stochastic Collocation method and with the same discretization parameters in time and space. In Figure 4.6 and Figure 4.7 we compare the Dual DO approximation error to the best approximation error as S increases and at different times. First of all, we observe that the approximation
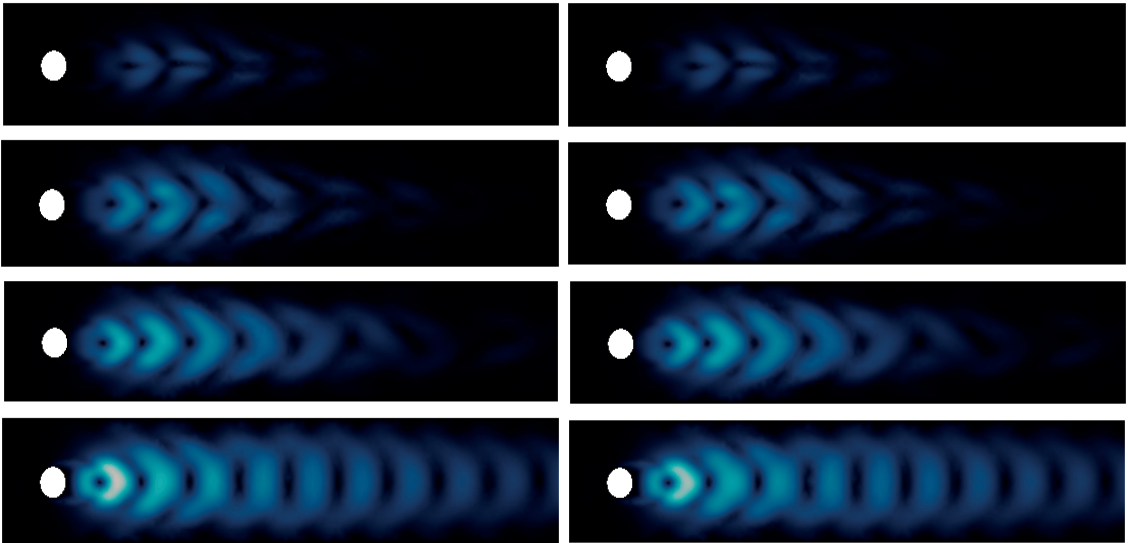
Figure 4.5 – The second deterministic mode of the Dual DO approximate solution with $S = 11$ (on the left) and the second KL eigen-mode of the reference solution (on the right) at $t = 0.6$, $t = 1.6$, $t = 2$, $t = 2.4$ and $t = 5.2$.

error increases in time, due to the intrinsic nature of the exact solution whose rank quickly increases until reaching a stable level when the vortex shedding is fully developed. In particular during the initial phase, for a fixed number of modes, both the DO and the Karhunen-Loève approximation error increase in time, which means that an increasing number of modes are needed to achieve a certain level of accuracy. We observe, however, in Figure 4.6 that the Dual DO approximation error stays very close the best approximation error and exhibits the same rate of convergence with respect to the number S of modes, until $T \approx 2.4$. On the other hand, Figure 4.7 shows that the difference between the Dual DO and the best approximation error tends to increase in time when the solution finally reaches the periodic phase. At this stage, we observe that the best approximation error stabilizes in time (or slightly decreases). On the other hand, the error of the Dual DO approximation is considerably larger and the convergence rate with respect to S seems to be worse than the one of the best approximation. This result is consistent with the quasi-optimal error estimate derived in [96] for linear parabolic problems in which the proportionality constant increases in time.

Time rescaling

The poor performance of the Dual DO method, as any other reduced order method, in efficiently approximate the problem in Section 4.5.1 for long times, is justified by the intrinsic nature of the solution whose structure is not apt to be well approximated in low rank format. This can be verified by looking at the evolution of the eigenvalues of the covariance operator in Figure 4.8 (Left): the eigenvalues increase fast and many of them reach not negligible values. Figure 4.7 shows the rate of convergence for the Karhunen-Loève approximation, i.e. the best
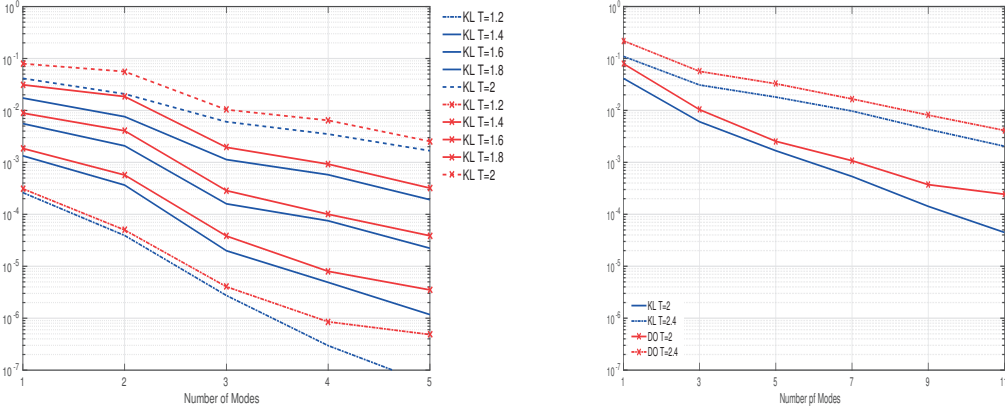
Figure 4.6 – The Dual DO approximation error (red) and the KL truncation error (blue) in norm $H^1(D) \otimes L^2(\Omega)$ with respect to the number of modes, at different time steps (transition phase).
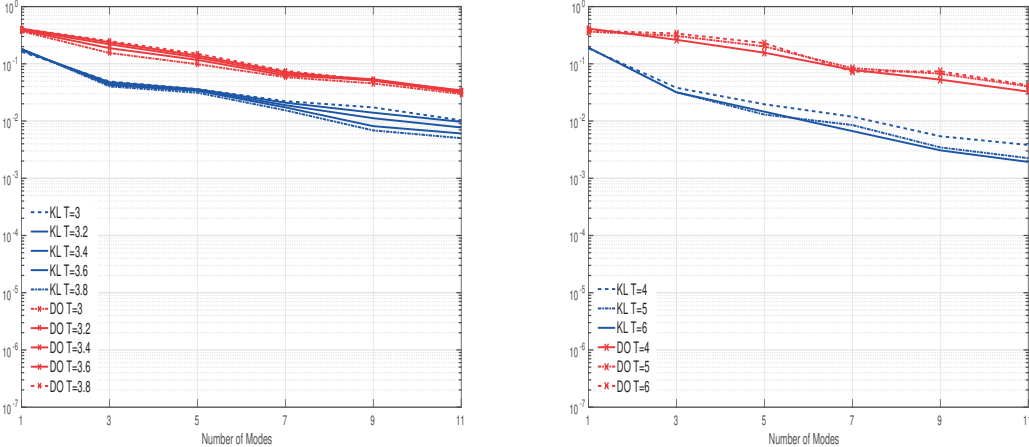
Figure 4.7 – The Dual DO approximation error (red) and the KL truncation error (blue) in norm $H^1(D) \otimes L^2(\Omega)$ with respect to the number of modes, at different time steps (from transition to periodic phase).
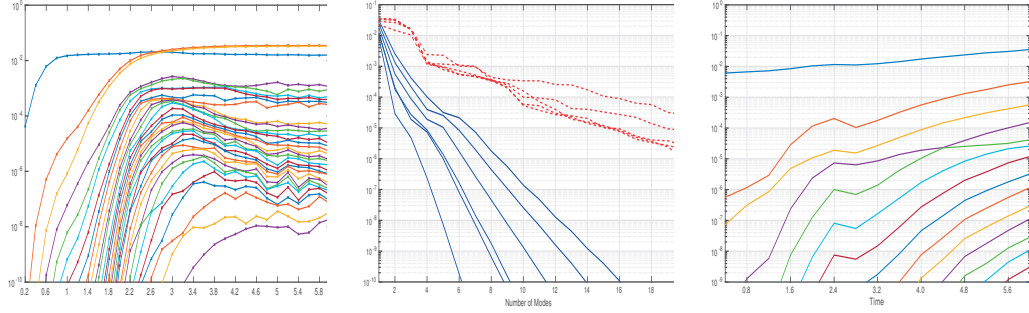
Figure 4.8 – Left: the time evolution of the eigenvalues of u^{KL}, i.e. the Karhunen-Loève decomposition of the reference solution, computed with the stochastic collocation method in a set of $N_y = 33$ Gauss Legendre collocation points. Middle: Decay of the eigenvalues of the Karhunen-Loève decomposition of the reference solution without (red dashed line) and with (blue solid line) time rescaling at different times. Right: the time evolution of the eigenvalues of u^{KL}, i.e. the Karhunen-Loève decomposition of the reference solution with time rescaling.

approximation with respect to the number of modes in norm $L^2(D) \otimes L^2(\Omega)$. We observe that the decay of the singular values is relatively slow. Moreover, we see that the decay significantly changes in time, meaning that the problem is not apt to be approximated by low-rank methods with fixed rank. To overcome these problems we propose here a strategy aiming at reformulating the original problem in a new coordinate system in order to obtain a solution that can be suitably approximated in low-rank format. First of all, we recall that, for deterministic values of the input parameter, namely for a fixed value of the Reynold's number, the flow features a periodic vortex shedding. In this case POD procedures from snapshots properly collected at different time instants, leads to accurate reconstructions of the solution with few modes. However, as numerically shown e.g. in [99, 85], the accuracy of these methods rapidly deteriorates as one slightly moves away from the parametric value used for the construction of the basis. This is because the input parameters affect the frequency and the length of the recirculation region. In particular, when the boundary conditions at the inflow are modeled as in (4.49), the solutions are velocity fields with varying vortex shedding frequencies which become more and more "out of phase" as time evolves (this explains the increasing rank of the solution in time). In light of that, our goal is to find a transformation which realigns all the solutions and keeps the rank small. For this purpose, we make use of an empirical formula [130] that linearly relates the vortex shedding frequency to the maximum velocity at the inflow: $f_s \propto \frac{U_{max}}{D}$ (where D here is the diameter of the cylinder). We recall that for the problem under consideration $U_{max} = U_m + \sigma z(\omega)$ so we claim that the frequency is linear in the random parameter z. Then, let us consider the fluid motion from a Lagrangian point of view. We define $X(x, s; t)$ the trajectory of the particle that at the instant $t = s$ passes through the point x, and we denote by $\tau = t - s$ the interval of time that the same particle needs to go from x to $x_1 = X(x, s; t)$. We recall that the Navier Stokes equations can be written in Lagrangian form as:

$$\begin{cases} \frac{D\mathbf{u}}{Dt} - \nu\Delta\mathbf{u} + \nabla p = \mathbf{f} \\ \nabla \cdot \mathbf{u} = 0 \end{cases} \tag{4.50}$$

Observe, however, that in our case the motion is a random field. This implies that, depending on the realization ω, the same particle will need a random interval of time to go from x to x_1. Because of that we define $X(x, s; \tau(\omega))$ the trajectory of the particle that for the realization ω was in x at time s. Observe that now the time is function of the random variable as also the period of the vortex shedding, defined $T = \frac{1}{f_s}$. Our purpose is to find an explicit formula relating T to ω and recover the Eulerian formulation of the motion expressed in terms of the new (random) time variable τ, with respect to which, the period of motion is almost deterministic. By using the empirical formula $f_s \propto \frac{U_m + \sigma z(\omega)}{D}$, we define the new time variable as:

$$(t, \omega) \to \tau(t, \omega) := \frac{U_m + \sigma z(\omega)}{U_m} t$$

(in the following we denote with $\alpha(\omega) = \frac{U_m + \sigma z(\omega)}{U_m}$) and we denote with $\hat{\mathbf{u}}$ the velocity field as a function of τ (instead of t):

$$\begin{cases} \hat{\mathbf{u}}(x, \tau(t, \omega)) = \frac{\partial X}{\partial \tau}(x, s; \tau(t, \omega)) \\ X(x, s; \tau(s, \omega)) = x(\omega) \end{cases}$$

Observe that $\hat{\mathbf{u}} = \frac{1}{\alpha(\omega)} \frac{\partial X}{\partial t} = \frac{1}{\alpha(\omega)}\mathbf{u}$. Then we can rewrite the first equation in (4.50) with respect to $\hat{\mathbf{u}}$ and τ and we obtain:

$$\alpha^2 \frac{D\hat{\mathbf{u}}}{D\tau} - \alpha\nu\Delta\hat{\mathbf{u}} + \nabla p = \mathbf{f}$$

or equivalently

$$\frac{\partial\hat{\mathbf{u}}}{\partial\tau} + \hat{\mathbf{u}} \cdot \nabla\hat{\mathbf{u}} - \frac{1}{\alpha}\nu\Delta\hat{\mathbf{u}} + \nabla\hat{p} = \hat{\mathbf{f}}$$

In conclusion the problem becomes:

$$\begin{cases} \frac{\partial\hat{\mathbf{u}}}{\partial\tau} + \hat{\mathbf{u}} \cdot \nabla\hat{\mathbf{u}} - \frac{1}{\alpha}\nu\Delta\hat{\mathbf{u}} + \nabla\hat{p} = \mathbf{0} \\ \nabla \cdot \hat{\mathbf{u}} = 0 \end{cases} \tag{4.51}$$

with deterministic boundary conditions at the inflow:

$$\hat{\mathbf{u}}(\mathbf{x}, \omega, \tau) = (4\bar{U}x_2(H - x_2)/H^2, 0) \quad \mathbf{x} = (x_1, x_2) \in \Gamma_{in}$$

Observe that now the diffusion coefficient is a random variable. We now apply the Dual DO method to problem (4.51) and we recover the approximate solution of the original problem as $\mathbf{u}_S(\mathbf{x}, \omega, t) = \alpha(\omega)\hat{\mathbf{u}}_S(\mathbf{x}, \omega, \tau(t, \omega))$. The advantage is that $\hat{\mathbf{u}}$ can be more easily approximated in low rank format. Indeed the rate of decay of the singular values of $\hat{\mathbf{u}}$ is significantly faster that the one of $\mathbf{u}$, see Figure 4.8 (Middle). Figure 4.8 (Right) shows instead the time evolution of the eigenvalues of the covariance function of $\hat{\mathbf{u}}$. In Figure 4.9 (Left), the Dual DO approximation is
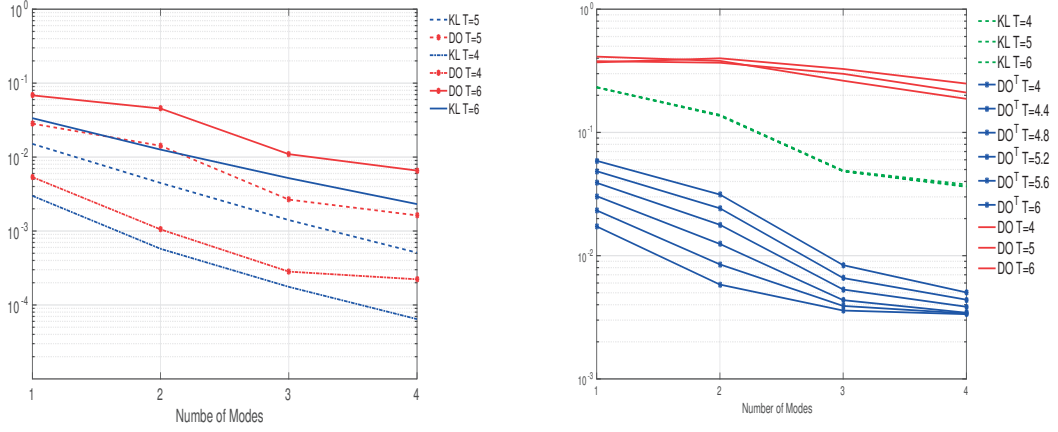
Figure 4.9 – Left: $H^1(D) \otimes L^2(\Omega)$ approximation error with time rescaling. The Dual DO approximation error (red) is compared to the best approximation error (blue) as the number of modes increases and at different time steps. Right: the Dual DO approximation error without (red) and with rescaling (blue, denoted by DO^T) and the best approximation error, all computed in norm $H^1(D) \otimes L^2(\Omega)$ and w.r.t. the reference solution in the original coordinates.

compared to the optimal one. The error is computed in $H^1(D) \otimes L^2(\Omega)$ norm for $\hat{\mathbf{u}}$, so before recovering the approximate solution in the original coordinates. We see that good levels of accuracy can be achieved with very few modes.

We conclude by analyzing the accuracy of the time rescaling technique, once the approximate solution $\mathbf{u}_S = \alpha\hat{\mathbf{u}}_S$ in the original coordinates is recovered. The performances of the time rescaled Dual DO are compared to the Dual DO method applied directly to the original problem (Section 6.2). In Figure 4.9 (Right) we compare the approximation error of the two approaches: Dual DO approximation without or with time rescaling. Both the approximate solutions are compared to the reference solution in the original coordinates and the error is computed in $H^1(D) \otimes L^2(\Omega)$ norm. We see that remarkable advantages are obtained by the second approach. Good levels of accuracy can be obtained with very few modes and the error appears to be also smaller than the optimal approximation error of $\mathbf{u}$ (solution without time rescaling) with the same number of modes. However, we remark that the error tends to increases for a longer time interval, probably due to the fact that we use an empirical formula to approximate the frequency of the solution, quantity that also is very sensitive to computational errors. However, this problem seems to be overcome by increasing the number of modes. For the example under consideration, 5 modes are enough to achieve very good levels of accuracy which remains approximately constant in time for the whole computational time interval.
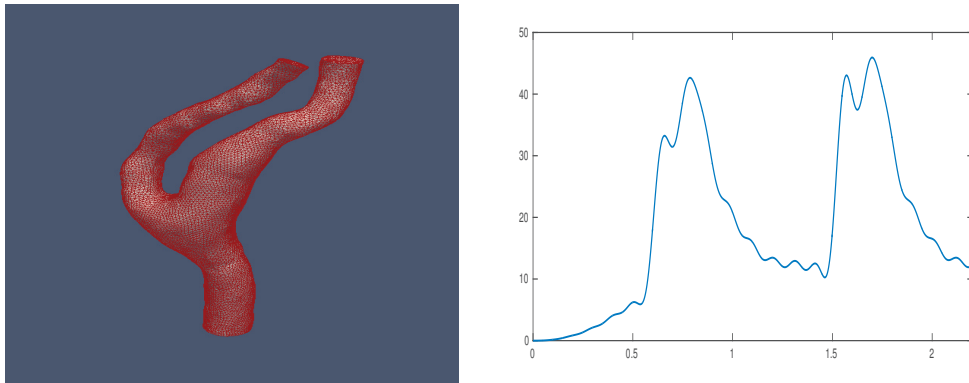
Figure 4.10 – Left: Computational mesh of the carotid artery, having 171123 cells and 34246 vertices. Right: The flow rate at the center of the inflow surface. The data correspond to two heart beats, with an initial quadratic ramp to go smoothly from zero flow rate to the physiological one.
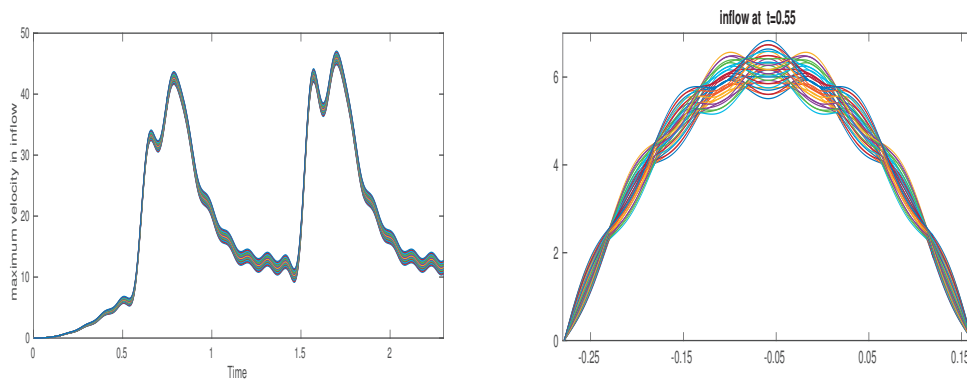
Figure 4.11 – Left: Time evolution of the maximum flow rate in the stochastic collocation points. It corresponds to two heart beat (plus an initial smoothing to agree with the uniform initial condition). Right: Inlet profile in the stochastic collocation points.
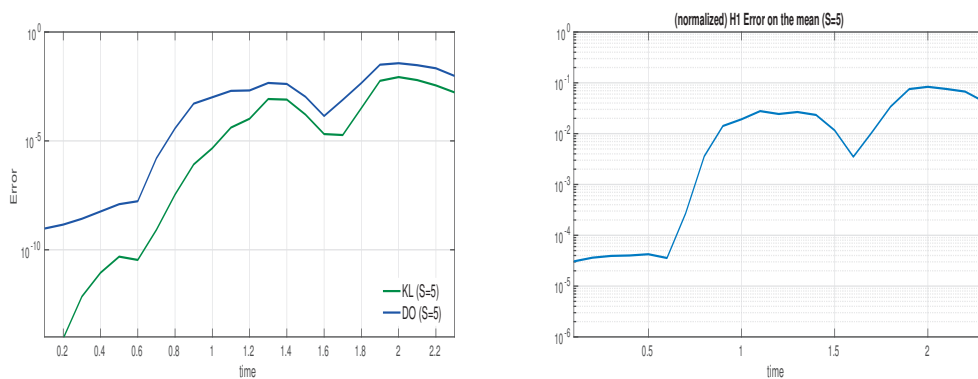
Figure 4.12 – Left: Dual DO approximation error (blue) compared to the best approximation error under boundary constraints. The error is computed in norm $[H^1(D)]^3 \otimes L^2(\Omega)$ with a number of modes $S = 5$. Right: Dual DO approximation error of the mean in norm $[H^1(D)]^3$ with a number of modes $S = 5$.
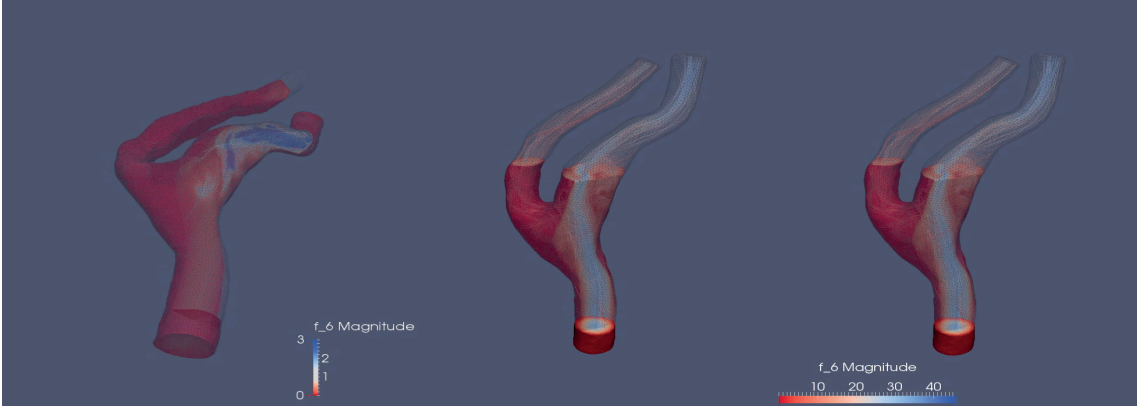
Figure 4.13 – On the left the standard deviation of the solution at time $t = 1.6$ during the second simulated heart beat. On the right we compare the mean of the Dual DO approximate solution computed with 5 modes (right) ($S = 5$) to the mean of the reference solution at the same time. We observe that the approximate solution effectively describes the dynamic and allow to accurately quantify the variability of the solutions.

4.5.2 Hemodynamic application

We now consider the Dual DO method for a hemodynamic problem with real data. Here the Dual DO method has been applied to simulate the blood flow in a realistic carotid artery reconstructed from MRI data (MRI images from the *Vascular-surgery and Radiology Divisions at Fondazione IRCSS Ca' Granda, Ospedale Maggiore Policlinico, Milan*): in Figure 4.10 (left) the mesh used in the numerical simulation, having 171123 cells and 34246 vertices; Figure 4.10 (right) shows the physiological pulse wave velocity imposed at the inlet, which corresponds to two heart beats.

We apply a non-homogeneous Dirichlet boundary condition at the inflow, a homogeneous Neumann condition at the outflow and non-slip conditions at the arterial walls. We assumed random inflow conditions due to possible errors in the Doppler measurements of the axial blood velocity at the inflow section. Specifically at the inlet we consider a parabolic velocity profile perturbed by two uniform and independent random variables Z_1 and Z_2 in $[-1, 1]$ that vary the maximum flow rate and slightly the shape:

$$\mathbf{u}_{|_{\Gamma_{in}}}(\mathbf{x}, t, \omega) = \left(0, 0, (f_b(t) + Z_1(\omega))(1 - (\frac{x^1 - x_c^1}{R})^2 - (\frac{x^2 - x_c^2}{R})^2) + Z_2(\omega) \cos(\frac{9(x^1 - x_c^1)}{2R}\pi) \cos(\frac{9(x^2 - x_c^2)}{2R})\right)$$
(4.52)

(x_c^1, x_c^2) are the coordinates of the center of the inflow section, R is the radius and f_b is the flow rate in figure 4.10 (right). In Figure 4.11, the maximum flow rate (left) and the inlet profile (right) for different values of Z_1 and Z_2 is shown. We refer to [108, 15] for the details about the typical numerical and physiological parameters.

We consider the Dual DO formulation with the isolation of the mean describe in Section 4. Let
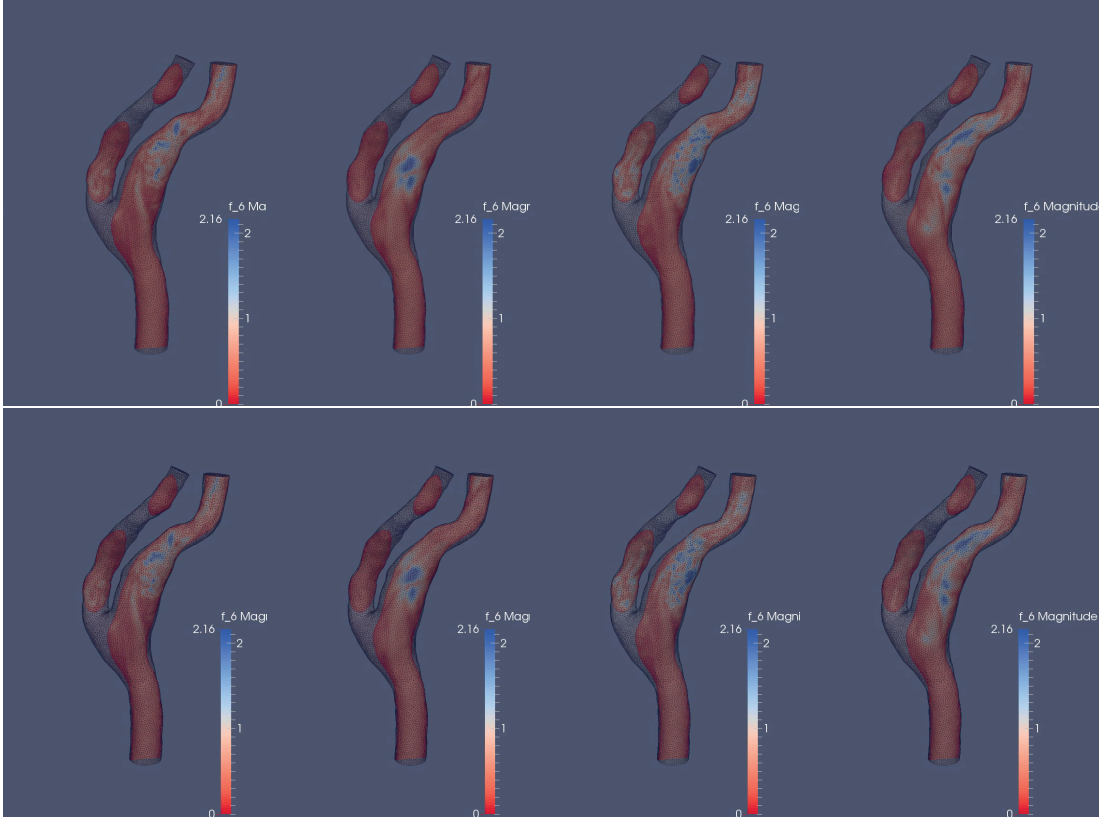
Figure 4.14 – The first deterministic mode of the Dual DO approximate solution with $S = 5$ (on the top) compared and the first eigen-mode of the best approximate solution (on the bottom) at different time

us write the boundary conditions at the inflow as:

$$\begin{aligned} \mathbf{u}(\mathbf{x}, t, \omega) &= \bar{\mathbf{g}}(\mathbf{x}, t) + \mathbf{g}_2(\mathbf{x}, t, \omega) \qquad \mathbf{x} \in \Gamma_{in} \\ \mathbf{g}_2(\mathbf{x}, t, \omega) &= \left(0, 0, (Z_1(\omega))(1 - (\tfrac{x^1 - x_c^1}{R})^2 - (\tfrac{x^2 - x_c^2}{R})^2) + Z_2(\omega) \cos(\tfrac{9(x^1 - x_c^1)}{2R}\pi) \cos(\tfrac{9(x^2 - x_c^2)}{2R})\right) \end{aligned}$$

and let $\mathcal{M}_{S,div}^{\mathbf{g}_2(t)}$ denote the manifold of all the divergence free S rank random fields that satisfy the boundary condition $\mathbf{g}_2(t)$ in Γ_D for a fixed $t \in [0, T]$. Hence the approximate solution is sought in the form:

$$\mathbf{u}_S(\mathbf{x}, t, \omega) = \bar{\mathbf{u}}_S(\mathbf{x}, t) + \underbrace{\sum_{i=1}^{S-2} \mathbf{U}_i(\mathbf{x}, t) Y_i(t, \omega) + \sum_{i=1}^{2} \mathbf{U}_i(\mathbf{x}, t) Z_i(\omega)}_{\mathbf{u}_S^*}$$

where $\mathbf{u}_S^*(\cdot, t, \cdot) = \mathbf{u}_S - \bar{\mathbf{u}}_S$ belongs to $\mathcal{M}_{S,div}^{\mathbf{g}_2(t)}$ and $\bar{\mathbf{u}}_S$ is equal to $\bar{\mathbf{g}}$ on Γ_D. The Dual DO reduced system is as in (4.5.1). By using the same discretization technique discussed in Section 4.5.1, at each time step we solve $S + 1$ decoupled deterministic PDEs and $S - 2$ ODEs. All implementations has been developed within the open source Finite Element library FEniCs,
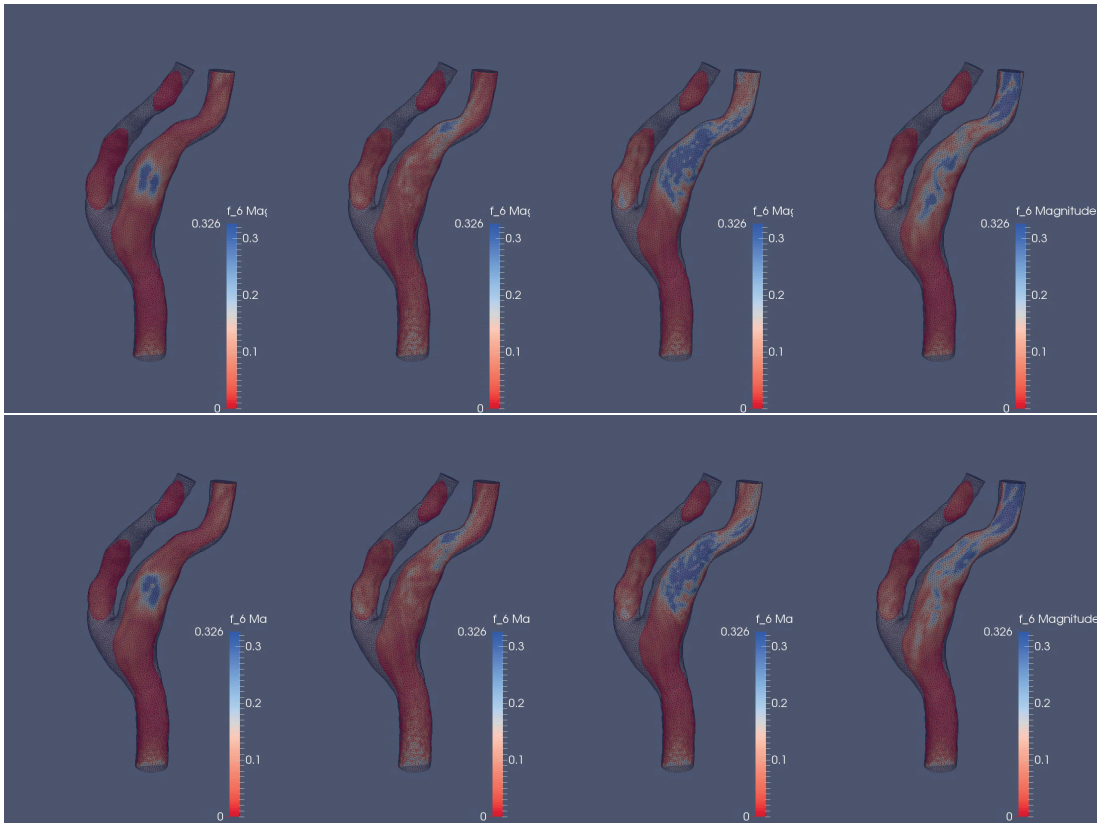
Figure 4.15 – The second deterministic mode of the Dual DO approximate solution with $S = 5$ (on the top) compared and the first eigen-mode of the best approximate solution (on the bottom) at different time

programmed and parallelized in Python. We report here the results obtained for $S = 5$. In Figure 4.12 (left) we compare the Dual DO and the best in norm $[H^1(D)]^3 \otimes L^2(\Omega)$ as time evolves. We observe that the two errors are proportional and the Dual DO approximate solution stays close to the best S rank approximation. The same conclusions can be drawn by comparing the Dual DO approximation of the mean to the mean of the reference solution, see Figure 4.12 (right) and Figure 4.13. Finally in Figure 4.14, 4.15, 4.16 we compare the Dual DO modes to the modes of the best approximation. We see that the Dual DO modes adapt properly to describe the variability of the solution. In conclusion for this case, the Dual DO method leads to very good results, in term of accuracy versus computational cost.

4.6 Conclusion

In this work, we have proposed a convenient strategy to strongly impose random Dirichlet boundary conditions in the dynamically low-rank approximation of parabolic PDEs with random parameters. We showed that the set of S rank random fields, constrained to satisfy an approximation of the boundary datum of the exact solution, can be equipped with the
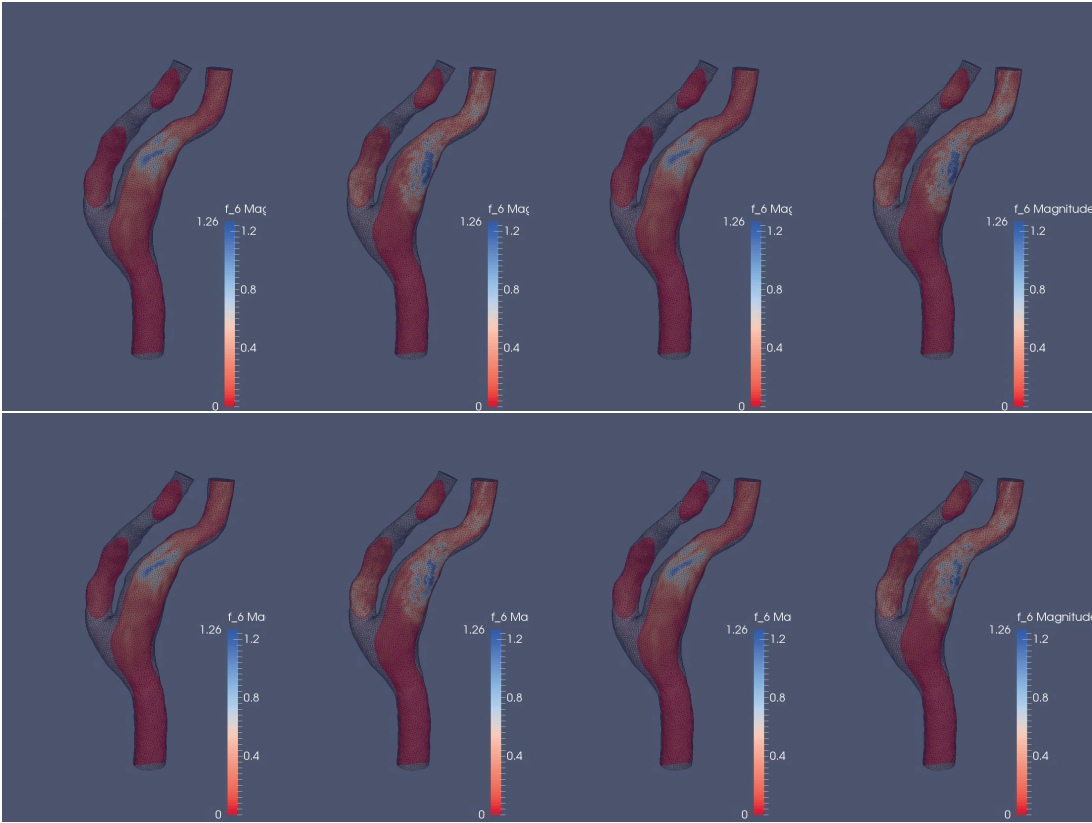
Figure 4.16 – The third deterministic mode of the Dual DO approximate solution with $S = 5$ (on the top) compared and the first eigen-mode of the best approximate solution (on the bottom) at different time

structure of a differential manifold, allowing for a parametrization of its tangent space in terms of dynamical constraints on the stochastic coefficients. To do so we proposed a Dual DO formulation in which the stochastic modes are kept orthonormal. Under the assumption that the boundary datum g can be properly approximated by a linear combination g_M of $M < S$ terms written in separable form, we fixed M stochastic modes in the approximate solution equal to those in the decomposition of g_M. This allowed us to identify the proper boundary conditions for each (time dependent) deterministic mode and guarantees that the boundary constraint is fulfilled at each time. We obtained a reduced system which consists of a set of S coupled PDEs for the evolution of the deterministic modes, M of which with non homogeneous boundary conditions, coupled with $S - M$ ODEs for the evolution of the stochastic modes. This resulted in an efficient dynamical low-rank approximation which accurately takes into account the randomness arising from the boundary data at the price of a slightly reduced flexibility in the evolution of the random modes. Furthermore, we observed that Dual DO formulation is also very convenient to include the incompressibility constraint when dealing with incompressible Navier-Stokes equations. Indeed we were able to effortlessly imposed the solenoidal constraint in each deterministic mode, facilitated by the fact that in the Dual DO formulation no numerical orthonormalization or dynamic constraint is required in the deterministic modes. In conclusion, Navier Stokes equations with random parameters, including random Dirichlet boundary conditions, has been reduced to S coupled deterministic PDEs of Navier-Stokes type and a system of $S - M$ stochastic ODEs.

We tested the potential and limitations of the proposed method on the classical benchmark 2D problem of an incompressible viscous fluid flowing around a cylindrical obstacle in a channel at moderate Reynold numbers Re$\in [80, 120]$, by adding some randomness in the inflow velocity. The numerical results obtained show good performance of the method, at least at the initial phase, but a loss of accuracy for long time integration. We observed that this is intrinsically due to the fact that the flow patterns become more and more out of phase one with respect to the others, as time evolves, requiring an increasing rank in time to keep a prescribed accuracy level. We numerically showed that a simple time rescaling based on an empirical linear relation between Reynolds number and shedding frequency considerably improves the performance of the method and allows to "rephase" all solutions. Finally, we highlighted the potentiality of the Dual DO method for biomedical applications, by simulating blood flows in a realistic carotid artery reconstructed from MRI data, with random inflow boundary conditions. The numerical results reported here, show that good level of accuracy can be achieved with only a few modes.

5 | Symplectic Dynamical Low-Rank approximation

In this Chapter, we propose a dynamical low-rank strategy for the approximation of second order wave equations with random parameters. The governing equation is rewritten in Hamiltonian form and the approximate solution is sought in the low dimensional manifold of all complex-valued random fields with rank equal to S. Recast in the real setting, the approximate solution is expanded over a set of $2S$ dynamical symplectic-orthogonal deterministic modes and satisfies the symplectic projection of the governing Hamiltonian system into the tangent space of the approximation manifold along the approximate trajectory. We recover a reduced order system for the evolution of both the stochastic coefficients and the deterministic basis. That guarantees the conservation of the average energy over the flow.

Introduction

Second order wave equations with random parameters, such as acoustic or elastic wave equations with uncertain random speed and/or source terms, appear in a large number of physical and engineering problems. Realistic applications are found for instance in seismology, where the propagation of the seismic waves strongly depends on source location, i.e. the epicenter of an earthquake, and density or elastic modulus of the medium. These parameters are inevitably affected by uncertainty.

In the context of low-rank approximation of wave equations, and more generally hyperbolic equations, a critical issue is how to construct reduced order systems which preserve the stability and the geometrical properties of the original problem. It has been reported in literature [109], that a POD-based reduced system may indeed become unstable even if the original hyperbolic systems was not. This calls for the development of "novel" reduced order techniques, built *ad hoc* to deal with hyperbolic problems. A possible strategy consists in designing low-rank techniques that preserve the underlying geometric structure of the full order system, which has generally a fundamental impact on the dynamic of the solutions. This leads to reduced order systems which enjoy the same conservation properties as the original full order system, and are better suited for long time integration and stability preservation. This direction has been initially investigated in [73] where the authors derive reduced order systems

which preserve the Lagrangian structure of the full order system. The same strategy is used in [24] and combined with the Gappy POD method to further reduce the computational cost. In the context of parametric Hamiltonian systems, recent works [107, 86] have proposed a reduced order method with symplectic bases, designed in analogy to the proper orthogonal decomposition, in which the standard Galerkin projection is replaced by a symplectic projection and the solution is approximated in a low dimensional symplectic space. The reduced order system consists then in a Hamiltonian system of small size which preserves the symplectic structure of the full order system, is energy conservative and preserve stability.

In this work, we propose a dynamical low-rank technique for the approximation of second order wave equations with random parameters. The aim is to exploit the flexibility of the DLR approximation, which, thanks to the use of dynamic modes, is well suited to approximate wave propagations. In this context, the challenge is to find a correct characterization of the manifold and parametrization of its tangent space at each point. Moreover, being the governing equation hyperbolic, special care is needed to preserve the stability of the reduced system. Our strategy, which we name Symplectic Dynamical Low Rank (Symplectic DO) method, is closely related to the Multi-Configuration Time-Dependent Hartree (MCTDH) method, used in [67, 39] for the approximation of deterministic Schrödinger equations and is based on recasting the governing equation in Hamiltonian form. In particular, the Symplectic DO method is designed as a combination of:

- the Dynamically Orthogonal (DO) method, introduced in [116, 117] for the approximation of parabolic equations with random coefficients;

- the symplectic reduced basis technique proposed in [107] in the context of parametric Hamiltonian systems.

The result is a reduced order method with symplectic dynamical deterministic bases which enjoys the conservation of energy, as the original problem. The approximate solution is expanded over a set of $2S$ (time dependent) symplectic-orthogonal deterministic modes, with (time dependent) stochastic coefficients and satisfies the symplectic projection of the Hamiltonian system into the tangent space to the approximation manifold along the approximate trajectory. We show that recast in the complex setting, this coincides with looking for a dynamical low-rank approximation of the governing complex-valued Hamiltonian system, into the low dimensional manifold of all the complex-valued random fields with S rank. Thanks to this analogy, we are able to obtain a proper parametrization of the tangent space, in terms of orthogonal constraints on the dynamics of the deterministic modes, and to derive the reduced dynamical system. The latter consists of a set of equations for the constrained dynamics of the deterministic modes, coupled with a reduced order Hamiltonian system for the evolution of the stochastic coefficients. The Symplectic DO shares with the symplectic order reduction the use of symplectic deterministic bases, and, as the "classic" DO approximation, allows both the stochastic and the deterministic modes to evolve in time.

5.1 Notation and problem setting

Let $\mathbb{F}$ stand for $\mathbb{R}$ or $\mathbb{C}$, and D be an open bounded subset of $\mathbb{R}^d$, $1 \le d \le 3$, with a smooth boundary ∂D. $L^2(D, \mathbb{F})$ (respectively $H^1(D, \mathbb{F})$) denotes the Hilbert space of square integrable functions (respectively with square square integrable partial derivatives) on D with values in $\mathbb{F}$. When $\mathbb{F}$ is omitted we always refer to $\mathbb{R}$. We denote by $\langle \cdot, \cdot \rangle$ the real inner product in $L^2(D, \mathbb{R})$, and with $\langle \cdot, \cdot \rangle_h$ the Hermitian inner product in $L^2(D, \mathbb{C})$, which is defined as:

$$\langle \hat{u}, \hat{v} \rangle_h := \langle u^q, v^q \rangle + \langle u^p, v^p \rangle + i(\langle u^p, v^q \rangle - \langle u^q, v^p \rangle) \qquad \forall \hat{u} = u^q + i u^p, \hat{v} = v^q + i v^p \in L^2(D, \mathbb{C}), \tag{5.1}$$

Hereafter, complex valued functions are denoted with the overhat symbol ($\hat{u}$), with real and complex components labeled with the apex q and p respectively ($\hat{u} = u^q + i u^p$).
We define the Stiefel manifold $St(S, H^1(D, \mathbb{F}))$, as the set of L^2-orthonormal frames of S functions in $H^1(D, \mathbb{F})$, i.e.:

$$St(S, H^1(D, \mathbb{F})) = \left\{ \mathbf{V} = (V_1, ..., V_S) : V_i \in H^1(D, \mathbb{F}) \text{ and } \langle V_i, V_j \rangle_* = \delta_{ij} \ \forall i, j = 1, ..., S \right\} \tag{5.2}$$

where $\langle V_i, V_j \rangle_*$ is the real L^2 product if $\mathbb{F} = \mathbb{R}$ and the hermitian product if $\mathbb{F} = \mathbb{C}$. We denote by $\mathscr{G}(S, H^1(D, \mathbb{F}))$ the Grassmann manifold of dimension S that consists of all the S-dimensional linear subspaces of $H^1(D, \mathbb{F})$. The definition of Stiefel and Grassmann manifold can be generalized to vector-valued functions in $[H^1(D, \mathbb{F})]^d$.
Let $(\Omega, \mathscr{A}, \mathscr{P})$ be a complete probability space, where Ω is the set of outcomes, $\mathscr{A}$ a σ-algebra and $P : \mathscr{A} \to [0, 1]$ a probability measure. Let $v : \Omega \to \mathbb{F}$ be an integrable random variable; we define the mean of v as:

$$\bar{y} = \mathbb{E}[y] = \int_\Omega y(\omega) d\mathscr{P}(\omega).$$

$L^2(\Omega, \mathbb{F})$ (respectively $L_0^2(\Omega, \mathbb{F})$) denotes the Hilbert space of square integrable random variables (respectively with zero mean), that is:

$$L^2(\Omega, \mathbb{F}) := \left\{ y : \Omega \to \mathbb{F} : \mathbb{E}[y^2] = \int_\Omega (y(\omega))^2 d\mathscr{P}(\omega) < \infty \right\}$$

We also recall that $L^2(D \times \Omega, \mathbb{F})$ denotes the space of all square integrable random fields, i.e.:

$$L^2(D \times \Omega, \mathbb{F}) := \left\{ u : D \times \Omega \to \mathbb{F} \text{ s.t. } \mathbb{E}\left[\|u\|_{L^2(D, \mathbb{F})}^2 \right] < \infty \right\}.$$

Observe that $L^2(D \times \Omega, \mathbb{F})$ is isometrically isomorphic to the tensor product space $L^2(D, \mathbb{F}) \otimes L^2(\Omega, \mathbb{F})$.

5.1.1 Wave equation with random parameters

We consider the following initial boundary value problem: find a random function $u : \bar{D} \times [0, T] \times \Omega \to \mathbb{R}$, such that P-almost everywhere in Ω (almost surely) the following holds:

$$\begin{cases} \ddot{u}(\mathbf{x},t,\omega) = \nabla \cdot \Big(c(\mathbf{x},\omega)\nabla u(\mathbf{x},t,\omega) \Big) + f(u(\mathbf{x},t,\omega),\omega) & \mathbf{x} \in D, \ t \in (0,T], \ \omega \in \Omega, \\ u(\mathbf{x},0,\omega) = p_0(\mathbf{x},\omega) & \mathbf{x} \in D, \ \omega \in \Omega, \\ \dot{u}(\mathbf{x},0,\omega) = q_0(\mathbf{x},\omega) & \mathbf{x} \in D, \ \omega \in \Omega, \\ u(\boldsymbol{\sigma},t,\omega) = 0 & \boldsymbol{\sigma} \in \partial D, \ t \in (0,T], \ \omega \in \Omega, \end{cases} \tag{5.3}$$

For convenience we restrict in this work to homogeneous Dirichlet boundary conditions, although the development hereafter generalizes easily to other types of boundary conditions, either homogeneous or non-homogeneous with deterministic forcing terms. The case of non-homogeneous stochastic boundary conditions can be treated as in Chapter 4 (or [95]) but will not be detailed in this work. For the well-posedness of problem (5.3), we assume that the random wave speed c is bounded and uniformly coercive [118, 94]:

$$0 < c_{min} \le c(\mathbf{x},\omega) \le c_{max} < \infty \qquad \forall \mathbf{x} \in D, \ a.s.,$$

and the initial data satisfy: $q_0 \in L^2(\Omega, H_0^1(D))$, $p_0 \in L^2(\Omega, L^2(D))$. Here the randomness may affect the wave speed c as well as the initial conditions p_0, q_0 and the (possibly non linear) source term f. Our goal is to find a dynamical low rank approximation of the solution of problem (5.3).

5.2 DO approximation

We recall that the Dynamically Orthogonal method (DO) [116, 117, 96] is a reduced basis technique used for the approximation of parabolic equations with random parameters. Consider the following general real valued problem:

$$\begin{cases} \dot{u}(\mathbf{x},t,\omega) = \mathcal{L}(u(\mathbf{x},t,\omega),t,\omega), & \mathbf{x} \in D, \ t \in (0,T], \ \omega \in \Omega, \\ u(\mathbf{x},0,\omega) = u_0(\mathbf{x},\omega) & \mathbf{x} \in D, \ \omega \in \Omega, \\ \mathcal{B}(u(\boldsymbol{\sigma},t,\omega),\omega) = g(\boldsymbol{\sigma},t) & \boldsymbol{\sigma} \in D, \ t \in (0,T], \ \omega \in \Omega, \end{cases} \tag{5.4}$$

where $\mathcal{L}$ is a linear or non-linear differential operator, $x \in D$ is the spatial coordinate and t is the time variable in $[0,T]$. Here $\omega \in \Omega$ represents a random elementary event which may affect the operator $\mathcal{L}$ (as e.g. a coefficient or a forcing term) or the initial conditions. Let us assume that the solution $u(\cdot,t,\omega)$ to problem (5.4) is in a certain real Hilbert space $\mathcal{H} \subset L^2(D)$ for (almost) all $t \in [0,T]$ and $\omega \in \Omega$ and that $\mathcal{L}(u,t,\omega) \in \mathcal{H}'$ for all $u \in \mathcal{H}$ and almost everywhere in $[0,T]$ and Ω. Hereafter, whenever no confusion arises, we may write simply $\mathcal{L}(u)$ instead of $\mathcal{L}(u,t,\omega)$. The approximation manifold consists of the collection of all S rank random fields, i.e functions that can be exactly expressed as linear combination of S linearly independent deterministic modes combined with S linearly independent stochastic modes.

Definition 5.2.1. *We define $\mathcal{M}_S \subset \mathcal{H} \otimes L^2(\Omega)$ the manifold of all S rank random fields, i.e.:*

$$\mathcal{M}_S = \left\{ u_S \in \mathcal{H} \otimes L^2(\Omega) : u_S = \sum_{i=1}^{S} U_i Y_i \,\middle|\, \begin{array}{l} span(U_1, ..., U_S) \in \mathcal{G}(S, \mathcal{H}), \\ span(Y_1, ..., Y_S) \in \mathcal{G}(S, L^2(\Omega)) \end{array} \right\} \tag{5.5}$$

The DO approximate solution is sought in $\mathcal{M}_S$ and satisfies the following variational principle:

DLR Variational Principle. *At each $t \in [0, T]$, find $u_S(t) \in \mathcal{M}_S$ such that: $u_S(0) = u_{0,S}$ and*

$$\mathbb{E}[\langle \dot{u}_S(\cdot, t, \cdot) - \mathcal{L}(u_S(\cdot, t, \cdot)), v \rangle] = 0, \quad \forall v \in \mathcal{T}_{u_S(t)}\mathcal{M}_S, \, t \in (0, T] \tag{5.6}$$

where $\mathcal{T}_{u_S(t)}\mathcal{M}_S$ is the tangent space to $\mathcal{M}_S$.

The variational principle (5.6) enforces the approximate solution u_S to satisfy the governing equation projected into the tangent to the approximation manifold along the solution trajectory. The initial datum $u_{0,S}$ is a suitable S rank approximation of u_0 by e.g. a truncated Karhunen-Loève expansion (best S rank approximation in the $L^2(D) \otimes L^2(\Omega)$ norm). In quantum mechanic this is known as Dirac-Frenkel time-dependent variational principle (see e.g. [80]) and leads to the MCTDH method [39, 67, 7] for the approximation of deterministic time-dependent Schrödinger equations.

There exist several possible parameterizations of a S rank random field. One option consists in assuming the deterministic modes orthonormal:

$$u_S(\mathbf{x}, \omega) = \sum_{i=1}^{S} Y_i(\omega) U_i(x) = \mathbf{U}\mathbf{Y} \tag{5.7}$$

where:

- $\mathbf{U} \in St(S, \mathcal{H})$ is a row vector of L^2−orthonormal deterministic functions,

- $\mathbf{Y}$ is a column vector of S random variables with full rank second moment matrix $\mathbf{C} = \mathbb{E}[\mathbf{Y}\mathbf{Y}^T]$.

In the following we denote by $B(S, L^2(\Omega))$ the set of all S frames of linearly independent random variables in $L^2(\Omega)$, i.e.:

$$B(S, L^2(\Omega)) = \{\mathbf{Y} = (Y_1, ..., Y_S) \in [L^2(\Omega)]^S \text{ s.t. } \text{rank}(\mathbb{E}[\mathbf{Y}\mathbf{Y}^T]) = S\}. \tag{5.8}$$

One easily sees that the representation (5.7) is not unique. For any orthogonal matrix $\mathbf{O} \in \mathcal{O}(S) \subset \mathbb{R}^{S \times S}$ one can always find a new couple of bases $\mathbf{W} = \mathbf{U}\mathbf{O} \in St(S, \mathcal{H})$ and $\mathbf{Z} = \mathbf{O}^T\mathbf{Y} \in B(S, L^2(\Omega))$ which represents the same S rank random field: $u_S = \mathbf{U}\mathbf{Y} = \mathbf{W}\mathbf{Z}$. The uniqueness of the decomposition (5.7), in terms of $\mathbf{U} \in St(S, \mathcal{H})$ and $\mathbf{Y} \in B(S, L^2(\Omega))$, is recovered by imposing the following constraint on the dynamics of $\mathbf{U}$ [66]:

$$\langle \dot{U}_i(t), U_j(t) \rangle = 0 \qquad i, j = 1, ..., S \tag{5.9}$$

This condition represents a quotientation of $St(S,\mathscr{H})$ with respect to the group of rotations $\mathscr{O}(S)$ and leads to the diffeomorphic identification of $\mathscr{M}_S$ with $\left(St(S,\mathscr{H})/\mathscr{O}(S)\right) \times B(S,L^2(\Omega))$. In particular (5.9) implies that the tangent bundle to $\left(St(S,\mathscr{H})/\mathscr{O}(S)\right)$ is parametrized in terms of the tangent vectors of $St(S,\mathscr{H})$ which are orthogonal to the equivalent classes of the quotientification. This procedure is based on classical results of fiber bundle theory. We refer interested readers to [74, 1, 64] for further details.

By means of (5.9), the tangent space to $\mathscr{M}_S$ at $u_S = \mathbf{U}\mathbf{Y}$ is parametrized as:

$$\mathscr{T}_{u_S}\mathscr{M}_\mathscr{S} = \left\{ \delta u = \sum_{i=1}^{S} \left(\delta Y_i U_i + Y_i \delta U_i\right) \in \mathscr{H} \otimes L^2(\Omega) : \langle U_i, \delta U_j \rangle = 0, \forall i, j = 1, ..., S \right\} \quad (5.10)$$

Then (5.6) leads to the DO reduced system [116, 117]: find $u_S(t) = \mathbf{U}(t)\mathbf{Y}(t)$, $t \in (0, T]$ such that

$$\begin{cases} \sum_{i=1}^{S} \dot{U}_i \mathbf{C}_{ij} = \mathscr{P}_{\mathbf{U}}^{\perp}\left(\mathbb{E}[\mathscr{L}(u_S)Y_j]\right) & \forall j = 1, ..., S \\ \dot{Y}_j = \langle \mathscr{L}(u_S), U_j \rangle & \forall j = 1, ..., S \end{cases} \quad (5.11)$$

where $\mathbf{C} = \mathbb{E}[\mathbf{Y}\mathbf{Y}^T] \in \mathbb{R}^{S \times S}$ and $\mathscr{P}_{\mathbf{U}}^{\perp}$ is the projection operator from the space $L^2(D)$ to the orthogonal complement of the S dimensional subspace $\mathscr{U} = \text{span}\{U_1, \cdots, U_S\}$, i.e. $\mathscr{P}_{\mathbf{U}}^{\perp}(v) = v - \mathscr{P}_{\mathbf{U}}(v) = v - \sum_{i=1}^{S} \langle v, U_i \rangle U_i$, $\forall v \in L^2(D)$. The initial condition is given by the truncated Karhunen-Loève expansion (the best S rank approximation in norm L^2) and the DO approximate solution is determined by solving (5.11).

The peculiarity of the DO method is that both the spatial and stochastic bases are computed on the fly and are free to evolve in time, thus adjusting at each time to the current structure of the random solution.

5.3 Symplectic Manifolds

Symplectic manifolds are the natural setting for Hamiltonians systems, due to the intrinsic symplectic structure of the canonical phase-space coordinates. We review in this section the main definitions and results concerning symplectic manifolds. For a comprehensive treatment see e.g. [88, 91].

Definition 5.3.1. *A symplectic manifold is a pair* $(\mathscr{V}, \vartheta)$ *consisting of a differential manifold* $\mathscr{V}$ *and a 2-form:*

$$\begin{aligned} \vartheta_u : \quad \mathscr{T}_u\mathscr{V} \times \mathscr{T}_u\mathscr{V} &\to \mathbb{R} \\ (y, z) &\to \vartheta_u(y, z) \end{aligned}$$

for any $u \in \mathscr{V}$, *which is:*

- *closed, i.e* $\mathbf{d}\vartheta = 0$ *where* $\mathbf{d}$ *is the exterior derivative.*

- *non-degenerate, i.e. for any* $u \in \mathscr{V}$ *and* $y \in \mathscr{T}_u\mathscr{V}$, $\vartheta_u(y, z) = 0$ *for all* $z \in \mathscr{T}_u\mathscr{V}$ *if and only if* $y = 0$.

The form ϑ is called symplectic form.

If $\mathcal{V}$ is a vector space, the requirement $\mathbf{d}\vartheta = 0$ is automatically satisfied since the ϑ_u is constant in u and Definition 5.3.1 is simplified as follows:

Definition 5.3.2. *Let $\mathcal{V}$ be a vector space and ϑ a bilinear map: $\vartheta : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ such that:*

- *ϑ is not degenerate, i.e. $\vartheta(y, z) = 0$ for all $z \in \mathcal{V}$ if and only if $y = 0$,*

- *ϑ is antisymmetric, i.e. $\vartheta(y, z) = -\vartheta(z, y)$ for any $y, z \in \mathcal{V}$.*

The pair $(\mathcal{V}, \vartheta)$ is called symplectic vector space.

Definition 5.3.3. *Let $(\mathcal{V}, \vartheta)$ be a symplectic vector space. A smooth submanifold $\mathcal{W} \subset \mathcal{V}$ is said symplectic if the restriction of ϑ to $\mathcal{W}$ is not-degenerate.*

Definition 5.3.4. *Let $(\mathcal{V}, \vartheta)$ be a symplectic vector space and $\mathcal{U}$ a subspace of $\mathcal{V}$. The symplectic complement of $\mathcal{U}$ is defined as:*

$$\mathcal{U}^{\perp, sym} = \left\{ z \in \mathcal{V} \text{ such that } \vartheta(z, y) = 0, \, \forall\, y \in \mathcal{U} \right\} \tag{5.12}$$

Unlike orthogonal complements, $\mathcal{U}^{\perp, sym} \cap \mathcal{U}$ is not necessary trivial. We start by recalling some properties of finite dimensional symplectic manifolds [77]. Afterwards, we look at the infinite dimensional case [129, 88].

Proposition 5.3.1. *All finite dimension symplectic vector spaces are even dimensional.*

This can be verified by observing that real skew-symplectic matrices of odd dimension must have a non trivial kernel. Since a symplectic form makes the tangent spaces into symplectic vector spaces, Proposition 5.3.1 actually applies to all finite dimension symplectic manifolds. Without further specification, in the following $\mathcal{V}_{2N}$ will always denote a finite dimensional manifold of dimension $2N$.

Theorem 5.3.1 (Darboux' theorem)**.** *Let $(\mathcal{V}_{2N}, \vartheta)$ be a symplectic manifold. For any $\mathbf{u} \in \mathcal{V}_{2N}$ there exists a neighborhood $\mathscr{B}_{\mathbf{u}} \subseteq \mathcal{V}_{2N}$ of $\mathbf{u}$ and a local coordinate chart in which ϑ is constant.*

Definition 5.3.5. *Let $(\mathcal{V}_{2N}, \vartheta)$ be a symplectic vector space. A basis $(\mathbf{e}_1, ..., \mathbf{e}_N, \mathbf{f}_1, ..., \mathbf{f}_N)$ of $\mathcal{V}_{2N}$ is said symplectic if:*

- *$\vartheta(\mathbf{e}_i, \mathbf{f}_j) = \delta_{ij} = -\vartheta(\mathbf{f}_j, \mathbf{e}_i), \, \forall\, i, j = 1, ..., N$,*

- *$\vartheta(\mathbf{e}_i, \mathbf{e}_j) = 0 = \vartheta(\mathbf{f}_i, \mathbf{f}_j), \, \forall\, i, j = 1, ..., N$.*

Darboux' theorem implies that, for any $\mathbf{u}$ in the symplectic manifold $(\mathcal{V}_{2N}, \vartheta)$ there is a neighborhood $\mathcal{B}_{\mathbf{u}} \subseteq \mathcal{V}_{2N}$ and a symplectic basis with respect to which the symplectic form is written as $\vartheta_{\mathbf{u}}(\mathbf{w}, \mathbf{v}) = \mathbf{w}^T \mathbf{J}_{2N} \mathbf{v}$ for all $\mathbf{w}, \mathbf{v} \in \mathcal{B}_{\mathbf{u}}$ (column vectors), where $\mathbf{J}_{2N} \in \mathbb{R}^{2N \times 2N}$ is the Poisson matrix, i.e.

$$\mathbf{J}_{2N} = \begin{pmatrix} 0 & \mathbb{I}_N \\ -\mathbb{I}_N & 0 \end{pmatrix}$$

and $\mathbb{I}_N$ is the identity matrix in $\mathbb{R}^{N \times N}$. It is easy to verify that $\mathbf{J}_{2N} \mathbf{J}_{2N}^T = \mathbf{J}_{2N}^T \mathbf{J}_{2N} = \mathbb{I}_{2N}$ and $\mathbf{J}_{2N} \mathbf{J}_{2N} = \mathbf{J}_{2N}^T \mathbf{J}_{2N}^T = -\mathbb{I}_{2N}$. When $\mathcal{V}_{2N}$ is a vector space, $\vartheta_{\mathbf{u}}$ is constant in $\mathbf{u}$ and $\mathcal{B}_{\mathbf{u}}$ corresponds to the whole space, namely $\vartheta(\mathbf{w}, \mathbf{v}) = \mathbf{w}^T \mathbf{J}_{2N} \mathbf{v}$ for all $\mathbf{v}, \mathbf{w} \in \mathcal{V}_{2N}$. If this symplectic basis coincides with the canonical basis of $\mathbb{R}^{2N}$ we call ϑ canonical symplectic form and we denote by $(\mathcal{V}_{2N}, \mathbf{J}_{2N})$ the corresponding symplectic manifold. A prototypical example of symplectic vector space arises from the identification of the complex space $\mathbb{C}^N$ with the real space $\mathbb{R}^{2N}$. Let us write elements of $\mathbb{C}^N$ as N-tuples of complex numbers $\hat{\mathbf{u}} = (\hat{u}_1, ..., \hat{u}_N)$, for each term $\hat{u}_i = u_i^q + i u_i^p$, the apex q and p denoting respectively the real and the complex components. Let $\mathbb{C}^N$ be equipped with the usual Hermitian inner product:

$$\langle \hat{\mathbf{u}}, \hat{\mathbf{v}} \rangle_h = \sum_{i=1}^{S} \hat{u}_i \hat{v}_i^* = \sum_{i=1}^{S} (u_i^q v_i^q + u_i^p v_i^p) + i \sum_{i=1}^{S} (v_i^q u_i^p - u_i^q v_i^p),$$

for any $\hat{\mathbf{u}}, \hat{\mathbf{v}} \in \mathbb{C}^N$. The realification, namely the identification of $\mathbb{C}^N$ with $\mathbb{R}^{2N}$, consists in associating to any $\hat{\mathbf{u}} \in \mathbb{C}^N$ the elements $\mathbf{u} = (\mathbf{u}^q, \mathbf{u}^p) \in \mathbb{R}^{2N}$, where $\mathbf{u}^q = (u_1^q, ..., u_N^q)$ and $\mathbf{u}^p = (u_1^p, ..., u_N^p)$. In the following we always use the overhat to distinguish complex elements $(\hat{\mathbf{u}})$ and corresponding real representations $(\mathbf{u})$. One can easily see that the canonical symplectic form of $\mathbb{R}^{2N}$ coincides with the imaginary part of the Hermitian product, with changed sign: $\mathbf{u}^T \mathbf{J}_{2N} \mathbf{v} = -\Im\langle \hat{\mathbf{u}}, \hat{\mathbf{v}} \rangle_h$, for all $\hat{\mathbf{u}}, \hat{\mathbf{v}} \in \mathbb{C}^N$.

We call symplectic matrix any $\mathbf{A} \in \mathbb{R}^{2N \times 2N}$ such that $\mathbf{A}^T \mathbf{J}_{2N} \mathbf{A} = \mathbf{J}_{2N}$. The collection of all symplectic matrices of $\mathbb{R}^{2N \times 2N}$ forms a group, called symplectic group.

Definition 5.3.6. *The symplectic group, denoted by $Sp(2N, \mathbb{R}^{2N})$, is the subset of $\mathbb{R}^{2N \times 2N}$ defined as:*

$$Sp(2N, \mathbb{R}^{2N}) := \{\mathbf{A} \in \mathbb{R}^{2N \times 2N} : \mathbf{A}^T \mathbf{J}_{2N} \mathbf{A} = \mathbf{J}_{2N}\}.$$

The unitary group is the subgroup of $Sp(2N, \mathbb{R}^{2N})$ of all unitary matrices

Definition 5.3.7. *The unitary group, denoted by $U(N, \mathbb{R}^{2N})$, is the subset of $\mathbb{R}^{2N \times 2N}$ defined as:*

$$U(N, \mathbb{R}^{2N}) := \{\mathbf{A} \in Sp(2N, \mathbb{R}^{2N}) : \mathbf{A}^T \mathbf{A} = \mathbf{A} \mathbf{A}^T = \mathbb{I}_{2N}\}.$$

In other words, $U(N, \mathbb{R}^{2N}) = Sp(2N, \mathbb{R}^{2N}) \cap \mathcal{O}(2N, \mathbb{R}^{2N})$, where $\mathcal{O}(2N, \mathbb{R}^{2N})$ denotes the group of orthogonal matrices in $\mathbb{R}^{2N \times 2N}$. Definitions 5.3.6 and 5.3.7 can be generalized to rectangular matrices $\mathbf{A} \in \mathbb{R}^{2N \times 2S}$ for any $0 < S < N$:

Definition 5.3.8. *We denote by $Sp(2S, \mathbb{R}^{2N})$ the sub-manifold of $\mathbb{R}^{2N \times 2S}$ defined as:*

$$Sp(2S, \mathbb{R}^{2N}) := \{\mathbf{A} \in \mathbb{R}^{2N \times 2S} : \mathbf{A}^T \mathbf{J}_{2N} \mathbf{A} = \mathbf{J}_{2S}\}$$

and by $U(S, \mathbb{R}^{2N})$, the submanifold of $\mathbb{R}^{2N \times 2S}$ defined as:

$$U(S, \mathbb{R}^{2N}) := \{\mathbf{A} \in Sp(2S, \mathbb{R}^{2N}) : \mathbf{A}^T \mathbf{A} = \mathbb{I}_{2S}\}.$$

We call symplectic (respectively unitary) matrix any $\mathbf{A} \in Sp(2S, \mathbb{R}^{2N})$ (respectively $\mathbf{A} \in U(S, \mathbb{R}^{2N})$).

Definition 5.3.9. *A linear map $\phi : \mathbb{R}^{2S} \to \mathbb{R}^{2N}$ defined as:*

$$
\begin{aligned}
\phi : \quad &\mathbb{R}^{2S} &\to \mathbb{R}^{2N} \\
&\mathbf{x} &\mapsto \phi(\mathbf{x}) := \mathbf{A}\mathbf{x}
\end{aligned}
\tag{5.13}
$$

is said symplectic if it preserves the canonical form, i.e

$$\mathbf{x}^T \mathbf{J}_{2S} \mathbf{x} = \phi(\mathbf{x})^T \mathbf{J}_{2N} \phi(\mathbf{x}) = (\mathbf{A}\mathbf{x})^T \mathbf{J}_{2N} \mathbf{A}\mathbf{x} \qquad \forall \mathbf{x} \in \mathbb{R}^{2S}$$

.

Observe that ϕ in (5.13) is symplectic if and only if $\mathbf{A} \in Sp(2S, \mathbb{R}^{2N})$.

In the same way as $\mathbb{R}^{2N}$ admits a (canonical) symplectic structure associated to the Euclidean product, all inner product vector spaces can be equipped with the symplectic form associated to their inner product, called again canonical form. In particular, consider the (possibly infinite dimensional) Hilbert space $\mathscr{H}$ and the product space $\mathcal{H} = [\mathscr{H}]^2$, for which we use the notation $\mathbf{u} = (u^q, u^p)$ to denote the first and second component of any $\mathbf{u} \in \mathcal{H}$. Let $\mathcal{H}$ be equipped with the usual inner product: $\langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{H}} := \langle u^q, v^q \rangle_{\mathscr{H}} + \langle u^p, v^p \rangle_{\mathscr{H}}$, for any $\mathbf{u}, \mathbf{v}$ in $\mathcal{H}$; we denote by $\mathscr{J}_2 : \mathcal{H} \to \mathcal{H}$ the following linear operator:

$$
\begin{aligned}
\mathscr{J}_2 : \quad &\mathcal{H} &\to \mathcal{H} \\
&\mathbf{u} &\mapsto \mathscr{J}_2(\mathbf{u}) := \begin{bmatrix} 0 & I_d \\ -I_d & 0 \end{bmatrix} \begin{bmatrix} u^q \\ u^p \end{bmatrix} = \begin{bmatrix} u^p \\ -u^q \end{bmatrix}
\end{aligned}
$$

where I_d is the identity operator in $\mathscr{H}$. Then, the canonical form of $\mathcal{H}$ is defined as

$$
\begin{aligned}
\vartheta : \quad &\mathcal{H} \times \mathcal{H} &\to \quad \mathbb{R} \\
&(\mathbf{u}, \mathbf{v}) &\mapsto \quad \langle \mathbf{u}, \mathscr{J}_2(\mathbf{v}) \rangle_{\mathcal{H}} = \langle u^q, v^p \rangle_{\mathscr{H}} - \langle u^p, v^q \rangle_{\mathscr{H}}
\end{aligned}
$$

The form ϑ is antisymmetric, being $\mathscr{J}_2 \circ \mathscr{J}_2(\mathbf{u}) = -\mathbf{u}$, for all $\mathbf{u} \in \mathcal{H}$, and non degenerate since $\vartheta(\mathscr{J}_2(\mathbf{u}), \mathbf{u}) = \langle \mathscr{J}_2(\mathbf{u}), \mathscr{J}_2(\mathbf{u}) \rangle_{\mathcal{H}} = \|\mathbf{u}\|^2_{\mathcal{H}}$ which is non zero for any $0 \neq \mathbf{u} \in \mathcal{H}$. Hence $(\mathcal{H}, \vartheta)$ is a symplectic vector space. We generally write $(\mathcal{V}, \mathscr{J}_2)$ to refer to a symplectic manifold $\mathcal{V} \subset \mathcal{H}$, when equipped with the canonical form of $\mathcal{H}$.

Proposition 5.3.2. *Let $\mathscr{H}^{\mathbb{C}}$ be a complex Hilbert space and $\mathscr{H} \times \mathscr{H}$ its realification. The*

Hermitian product of $\mathcal{H}^{\mathbb{C}}$, defined as:

$$\langle \hat{u}, \hat{v} \rangle_{\mathcal{H}^{\mathbb{C}}} := \langle u^q, v^q \rangle_{\mathcal{H}} + \langle u^p, v^p \rangle_{\mathcal{H}} + i(\langle u^p, v^q \rangle_{\mathcal{H}} - \langle v^p, u^q \rangle_{\mathcal{H}}) \qquad \forall \hat{u} = u^q + i u^p, \hat{v} = v^q + i v^p \in \mathcal{H}^{\mathbb{C}},$$

satisfies:

$$\langle (u^q, u^p), \mathscr{J}_2(v^q, v^p) \rangle_{\mathcal{H} \times \mathcal{H}} = -\Im(\langle \hat{u}, \hat{v} \rangle_{\mathcal{H}^{\mathbb{C}}})$$

for any $\mathbf{u} = (u^q, u^p), \mathbf{v} = (v^q, v^p) \in \mathcal{H} \times \mathcal{H}$ *and* $\hat{u} = u^q + i u^p, \hat{v} = v^q + i v^p \in \mathcal{H}^{\mathbb{C}}$.

This construction applies straightforwardly to $\mathcal{H} = [L^2(D)]^2$ equipped with the $[L^2(D)]^2$ inner product and to $\mathcal{H} = [H^1(D)]^2$ equipped either with the $L^2(D) \times L^2(D)$ or the $[H^1(D)]^2$ inner product. The identification in complex setting leads to $\mathcal{H}^{\mathbb{C}} = L^2(D, \mathbb{C})$ and $\mathcal{H}^{\mathbb{C}} = H^1(D, \mathbb{C})$ for $\mathcal{H}$ respectively equal to $\mathcal{H} = [L^2(D)]^2$ or $\mathcal{H} = [H^1(D)]^2$.

In view of the Symplectic Dynamical Low Rank approximation of wave equations we need to recast problem (5.3) into a Hamiltonian system, in terms of the phase-space coordinates $(u, \dot{u}) \in H^1(D) \times L^2(D)$. For this aim, we are interested to equip $\mathcal{H} = H^1(D) \times L^2(D)$ with the symplectic form associated to the $L^2(D) \times L^2(D)$ inner product and verify that what we obtain is still a symplectic space. The issue is due to the fact that now $\mathcal{H}$ is a product of two different Hilbert spaces. With a little abuse of notation, we use the same symbol $\mathscr{J}_2$ to denote the restriction of $\mathscr{J}_2$ to $H^1(D) \times L^2(D)$, i.e. the linear operator:

$$
\begin{aligned}
\mathscr{J}_2 : \quad & H^1(D) \times L^2(D) \quad \rightarrow \quad L^2(D) \times H^1(D) \\
& \mathbf{u} = (u^q, u^p)^T \quad \mapsto \quad \mathscr{J}_2(\mathbf{u}) := \begin{bmatrix} 0 & I_d \\ -I_d & 0 \end{bmatrix} \begin{bmatrix} u^q \\ u^p \end{bmatrix} = \begin{bmatrix} u^p \\ -u^q \end{bmatrix}
\end{aligned}
\tag{5.14}
$$

where I_d is the identity operator defined in $L^2(D)$ or restricted to $H^1(D)$. Then the bilinear form associated to $\mathscr{J}_2$ is clearly antisymmetric and non degenerate in $H^1(D) \times L^2(D)$, thanks to the fact that $H^1(D)$ is dense in $L^2(D)$. This allows us to conclude that $H^1(D) \times L^2(D)$ is a symplectic (pre-Hilbert) vector space when endowed with the canonical form associated to the $[L^2(D)]^2$ inner product. Conversely, in this case we loose the identification in complex setting, namely Proposition 5.3.2 does not apply to $H^1(D) \times L^2(D)$ since we are dealing with the cartesian product of two different spaces. We denote by ϑ_D the symplectic form of $H^1(D) \times L^2(D)$ associated to the $L^2(D) \times L^2(D)$ inner product, i.e:

$$\vartheta_D(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u}, \mathscr{J}_2 \mathbf{v} \rangle_{[L^2(D)]^2}, \qquad \mathbf{u}, \mathbf{v} \in H^1(D) \times L^2(D). \tag{5.15}$$

Hereafter, when confusion does not arise, we omit the subscript and we write $< \cdot, \cdot >$ to indicate the $L^2(D) \times L^2(D)$-product in $H^1(D) \times L^2(D)$ (or any other Sobolev space $\mathcal{H} \subset [L^2(D)]^2$). The same considerations apply to $(H^1(D) \times L^2(D)) \otimes L^2(\Omega)$ and $[H^1(D)]^S \times [L^2(D)]^S$, for any $S > 0$. In analogy with (5.2), one can define the Stiefel manifold $Sp(2S, H^1(D) \times L^2(D))$ of all possible $2S$ dimensional symplectic bases in $H^1(D) \times L^2(D)$ with respect to the symplectic form ϑ_D.

Definition 5.3.10. *We denote with* $Sp(2S, H^1(D) \times L^2(D))$ *the Stiefel manifold of all S dimen-*

sional symplectic bases of $(H^1(D) \times L^2(D), \vartheta_D)$, *i.e.:*

$$Sp(2S, H^1(D) \times L^2(D)) := \{\mathbf{U} \quad = (\mathbf{U}_1, ..., \mathbf{U}_{2S}) \in [H^1(D) \times L^2(D)]^{2S}, \text{ such that}$$
$$\vartheta_D(\mathbf{U}_i, \mathbf{U}_j) = (\mathbf{J}_{2S})_{ij}, \forall i, j = 1, ..., 2S\}. \tag{5.16}$$

We denote by $\mathscr{U}^{sym} \subset H^1(D) \times L^2(D)$ the subspace spanned by $\mathbf{U}$, for any $\mathbf{U} \in Sp(2S, H^1(D) \times L^2(D))$, and we call $\mathbf{U}$ a symplectic basis of $\mathscr{U}^{sym}$. Note that the symplectic form ϑ_D, when restricted to $\mathscr{U}^{sym}$, can be identified with the canonical form of $\mathbb{R}^S$, that is for any $\mathbf{R}, \mathbf{G} \in \mathbb{R}^S$ and $\mathbf{u} = \mathbf{UR}$, $\mathbf{v} = \mathbf{UG} \in \mathscr{U}^{sym}$: $\vartheta_D(\mathbf{u}, \mathbf{v}) = \sum_{i,j=1}^{2S} \mathbf{R}_i \langle \mathbf{U}_i, \mathscr{J}_2 \mathbf{U}_j \rangle \mathbf{G}_j = \mathbf{R}^T \mathbf{J}_{2S} \mathbf{G}$. This implies that ϑ_D is non degenerate in $\mathscr{U}^{sym}$ and $\mathscr{U}^{sym}$, is a symplectic submanifold of $H^1(D) \times L^2(D)$. We define in $Sp(2S, H^1(D) \times L^2(D))$ the following equivalence relation:

$$\mathbf{W} \sim \mathbf{U} \iff \mathscr{W}^{sym} = \mathscr{U}^{sym}$$

meaning that two equivalent elements span the same symplectic subspace.

Lemma 5.3.1. *Two symplectic bases* $\mathbf{W}, \mathbf{U} \in Sp(2S, H^1(D) \times L^2(D))$ *are equivalent if and only if there exists a symplectic matrix* $\mathbf{B} \in Sp(2S, \mathbb{R}^{2S})$ *such that* $\mathbf{W} = \mathbf{UB}$.

Proof. The sufficient condition is obvious: $\mathbf{B}^T \mathbf{J}_{2S} \mathbf{B} = \mathbf{J}_{2S}$ implies $\langle (\mathbf{W})_i, (\mathscr{J}_2 \mathbf{W})_j \rangle_{L^2(D)} = (\mathbf{J}_{2S})_{ij}$. On the hand, if $\mathbf{U} \in Sp(2S, H^1(D) \times L^2(D))$, then $\mathbf{U}_1, ..., \mathbf{U}_{2S}$ are linearly independent. Hence, if $\mathbf{W}, \mathbf{U} \in Sp(2S, H^1(D) \times L^2(D))$ span the same subspace, there necessarily exists a (unique) full rank matrix $\mathbf{B} \in \mathbb{R}^{2S \times 2S}$ such that $\mathbf{W} = \mathbf{UB}$. Then $\langle (\mathbf{W})_i, (\mathscr{J}_2 \mathbf{W})_j \rangle_{L^2(D)} = (\mathbf{J}_{2S})_{ij}$ implies that $\mathbf{B}$ belongs to $Sp(2S, \mathbb{R}^{2S})$. $\square$

5.4 Hamiltonian formulation of wave equations with random parameters

From a physics point of view, a Hamiltonian, denoted in the following by H, is a smooth function which expresses the total energy of a dynamical system in terms of the position and the momentum of its particles. In more abstract setting we can state the following [87]:

Definition 5.4.1. *Let* $(\mathscr{V}, \vartheta)$ *be a symplectic manifold. A vector field* X_H *on* $\mathscr{V}$ *is called Hamiltonian if there is a function* $H : \mathscr{V} \to \mathbb{R}$ *such that:*

$$\vartheta_{\mathbf{u}}(X_H(\mathbf{u}), \mathbf{v}) = \mathbf{d}H(\mathbf{u}) \cdot \mathbf{v}$$

where $\mathbf{d}H(\mathbf{u}) \cdot \mathbf{v}$ *is the directional derivative of* H *along* $\mathbf{v}$. *Hamilton's equations are the evolution equations:*

$$\dot{\mathbf{u}} = X_H(\mathbf{u}) \tag{5.17}$$

If $(\mathscr{V}_{2N}, \vartheta)$ is a symplectic vector space and $(\mathbf{q}, \mathbf{p}) = (q_1, ..., q_N, p_1, ..., p_N)$ denote the canonical

coordinates with respect to which ϑ has matrix $\mathbf{J}_{2N}$, the Hamiltonian equations become:

$$\dot{\mathbf{u}} = \mathbf{J}_{2N} \nabla H(\mathbf{u}).$$

Let ϕ_t denote the flow of the Hamiltonian X_H, that is $\phi_t(\mathbf{u}_0)$ is the solution to (5.17) with initial condition $\mathbf{u}_0 \in \mathcal{V}$, we have that ϕ_t conserves the energy of H.

Proposition 5.4.1. *Let ϕ_t be the flow of X_H on the symplectic manifold $(\mathcal{V}, \vartheta)$. Then $H \circ \phi_t = H$, where defined.*

Proof.

$$\begin{aligned}
\tfrac{d}{dt}\big(H \circ \phi_t(\mathbf{u})\big) &= \mathbf{d}H(\phi_t(\mathbf{u})) \cdot X_H(\phi_t(\mathbf{u})) \\
&= \vartheta_{\phi_t(\mathbf{u})}(X_H(\phi_t(\mathbf{u})), X_H(\phi_t(\mathbf{u}))) = 0
\end{aligned}$$

$\square$

The flow ϕ_t of a Hamiltonian vector field consists of symplectic transformations, namely ϕ_t (whenever it is defined) preserves the symplectic form ϑ. Formally, for all $\mathbf{u} \in \mathcal{V}$ and $\mathbf{v}, \mathbf{z} \in \mathcal{T}_{\mathbf{u}}\mathcal{V}$, we have:

$$\vartheta_{\mathbf{u}}(\mathbf{v}, \mathbf{z}) = \vartheta_{\phi_t(\mathbf{u})}(D_{\mathbf{u}}[\phi_t](\mathbf{v}), D_{\mathbf{u}}[\phi_t](\mathbf{z})) \tag{5.18}$$

where $D_{\mathbf{u}}[\phi_t]$ is the differential of ϕ_t at $\mathbf{u}$. It follows from Poincaré lemma [88, 56] that the flow ϕ_t of a vector field X is symplectic if and only if it is locally Hamiltonian, that is there locally exists a Hamiltonian function H such that $\vartheta_{\mathbf{u}}(X(\mathbf{u}), \mathbf{v}) = \mathbf{d}H(\mathbf{u}) \cdot \mathbf{v}$. The link between symplecticity and energy preservation has been widely studied and exploited to derive numerical time discretization schemes that share the same symplectic structure of the original system, in order to preserve the geometric properties. The same idea can be used to formulate reduced order methods which preserve the symplectic structure underlying the original full order Hamiltonian system, thus being energy conservative and preserving stability.

We start by looking for a suitable Hamiltonian formulation for wave equations with random parameters. As shown in literature [118, 94], problem (5.3) admits a unique solution $u \in L^{\infty}([0, T], H_0^1(D) \otimes L^2(\Omega))$ with time derivative $\dot{u} \in L^{\infty}([0, T], L^2(D) \otimes L^2(\Omega))$, provided that the random wave speed c is bounded and uniformly coercive and the initial data (q_0, p_0) belong to $\big(H_0^1(D) \otimes L^2(\Omega)\big) \times \big(L^2(D) \otimes L^2(\Omega)\big)$. Let us introduce the phase space variables $(p, q) = (u, \dot{u})$, then problem (5.3) can be rewritten into a first order system in $H^1(D) \times L^2(D)$ for almost all

$\omega \in \Omega$:

$$\begin{cases} \dot{q}(\mathbf{x},t,\omega) = p(\mathbf{x},t,\omega) & \mathbf{x} \in D, \ t \in (0,T], \ \omega \in \Omega, \\ \dot{p}(\mathbf{x},t,\omega) = \nabla \cdot \big(c(\mathbf{x},\omega)\nabla q(\mathbf{x},t,\omega)\big) - f(q(\mathbf{x},t,\omega),\omega) & \mathbf{x} \in D, \ t \in (0,T], \ \omega \in \Omega, \\ q(\mathbf{x},0,\omega) = q_0(\mathbf{x},\omega) & \mathbf{x} \in D, \ \omega \in \Omega, \\ p(\mathbf{x},0,\omega) = p_0(\mathbf{x},\omega) & \mathbf{x} \in D, \ \omega \in \Omega, \\ q(\mathbf{x},t,\omega) = 0 & \mathbf{x} \in \partial D, \ t \in (0,T], \ \omega \in \Omega, \end{cases} \tag{5.19}$$

analogously written in matrix form as:

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \mathscr{J}_2 \begin{pmatrix} -\nabla \cdot \big(c\nabla \cdot q\big) + f(q) \\ p \end{pmatrix}$$

Problem (5.19) can be interpreted as a Hamiltonian system in the symplectic space $(H_0^1(D) \times L^2(D), \vartheta_D)$ with symplectic form ϑ_D defined in (5.15). In this case, the Hamiltonian energy associated to (5.19) is defined pointwise in ω as:

$$H_\omega(q,p) = \frac{1}{2}\int_D \Big(|p|^2 + c(\omega)|\nabla q|^2 + F(q)\Big), \qquad F'(q) = f(q).$$

Thus, by denoting with $\nabla_q H_\omega$, $\nabla_p H_\omega$ the functional derivatives of H_ω with respect to q and p respectively, i.e.:

$$\begin{aligned} \langle \nabla_q H_\omega, \delta q \rangle \ &= \int_D c\nabla q \nabla \delta q + \int_D f(q)\delta q \qquad \text{and} \qquad \langle \nabla_p H_\omega, \delta p \rangle = \int_D p\delta p. \\ &= \int_D \big(-\nabla \cdot (c\nabla \cdot q) + f(q)\big)\delta q, \end{aligned}$$

for any $\delta q \in H_0^1(D)$, $\delta p \in L^2(D)$, where the term $\int_D -\nabla \cdot (c\nabla \cdot q)\delta q$ should be interpreted in distributional sense, equation (5.3) is recast into the following canonical Hamiltonian system, written with respect to $\mathbf{u} = (q,p)$:

$$\begin{cases} \dot{\mathbf{u}}(\mathbf{x},t,\omega) = \mathscr{J}_2 \nabla H_\omega\big(\mathbf{u}(\mathbf{x},t,\omega),\omega\big), \\ \mathbf{u}(\mathbf{x},0,\omega) = (q_0(\mathbf{x},\omega), p_0(\mathbf{x},\omega))^T \end{cases} \tag{5.20}$$

for almost every $\mathbf{x} \in D$ and $\omega \in \Omega$. Observe that both the flow of the solutions and the Hamiltonian depend on the random input, and that the conservation of energy applies point-wise in the parameter space, which means that, for any realization ω, the flow ϕ_t of (5.20) with initial conditions evaluated in ω, conserves the Hamiltonian evaluated in ω. This immediately implies that the expected value, and generally any finite moment of H_ω, are constant along the flow of the solutions.

Alternatively, in a setting more suited to our context, the conservation of energy can be derived directly in $\big(H^1(D) \otimes L^2(\Omega)\big) \times \big(L^2(D) \otimes L^2(\Omega)\big) = \big(H^1(D) \times L^2(D)\big) \otimes L^2(\Omega)$ equipped with the

following symplectic form:

$$
\begin{aligned}
\vartheta(\mathbf{u},\mathbf{v}) \quad &= \mathbb{E}[\langle \mathbf{u}, \mathscr{J}_2\mathbf{v}\rangle_{[L^2(D)]^2}], \\
&= \mathbb{E}[\langle u^q, v^p\rangle_{L^2(D)}] - \mathbb{E}[\langle u^p, v^q\rangle_{L^2(D)}]
\end{aligned}
\tag{5.21}
$$

for any $\mathbf{u} = (u^q, u^p)$, $\mathbf{v} = (v^q, v^p) \in \big(H^1(D) \otimes L^2(\Omega)\big) \times \big(L^2(D) \otimes L^2(\Omega)\big)$, with $u^q, v^q \in H^1(D) \otimes L^2(\Omega)$, $u^p, v^p \in L^2(D) \otimes L^2(\Omega)$. The pair $\big((H^1(D) \otimes L^2(\Omega)) \times (L^2(D) \otimes L^2(\Omega))\big)$ is the symplectic space that will be used in Section 5.6 to derive the Symplectic Dynamical Low Rank method. In this setting the Hamiltonian energy associated to (5.19) is defined as:

$$
H(q,p) = \frac{1}{2}\mathbb{E}\Big[\int_D \big(|p|^2 + c(\omega)|\nabla q|^2 + F(q)\big)\Big], \qquad F'(q) = f(q).
$$

In particular, if $X_{H(\omega)}$ denotes the Hamiltonian vector field associated to (5.20), for $\mathbf{u}$ sufficiently smooth, system (5.20) can be rewritten as $\dot{\mathbf{u}} = X_{H(\omega)}(\mathbf{u})$ and the conservation of mean energy along the flow of the solutions can be rederived in terms of the symplectic form (5.21) as:

$$
\begin{aligned}
\frac{d}{dt} H(\mathbf{u}(t)) \quad &= \langle \nabla H(\mathbf{u}(t)), \dot{\mathbf{u}}(t)\rangle \\
&= \langle \nabla H(\mathbf{u}(t)), X_{H(\omega)}(\mathbf{u}(t))\rangle \\
&= -\vartheta\big(X_{H(\omega)}(\mathbf{u}(t)), X_{H(\omega)}(\mathbf{u}(t))\big) = 0
\end{aligned}
\tag{5.22}
$$

thanks to the antisymmetry of ϑ.

5.5 Symplectic Order Reduction

We recall here the symplectic order reduction for parametric Hamiltonian systems proposed in [107]. This method is designed in analogy to the proper orthogonal decomposition where the standard inner product is replaced by the symplectic form and leads to approximate solutions which belong to a low dimensional symplectic space. This method has the desirable property of preserving the symplectic structure of the full order system, which allows deriving conservative schemes.

Definition 5.5.1. *Let* $\mathbf{U} \in Sp(2S, H^1(D) \times L^2(D))$, *the symplectic inverse of* $\mathbf{U}$, *denoted by* $\mathbf{U}^+$, *is the* $2S$ *vector function written as:*

$$
\mathbf{U}^+ := \mathscr{J}_2^T \mathbf{U} \mathbf{J}_{2S} \in Sp(2S, L^2(D) \times H^1(D)).
\tag{5.23}
$$

If we write $\mathbf{U}$ component-wise, with $\mathbf{U}_i = (U_i^q, U_i^p)^T \in H^1(D) \times L^2(D)$:

$$
\mathbf{U} = \begin{bmatrix} U_1^q & \cdots & U_S^q & U_{S+1}^q & \cdots & U_{2S}^q \\ U_1^p & \cdots & U_S^p & U_{S+1}^p & \cdots & U_{2S}^p \end{bmatrix}
\qquad
\begin{aligned}
\mathbf{Q}_{I,S} &= U_1^q, ..., U_S^q, & \mathbf{Q}_{II,S} &= U_{S+1}^q, ..., U_{2S}^q, \\
\end{aligned}
$$
$$
= \begin{bmatrix} \mathbf{Q}_{I,S} & \mathbf{Q}_{II,S} \\ \mathbf{P}_{I,S} & \mathbf{P}_{II,S} \end{bmatrix},
\qquad
\begin{aligned}
\mathbf{P}_{I,S} &= U_1^p, ..., U_S^p, & \mathbf{P}_{II,S} &= U_{S+1}^p, ..., U_{2S}^p,
\end{aligned}
$$

then $\mathbf{U}^+$ is explicitly given by:

$$\mathbf{U}^+ = \begin{bmatrix} \mathbf{P}_{II,S} & -\mathbf{P}_{I,S} \\ -\mathbf{Q}_{II,S} & \mathbf{Q}_{I,S} \end{bmatrix}$$

It is straightforward to verify that $\langle \mathbf{U}_i, \mathbf{U}_j^+ \rangle = \delta_{ij}$, $\forall i, j = 1,...,2S$. The notion of symplectic inverse is used to define the symplectic Galerkin projection. Precisely:

Definition 5.5.2. *Let* $\mathbf{v} = (v^q, v^p)^{\mathrm{T}}$ *be a square integrable random field in* $\left(H^1(D) \times L^2(D) \right) \otimes L^2(\Omega)$ *and* $\mathbf{U} \in Sp(2S, H^1(D) \times L^2(D))$ *a symplectic basis, spanning* $\mathscr{U}^{sym}$. *The symplectic projection of* $\mathbf{v}$ *into* $\mathscr{U}^{sym} \otimes L^2(\Omega)$ *is defined as:*

$$\mathbf{v}_S(\mathbf{x}, \omega) = \mathscr{P}_{\mathbf{U}}^{sym}[\mathbf{v}] = \mathscr{P}_{\mathbf{U}^+}[\mathbf{v}] = \mathbf{U}(\mathbf{x})\mathbf{Y}(\omega) \tag{5.24}$$

where

- $\mathbf{U}^+ \in Sp(2S, L^2(D) \times H^1(D))$ *is the symplectic inverse of* $\mathbf{U}$*, defined in (5.23);*

- $\mathbf{Y} = Y_1, ..., Y_{2S}$ *is a vector of* $2S$ *square integrable random variables defined as* $Y_i = \langle \mathbf{v}, \mathbf{U}_i^+ \rangle$.

Moreover we say that $\mathbf{v}$ *is in the subspace spanned by* $\mathbf{U}$ *if* $\mathbf{v} = \mathscr{P}_{\mathbf{U}}^{sym}[\mathbf{v}]$*, or namely if there exists a vector of (square integrable) random variables* $\mathbf{Y}$*, such that* $\mathbf{v} = \mathbf{U}\mathbf{Y}$*. Observe that* $\mathbf{Y}$ *is uniquely determined by* $\mathbf{U}$ *by means of symplectic projection as* $Y_i = \langle \mathbf{v}, \mathbf{U}_i^+ \rangle$*. On the contrary, we say that* $\mathbf{v} \in \left(H^1(D) \times L^2(D) \right) \otimes L^2(\Omega)$ *is in the symplectic orthogonal complement of* $\mathbf{U}$ *if* $\mathscr{P}_{\mathbf{U}}^{sym}[\mathbf{v}] = 0$*. We denote by* $\mathscr{P}_{\mathbf{U}}^{sym,\perp}[\cdot] = \mathbb{I} - \mathscr{P}_{\mathbf{U}}^{sym}[\cdot]$ *the projection into the symplectic orthogonal complement of* $\mathbf{U}$*.*

The Symplectic Order Reduction method consists of two steps:

- an off-line stage for computing the basis functions $\mathbf{U} = (\mathbf{U}_1, ..., \mathbf{U}_{2S})$. They can be extracted by means of Principal Symplectic Decomposition (PSD) procedures from snapshots $\mathbf{u}(\cdot, t_j, \omega_k)$ collected at different times and for different values of the parameters [107], or following a greedy-PSD approach as described in [86].

- an on-line stage which consists in low-cost reduced-order simulations for computing the coefficients $\mathbf{Y} = (Y_1, ..., Y_{2S})$ at each time and for different values of the parameters. The reduced order system is obtained by preforming a symplectic Galerkin projection of the governing Hamiltonian equations in the subspace spanned by $\mathbf{U}$.

The use of the symplectic Galerkin projection aims to preserve the symplectic structure of the original problem, in order to ensure the stability of the reduced order system [73]. More precisely, the approximate solution $\mathbf{u}_S = (q_S, p_S)^T$ to problem (5.20), which is written as:

$$\mathbf{u}_S(\mathbf{x}, t, \omega) = \mathbf{U}(\mathbf{x})\mathbf{Y}(\omega, t),$$

satisfies the following variational principle at each time and for any $\omega \in \Omega$:

$$\left\langle \dot{\mathbf{u}}_S - \mathcal{J}_2 \nabla H_\omega(\mathbf{u}_S, \omega), \mathcal{J}_2^T \mathbf{v} \right\rangle = 0, \quad \forall \mathbf{v} \in \mathcal{U}^{sym}, \tag{5.25}$$

where $\mathcal{U}^{sym}$ is the subspace spanned by $\mathbf{U} \in Sp(2S, H^1(D) \times L^2(D))$. This can be written formally as a symplectic projection of the governing equation (5.20) into $\mathcal{U}^{sym}$:

$$\dot{\mathbf{u}}_S(t) = \mathcal{P}_{\mathbf{U}}^{sym}[\mathcal{J}_2 \nabla H_\omega(\mathbf{u}_S(t), \omega)], \qquad \forall t, \omega \in (0, T] \times \Omega$$

where the definition of $\mathcal{P}_{\mathbf{U}}^{sym}[\cdot]$ is properly extended to all $\mathbf{v} \in \left(L^2(D) \times H^{-1}(D)\right) \otimes L^2(\Omega)$ as $\mathcal{P}_{\mathbf{U}}^{sym}[\mathbf{v}] = \sum\limits_{i=1}^{2S} \mathbf{U}_i \langle \mathbf{v}, \mathbf{U}_i^+ \rangle$ and $\langle \cdot, \cdot \rangle$ denoting the H_0^1-H^{-1} duality pair. If we write the solution component-wise, the position and momentum are respectively approximated as:

$$q(\mathbf{x}, t, \omega) \approx q_S(\mathbf{x}, t, \omega) = \sum_{i=1}^{2S} U_i^q(\mathbf{x}) Y_i(\omega, t), \qquad p(\mathbf{x}, t, \omega) \approx p_S(\mathbf{x}, t, \omega) = \sum_{i=1}^{2S} U_i^p(\mathbf{x}) Y_i(\omega, t)$$

The stochastic coefficients $\mathbf{Y} = (Y_1, ..., Y_{2S})$ belong to $[L^2(\Omega)]^{2S}$ and satisfy the following system of ordinary differential equations (ODEs):

$$\begin{aligned} \dot{\mathbf{Y}}(\omega) &= \langle \mathcal{P}_{\mathbf{U}}^{sym}[\mathcal{J}_2 \nabla H_\omega(\mathbf{u}_S, \omega)], \mathbf{U}^+ \rangle \\ &= \langle \mathcal{J}_2 \nabla H_\omega(\mathbf{u}_S, \omega), \mathbf{U}^+ \rangle \\ &= \langle \nabla H_\omega(\mathbf{u}_S, \omega), \mathbf{U} \mathbf{J}_{2S}^T \rangle = \mathbf{J}_{2S} \nabla_{\mathbf{Y}} \tilde{H}_\omega(\mathbf{Y}, \omega) \end{aligned} \tag{5.26}$$

with initial conditions $Y_i(0) = \langle (q_0, p_0)^T, \mathbf{U}_i^+ \rangle$ for all $i = 1, ..., 2S$, obtained by performing a symplectic projection of the initial datum (5.19) on $\mathbf{U}$.

Remark 5.5.1. *The reduced system (5.26) consists of Hamiltonian equations in the symplectic Hilbert space $[L^2(\Omega)]^{2S}$ equipped with the canonical form: $\mathbb{E}[\mathbf{Y}^T \mathbf{J}_{2S} \mathbf{Z}], \forall \mathbf{Y}, \mathbf{Z} \in [L^2(\Omega)]^{2S}$.*

Lemma 5.5.1 (from [107])**.** *Let $\mathbf{U}$ belong to $Sp(2S, H^1(D) \times L^2(D))$ and $\phi_{\mathbf{U}}$ be the linear map associated to $\mathbf{U}$, defined as:*

$$\begin{aligned} \phi_{\mathbf{U}} \quad &: [L^2(\Omega)]^{2S} \quad \rightarrow \quad [H^1(D) \times L^2(D)] \otimes L^2(\Omega) \\ &\qquad \mathbf{Y} \qquad \mapsto \quad \phi_{\mathbf{U}}(\mathbf{Y}) := \mathbf{U}\mathbf{Y}. \end{aligned}$$

Then $\phi_{\mathbf{U}}$ is a symplectic linear map between $\left([L^2(\Omega)]^{2S}, \mathbf{J}_{2S}\right)$ and $\left([H^1(D) \times L^2(D)] \otimes L^2(\Omega), \mathcal{J}_2\right)$, i.e. $\phi_{\mathbf{U}}$ preserves the symplectic form:

$$\mathbb{E}[\mathbf{Y}^T \mathbf{J}_{2S} \mathbf{Z}] = \mathbb{E}[\langle \mathbf{U}\mathbf{Z}, \mathcal{J}_2 \mathbf{U}\mathbf{Y} \rangle]$$

for any $\mathbf{Y}, \mathbf{Z} \in [L^2(\Omega)]^{2S}$. Moreover, the function $\tilde{H}_\omega$ in (5.26), which can be written as:

$$\begin{aligned} \tilde{H}_\omega := H_\omega \circ \phi_{\mathbf{U}} \quad &: [L^2(\Omega)]^{2S} \rightarrow L^2(\Omega) \\ &\mathbf{Y} \rightarrow H_\omega\left(\sum_{i=1}^{2S} \mathbf{U}_i Y_i, \omega \right), \end{aligned}$$

is a first integral, pointwise in ω, of $\mathbf{Y}(t)$, *that is the flow of* (5.26) *preserves the energy of* $\tilde{H}_\omega$ *at each time and for any* ω.

In conclusion, the original problem (5.19), set in $(H^1(D) \times L^2(D)) \otimes L^2(\Omega)$, is reduced to a Hamiltonian ODE system of dimension $2S$, set in $[L^2(\Omega)]^{2S}$, describing the evolution of the random coefficients $Y_1, ..., Y_{2S}$. To verify that $\tilde{H}_\omega$ is conserved by the solution of (5.26), note that $\frac{d}{dt} \tilde{H}_\omega(\mathbf{Y}(t)) = \sum_{i=1}^{2S} \nabla_{Y_i(t)} \tilde{H}_\omega(\mathbf{Y}(t)) \cdot \dot{Y}_i(t) = (\nabla_{\mathbf{Y}} \tilde{H}_\omega(\mathbf{Y}(t)))^T \mathbf{J}_{2S} \nabla_{\mathbf{Y}} \tilde{H}_\omega(\mathbf{Y}(t)) = 0$ a.s. in Ω. The energy of the approximate solution, that is $\tilde{H}_\omega(\mathbf{Y}(t)) = H_\omega(\mathbf{U}\mathbf{Y}(t)) = H_\omega(\mathbf{u}_S(t))$, is not necessary equal to the exact one, namely the energy of the exact solution $H_\omega(\mathbf{u}(t))$, but the discrepancy between the exact and the approximate energy remains constant in time and can be evaluated at initial time. The drawback of the Symplectic Reduced Order approach with a fixed basis $\mathbf{U}$, is that if the solution manifolds $\mathscr{M}(t) = \{u(\cdot, t, \omega), \omega \in \Omega\}$ significantly change during the time evolution, as it typically happens in wave propagation phenomena, the *fixed* reduced basis $\mathbf{U} = (\mathbf{U}_1, \dots, \mathbf{U}_{2S})$ has to be sufficiently rich to be able to approximate such manifolds for all $t \in [0, T]$. This leads to a fairly large reduced model thus compromising its efficiency.

5.6 Symplectic Dynamical Low Rank approximation

In this paper we propose the Symplectic Dynamical Low Rank (Symplectic DO) approximation for wave equations with random parameters which combines the Dynamically Orthogonal approach described in Section 5.2 with the Symplectic Order Reduction strategy summarized in Section 5.5. This method shares with the symplectic order reduction the use of symplectic deterministic bases, and, as the "classic" DO approximation, allows both the stochastic and the deterministic modes to evolve in time. This aims to both preserve the Hamiltonian structure of the original problem and guarantee more flexibility to the approximation. The approximate solution, indeed, preserves the (approximated) mean Hamiltonian energy and continuously adapts in time to the structure of the solution. The reduced dynamical system consists of a set of equations for the constrained dynamics of the deterministic modes in a submanifold of $Sp(2S, H^1(D) \times L^2(\Omega))$, coupled with a reduced order Hamiltonian system for the evolution of the stochastic coefficients.

Definition 5.6.1. *We denote* $U(S, [H^1(D)]^2)$ *the submanifold of* $Sp(2S, H^1(D) \times L^2(D))$ *consisting of all* L^2-*orthonormal symplectic bases in* $[H^1(D)]^2$, *i.e.:*

$$U(S, [H^1(D)]^2) := \{\mathbf{U} = (\mathbf{U}_1, ..., \mathbf{U}_{2S}) \in [H^1(D) \times H^1(D)]^{2S} \text{ such that}$$
$$\vartheta_D(\mathbf{U}_i, \mathbf{U}_j) = (\mathbf{J}_{2S})_{ij} \text{ and } \langle \mathbf{U}_j, \mathbf{U}_i \rangle_{L^2(D)} = \delta_{ij}, \forall i, j = 1, ..., 2S\},$$

with ϑ_D *defined in* (5.15).

The advantage in restricting $Sp(2S, H^1(D) \times L^2(D))$ to $U(S, [H^1(D)]^2)$ is the possibility to identify the latter with the Stiefel manifold $St(S, H^1(D, \mathbb{C}))$ of all S-dimensional orthonormal complex bases in $H^1(D, \mathbb{C})$ (while the same clearly does not applies to $Sp(2S, H^1(D) \times L^2(D))$). We

149

postpone this discussion to Section 5.6.1, and we go forward here with the construction of the approximation manifold.

Proposition 5.6.1. *The following properties hold for any* $\mathbf{U} \in U(S, [H^1(D)]^2)$:

a) *let* $\mathbf{U} \in Sp(2S, H^1(D) \times L^2(D))$, *then* $\mathbf{U} \in U(S, [H^1(D)]^2)$ *if and only if:*

$$\mathbf{U}^+ = \mathscr{J}_2^T \mathbf{U} \mathbf{J}_{2S} = \mathscr{J}_2 \mathbf{U} \mathbf{J}_{2S}^T = \mathbf{U}; \tag{5.27}$$

b) $\mathbf{U} \in U(S, [H^1(D)]^2)$ *if and only if:*

$$\mathbf{U} = \begin{pmatrix} \mathbf{Q} & -\mathbf{P} \\ \mathbf{P} & \mathbf{Q} \end{pmatrix} \tag{5.28}$$

with $\mathbf{Q}, \mathbf{P} \in [H^1(D)]^S$ *row vector functions such that:*

$$\langle P_i, Q_j \rangle = \langle Q_i, P_j \rangle \quad and \quad \langle Q_i, Q_j \rangle + \langle P_i, P_j \rangle = \delta_{ji}, \tag{5.29}$$

for all $i, j = 1, ..., S$.

Proof. Here we use the notation $\ll \mathbf{U}, \mathbf{V} \gg$ to denote the $2S \times 2S$ matrix with entries $\ll \mathbf{U}, \mathbf{V} \gg_{ij} = \langle \mathbf{U}_j, \mathbf{V}_i \rangle$, for all $\mathbf{U}, \mathbf{V} \in [H^1(D) \times H^1(D)]^{2S}$ (Analogous definition for $\mathbf{U}, \mathbf{V} \in [H^1(D)]^S$).

a) If (5.27) then $\ll \mathbf{U}, \mathbf{U} \gg = \ll \mathscr{J}_2^T \mathbf{U} \mathbf{J}_{2S}, \mathbf{U} \gg = \ll \mathbf{U}^+, \mathbf{U} \gg = \mathbb{I}_{2S}$ implies $\mathbf{U} \in U(S, [H^1(D)]^2)$. If $\mathbf{U} \in U(S, [H^1(D)]^2)$, then $\mathbf{U}^+ \in U(S, [H^1(D)]^2)$. Since both $\mathbf{U}, \mathbf{U}^+$ are vectors of linearly independent functions

$$\ll \mathbf{U}, \mathbf{U} \gg = \mathbb{I}_{2S} = \ll \mathbf{U}^+, \mathbf{U} \gg \quad \Rightarrow \quad \mathbf{U}^+ = \mathscr{J}_2^T \mathbf{U} \mathbf{J}_{2S} = \mathbf{U}.$$

b) If $\mathbf{U} \in U(S, [H^1(D)]^2)$ then $\ll \mathbf{U}, \mathscr{J}_2 \mathbf{U} \gg = \mathbf{J}_{2S}$. Block-wise, this is written as:

$$\mathbf{U} = \begin{bmatrix} \mathbf{Q}_{I,S} & \mathbf{Q}_{II,S} \\ \mathbf{P}_{I,S} & \mathbf{P}_{II,S} \end{bmatrix} \tag{5.30}$$

with $\mathbf{Q}_{I,S}, \mathbf{Q}_{II,S}, \mathbf{P}_{I,S}, \mathbf{P}_{II,S} \in [H^1(D)]^S$ such that:

$$\ll \mathbf{P}_{II,S}, \mathbf{Q}_{I,S} \gg - \ll \mathbf{Q}_{II,S}, \mathbf{P}_{I,S} \gg = \mathbb{I}_S, \tag{5.31}$$
$$\ll \mathbf{P}_{I,S}, \mathbf{Q}_{I,S} \gg = \ll \mathbf{Q}_{I,S}, \mathbf{P}_{I,S} \gg \quad and \quad \ll \mathbf{P}_{II,S}, \mathbf{Q}_{II,S} \gg = \ll \mathbf{Q}_{II,S}, \mathbf{P}_{II,S} \gg$$

From a) we have that $\mathbf{U}^+ = \mathbf{U}$, i.e.:

$$\mathbf{U}^+ = \begin{bmatrix} \mathbf{P}_{II,S} & -\mathbf{P}_{I,S} \\ -\mathbf{Q}_{II,S} & \mathbf{Q}_{I,S} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{I,S} & \mathbf{Q}_{II,S} \\ \mathbf{P}_{I,S} & \mathbf{P}_{II,S} \end{bmatrix} = \mathbf{U}$$

which implies $\mathbf{Q} := \mathbf{Q}_{I,S} = \mathbf{P}_{II,S}$ and $\mathbf{P} := \mathbf{P}_{I,S} = -\mathbf{Q}_{II,S}$. The proof is concluded by combining the last relation with (5.31). The other implication is obvious.

$\square$

From Proposition 5.6.1 it follows that the symplectic projection coincides with the standard projection:

Proposition 5.6.2. *For any* $\mathbf{v} = (v^q, v^p)^{\mathrm{T}} \in \left(H^1(D) \times L^2(D) \right) \otimes L^2(\Omega)$ *and* $\mathbf{U} \in U(S, [H^1(D)]^2)$ *it holds that:*

$$\mathscr{P}_{\mathbf{U}}^{sym}[\mathbf{v}] = \mathscr{P}_{\mathbf{U}^+}[\mathbf{v}] = \mathscr{P}_{\mathbf{U}}[\mathbf{v}] \tag{5.32}$$

Additionally the following properties hold:

Proposition 5.6.3. *Let* $\mathbf{v}$ *be a square integrable random field* $\mathbf{v} = (v^q, v^p)^{\mathrm{T}} \in (H^1(D) \times L^2(D)) \otimes L^2(\Omega)$. *For any* $\mathbf{U} \in U(S, [H^1(D)]^2)$ *we have that:*

$$\mathscr{P}_{\mathbf{U}}^{sym}[\mathbf{v}] = \mathscr{P}_{\mathbf{U}}[\mathbf{v}] = \mathscr{P}_{\mathscr{J}_2\mathbf{U}}[\mathbf{v}] = \mathscr{P}_{\mathscr{J}_2\mathbf{U}}^{sym}[\mathbf{v}]; \tag{5.33}$$

where $\mathscr{P}_{\mathbf{U}}, \mathscr{P}_{\mathscr{J}_2\mathbf{U}}$ (respectively $\mathscr{P}_{\mathbf{U}}^{sym}, \mathscr{P}_{\mathscr{J}_2\mathbf{U}}^{sym}$) are the standard (respectively symplectic) projections in the subspace spanned by $\mathbf{U}$ and $\mathscr{J}_2\mathbf{U}$ respectively.

Proposition 5.6.4. *Let* $\mathbf{v}$ *be a square integrable random field* $\mathbf{v} = (v^q, v^p)^{\mathrm{T}} \in (H^1(D) \times L^2(D)) \otimes L^2(\Omega)$. *For any* $\mathbf{U} \in U(S, [H^1(D)]^2)$ *we have that:*

$$\mathscr{J}_2\mathscr{P}_{\mathbf{U}}^{sym}[\mathbf{v}] = \mathscr{J}_2\mathscr{P}_{\mathbf{U}}[\mathbf{v}] = \mathscr{P}_{\mathbf{U}}[\mathscr{J}_2\mathbf{v}] = \mathscr{P}_{\mathbf{U}}^{sym}[\mathscr{J}_2\mathbf{v}]. \tag{5.34}$$

The same property is satisfied by the projector into the symplectic -orthogonal complement of $\mathbf{U}$:

$$\mathscr{J}_2\mathscr{P}_{\mathbf{U}}^{sym,\perp}[\mathbf{v}] = \mathscr{J}_2\mathscr{P}_{\mathbf{U}}^{\perp}[\mathbf{v}] = \mathscr{P}_{\mathbf{U}}^{\perp}[\mathscr{J}_2\mathbf{v}] = \mathscr{P}_{\mathbf{U}}^{sym,\perp}[\mathscr{J}_2\mathbf{v}].$$

The Symplectic DO approximate solution of problem (5.19) is sought in the approximation manifold defined as follows:

Definition 5.6.2. *We call symplectic manifold of rank S, denoted by* $\mathscr{M}_S^{sym}$, *the collection of all random fields* $\mathbf{u}_S = (q_S, p_S)^T \in [H^1(D)]^2 \otimes L^2(\Omega)$ *that can be written as:* $\mathbf{u}_S = \mathbf{U}\mathbf{Y}$ *where*

- $\mathbf{U} \in U(S, [H^1(D)]^2)$,

- $\mathbf{Y} = Y_1, ..., Y_{2S}$ *is a* $2S$ *dimensional vector of square integrable random variables* $Y_i \in L^2(\Omega)$, *such that* $rank(\mathbb{E}[\mathbf{Y}\mathbf{Y}^T] + \mathbf{J}_{2S}^T\mathbb{E}[\mathbf{Y}\mathbf{Y}^T]\mathbf{J}_{2S}) = 2S$.

We call symplectic S rank random field any function $\mathbf{u}_S \in \mathcal{M}_S^{sym}$. *This can be written component-wise as follows:*

$$q_S(\mathbf{x}, \omega) = \sum_{i=1}^{S} Q_i(\mathbf{x}) Y_i(\omega) - \sum_{i=1}^{S} P_i(\mathbf{x}) Y_{S+i}(\omega), \quad p_S(\mathbf{x}, \omega) = \sum_{i=1}^{S} P_i(\mathbf{x}) Y_i(\omega) + \sum_{i=1}^{S} Q_i(\mathbf{x}) Y_{S+i}(\omega) \quad (5.35)$$

In the following we denote by $B^{sym}(2S, L^2(\Omega)) \subset [L^2(\Omega)]^{2S}$ the set of all $2S$-vectors $\mathbf{Z} = (Z_1, ..., Z_{2S}) \in [L^2(\Omega)]^{2S}$, that satisfy the full rank condition on $\mathbb{E}[\mathbf{Z}\mathbf{Z}^T] + \mathbf{J}_{2S}^T \mathbb{E}[\mathbf{Z}\mathbf{Z}^T] \mathbf{J}_{2S}$. Observe (5.28) implies that the first S components of $\mathbf{U} \in U(S, [H^1(D)]^2)$ characterize the whole vectors $\mathbf{U}$ and $\mathbf{U}^+$, for all $\mathbf{U} \in U(S, [H^1(D)]^2)$ (which motivates the name of symplectic S rank random field). Hence, for (5.28) to be verified, the same regularity has to be assumed for both the position and momentum components. This means that, when we look for an approximate solution of problem (5.19) in $\mathcal{M}_S^{sym}$, we necessary have to assume some extra-regularity on the approximate momentum. In other words, the orthonormality combined to the symplectic condition forces to set the approximation problem in $[H^1(D)]^2 \otimes L^2(\Omega)$ while the natural setting would be $(H^1(D) \times L^2(D)) \otimes L^2(\Omega)$.

Remark 5.6.1. *The representation of* $\mathbf{u}_S \in \mathcal{M}_S^{sym}$ *in terms of* $\mathbf{U} \in U(S, [H^1(D)]^2)$ *and* $\mathbf{Y} \in B^{sym}(2S, L^2(\Omega))$ *(decomposition (5.35)) is not unique. Let* $\mathbf{u}_S = \mathbf{U}\mathbf{Y} \in \mathcal{M}_S^{sym}$ *with* $\mathbf{U} \in U(S, [H^1(D)]^2)$, $\mathbf{Y} \in B^{sym}(2S, L^2(\Omega))$, *then for any* $\mathbf{B} \in U(S, \mathbb{R}^{2S})$ *we have that* $\mathbf{W} = \mathbf{U}\mathbf{B} \in U(S, [H^1(D)]^2)$, $\mathbf{Z} = (\mathbf{B}^+)^T \mathbf{Y} = \mathbf{B}^T \mathbf{Y} \in B^{sym}(2S, L^2(\Omega))$ *leads to* $\mathbf{W}\mathbf{Z} = \mathbf{u}_S$. *Indeed:*

- $\mathbf{W} \in U(S, [H^1(D)]^2)$:

$$\begin{aligned} \langle \mathbf{W}_j, (\mathcal{J}_2 \mathbf{W})_i \rangle &= \langle (\mathbf{U}\mathbf{B})_j, (\mathcal{J}_2 \mathbf{U}\mathbf{B})_i \rangle \\ &= \mathbf{B}_{kj} \langle \mathbf{U}_k, (\mathcal{J}_2 \mathbf{U})_l \rangle \mathbf{B}_{li} \\ &= \mathbf{B}_{jk}^T (\mathbf{J}_{2S})_{kl} \mathbf{B}_{li} = (\mathbf{J}_{2S})_{ij} \end{aligned}$$

$$\begin{aligned} \langle \mathbf{W}_j, \mathbf{W}_i \rangle &= \langle \mathbf{U}_l \mathbf{B}_{li}, \mathbf{U}_s \mathbf{B}_{sj} \rangle \\ &= \mathbf{B}_{il}^T \langle \mathbf{U}_l, \mathbf{U}_s \rangle \mathbf{B}_{sj} \quad ; \\ &= \mathbf{B}_{il}^T \delta_{ls} \mathbf{B}_{sj} = \delta_{ij} \end{aligned}$$

Here the Einstein notation is used.

- $\mathbb{E}[\mathbf{Z}\mathbf{Z}^T] + \mathbf{J}_{2S} \mathbb{E}[\mathbf{Z}\mathbf{Z}^T] \mathbf{J}_{2S}$ *is full rank since:*

$$\begin{aligned} \mathbb{E}[\mathbf{Z}\mathbf{Z}^T] + \mathbf{J}_{2S} \mathbb{E}[\mathbf{Z}\mathbf{Z}^T] \mathbf{J}_{2S} &= \mathbf{B}^T \mathbf{C} \mathbf{B} + \mathbf{J}_{2S}^T \mathbf{B}^T \mathbf{C} \mathbf{B} \mathbf{J}_{2S} \\ &= \mathbf{B}^T \mathbf{C} \mathbf{B} + \mathbf{B}^T \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S} \mathbf{B} \\ &= \mathbf{B}^T (\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S}) \mathbf{B}. \end{aligned}$$

and $\mathbf{B}$ *and* $(\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S})$ *are the both full rank.*

A necessary condition for $\mathbf{Z}$ to belong to $B^{sym}(2S, L^2(\Omega))$ is that the second moments matrix $\mathbb{E}[\mathbf{Z}\mathbf{Z}^T]$ has rank at least equal to S. Indeed, since $\mathbf{J}_{2S}$ is a full rank matrix, the rank of

$\mathbf{J}_{2S}^T\mathbb{E}[\mathbf{Z}\mathbf{Z}^T]\mathbf{J}_{2S}$ is equal to the rank of $\mathbb{E}[\mathbf{Z}\mathbf{Z}^T]$. Then the conclusion is drawn by recalling that the sum of ranks is greater or equal to the rank of the sum (i.e. $\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) \geq \text{rank}(\mathbf{A} + \mathbf{B})$, $\forall \mathbf{A}, \mathbf{B}$).

Remark 5.6.2. *We recall that in the standard DO approach for parabolic equations (Section 5.2), one assumes that the second moments matrix $\mathbf{C} = \mathbb{E}[\mathbf{Y}\mathbf{Y}^T]$ is full rank ($\text{rank}(\mathbf{C}) = 2S$). Here we need the weaker assumption $\text{rank}(\mathbf{C} + \mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}) = 2S$. The motivation is related to the fact that we work in the phase-space coordinates: we need to uniquely determine the couple (q_S, p_S) and not the position and the momentum separately. Namely asking $Y_1, ..., Y_{2S}$ linearly independent is a too strong assumption in our model, as emphasized by the following example:*

$$\begin{cases} \ddot{q}(x,t,\omega) = \Delta q(x,t,\omega) & x \in (0,2\pi),\ t \in (0,T],\ \omega \in \Omega \\ q(0,t,\omega) = q(2\pi,t,\omega) = 0 & t \in (0,T],\ \omega \in \Omega \\ q(x,0,\omega) = q_0(x,\omega) = Z_1(\omega)\frac{1}{\sqrt{\pi}}\sin(x) & x \in [0,2\pi],\ \omega \in \Omega \\ \dot{q}(x,0,\omega) = p_0(x,\omega) = 0 & x \in [0,2\pi],\ \omega \in \Omega \end{cases} \qquad (5.36)$$

Here Z_1 is a square integrable random variable. We start looking for a symplectic decomposition of the initial data (q_0, p_0) in $\mathcal{M}_S^{sym}$. Problem (5.36) is linear with only one random variable which multiplies the initial data, which suggests to set $S = 1$. Hence we look for $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2) \in U(1, [H^1(D)]^2)$ and $\mathbf{Y} = (Y_1, Y_2)^T \in B^{sym}(2, L^2(\Omega))$ such that $\sum_{i=1}^2 \mathbf{U}_i Y_i = (Z_1(\omega)\frac{1}{\sqrt{\pi}}\sin(x), 0)^T$ (observe that, by working with symplectic bases, we can not decrease further the number of modes). The solution can be obtained by setting $Y_1 = Z_1$ and $Y_2 = 0$. Then the deterministic basis $\mathbf{U} \in U(1, [H^1(D)]^2)$ is uniquely determined by:

$$\mathbf{U} = \frac{1}{\sqrt{\pi}}\begin{pmatrix} \sin(x) & 0 \\ 0 & \sin(x) \end{pmatrix} \qquad (5.37)$$

Conversely, we can not find any symplectic basis $\mathbf{U} \in U(1, [H^1(D)]^2)$, or more generally $\mathbf{U} \in Sp(2, H^1(D) \times L^2(D))$, if we assume that $\mathbb{E}[\mathbf{Y}\mathbf{Y}^T]$ is full rank. Let us write $\mathbf{U} = \begin{bmatrix} Q_{I,1} & Q_{II,1} \\ P_{I,1} & P_{II,1} \end{bmatrix}$; if Y_1, Y_2 are linearly independent, then $P_{I,1}(x)Y_1(\omega) + P_{II,1}(x)Y_2(\omega) = p_0(x,\omega) = 0$ implies $P_{I,1} = P_{II,1} = 0$ and hence $\mathbf{U}^T \mathcal{J}_2 \mathbf{U} \neq \mathbf{J}_2$. Generally any symplectic basis $\mathbf{U} \in Sp(2S, H^1(D) \times L^2(D))$ consists of $2S$ linearly independent functions, which implies that we can not have $\mathbf{P}_{I,S} = \mathbf{P}_{II,S} = 0$. On the other hand we have seen that the symplectic decomposition of (q_0, p_0) in $\mathcal{M}_1^{sym}$ is well defined when the assumption of linear independence of Y_1, Y_2 is relaxed to $\mathbf{C} + \mathbf{J}_2^T\mathbf{C}\mathbf{J}_2$ full rank. This consideration generally applies to the solution of (5.36) at any time. The solution, given by $\mathbf{u}_1(t) = (q(t), \dot{q}(t)) = (Z_1(\omega)\cos(t)\frac{1}{\sqrt{\pi}}\sin(x), -Z_1(\omega)\sin(t)\frac{1}{\sqrt{\pi}}\sin(x)) \in \mathcal{M}_1^{sym}$, is characterized by a covariance matrix $\mathbf{C}(t) = \mathbb{E}[\mathbf{Y}(t)\mathbf{Y}^T(t)]$ with $Y_1(t) = Z_1\cos(t)$ and $Y_1(t) = -Z_1\sin(t)$ of defective rank while $\mathbf{C}(t) + \mathbf{J}_{2S}^T\mathbf{C}(t)\mathbf{J}_{2S}$ is full rank at any time. (We will see in the following that for this particular case the symplectic DO approximation degenerates to the Symplectic Proper Decomposition described in Section 5.5: the deterministic basis does not evolve, the coefficients evolve according to (5.26) and the approximation (with $S = 1$) is exact).

We emphasize that the full rank condition for $\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S}$ guarantees the uniqueness of the representation on $\mathbf{U}$ once $\mathbf{Y}$ is fixed. Namely, let $\mathbf{u}_S$ be in $\mathcal{M}_S^{sym}$; if $\mathbf{u}_S = \mathbf{U}\mathbf{Y} = \mathbf{W}\mathbf{Y}$ with $\mathbf{U}, \mathbf{W} \in U(S, [H^1(D)]^2)$ and $\mathbf{Y} \in \mathrm{B}^{sym}(2S, L^2(\Omega))$, then necessarily $\mathbf{U} = \mathbf{W}$. Indeed:

$$
\begin{aligned}
\mathbf{0} \quad &= (\mathbf{U} - \mathbf{W})\mathbf{Y} &&\Rightarrow (\mathbf{U} - \mathbf{W})\mathbf{C} = \mathbf{0} \\
&= \mathscr{J}_2(\mathbf{U} - \mathbf{W})\mathbf{J}_{2S}^T)\mathbf{Y} &&\Rightarrow (\mathbf{U} - \mathbf{W})\mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S} = \mathbf{0}
\end{aligned}
\tag{5.38}
$$

By summing the two equations on the right, we get $(\mathbf{U} - \mathbf{W})(\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S}) = \mathbf{0}$, which implies $\mathbf{U} = \mathbf{W}$ thanks to the full rank condition on $\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S}$. The same result does not apply if we extend the submanifold $U(S, [H^1(D)]^2)$ to the whole $Sp(2S, H^1(D) \times L^2(D))$. Consider for instance the random field $\mathbf{u}_1 = (q, p) = (Z(\omega) \frac{1}{2\sqrt{\pi}} \sin(x), Z(\omega) \frac{1}{\sqrt{\pi}} \sin(2x))$ with $x \in [0, 2\pi]$ and $Z \in L^2(\Omega)$. This can be represented, for instance, as:

$$
\mathbf{U} = \frac{1}{\sqrt{\pi}} \begin{pmatrix} \frac{1}{2}\sin(x) & -\sin(2x) \\ \sin(2x) & 0 \end{pmatrix} \qquad \mathbf{Y} = \begin{pmatrix} Z \\ 0 \end{pmatrix}
\tag{5.39}
$$

or equivalently as:

$$
\mathbf{W} = \frac{1}{\sqrt{\pi}} \begin{pmatrix} \frac{1}{2}\sin(x) & 0 \\ \sin(2x) & 2\sin(x) \end{pmatrix} \qquad \mathbf{Y} = \begin{pmatrix} Z \\ 0 \end{pmatrix}
\tag{5.40}
$$

where $\mathbf{U}, \mathbf{W} \in Sp(2, H^1(D) \times L^2(D))$ and $\mathbf{Y} \in \mathrm{B}^{sym}(2, L^2(\Omega))$. This implies that, if we replace $U(1, [H^1(D)]^2)$ with $Sp(2, H^1(D) \times L^2(D))$ in Definition 5.6.2, what we get is not a manifold anymore. Indeed, to get a manifold we need that the decomposition of $\mathbf{u}_S \in \mathcal{M}_S$, even though it is not unique in terms of $\mathbf{U}, \mathbf{Y}$, is uniquely characterized when one of the two bases is fixed. This implies that a stronger condition on $\mathbf{C}$ should be required when $U(1, [H^1(D)]^2)$ is extended to $Sp(2, H^1(D) \times L^2(D))$.

5.6.1 Parametrization of the tangent space by means of complex representation

In this section we discus how to equip $\mathcal{M}_S^{sym}$ with a differential manifold structure and parametrize the tangent space. This is achieved by identifying $\mathcal{M}_S^{sym}$, i.e. the manifold of all real valued symplectic random fields of rank S, with the manifold of all complex valued functions of rank S. To do so let us introduce the complex variable $\hat{v} = q + ip$ and its complex conjugate $\hat{v}^* = q - ip$. The Hamiltonian system (5.20), written in terms of the new variables $(\hat{v}, \hat{v}^*)$, becomes:

$$
\begin{aligned}
i\dot{\hat{v}} \quad &= 2\partial_{\hat{v}^*} H(\hat{v}, \hat{v}^*, \omega) \\
i\dot{\hat{v}}^* \quad &= -2\partial_{\hat{v}} H(\hat{v}, \hat{v}^*, \omega)
\end{aligned}
\tag{5.41}
$$

Observe that the second equation can be obtained from the first one by complex conjugation, thus it is redundant. The Hamiltonian function in (5.41) is now expressed with respect to the

new complex variables $\hat{v}$ and $\hat{v}^*$ and satisfies the reality condition:

$$H(\hat{v}, \hat{v}^*, \omega) = \big(H(\hat{v}, \hat{v}^* \omega)\big)^* =: H^*(\hat{v}^*, \hat{v}, \omega)$$

where with the symbol * we always denote the complex conjugate. We emphasize that the solution of (5.41), which is completely characterized by solving only one of the two equations in (5.41), is a complex valued function $\hat{v} : \bar{D} \times [0, T] \times \Omega \to \mathbb{C}$, whose real and imaginary parts correspond respectively to the position and momentum in system (5.20). In what follows, complex functions will be written as $\hat{v} = v^q + i v^p$, according to which the apex q and p will denote respectively the real and the imaginary part.

Definition 5.6.3. *We call complex S rank random field any function $\hat{u}_S \in H^1(D, \mathbb{C}) \otimes L^2(\Omega, \mathbb{C})$ which can be exactly expressed as:*

$$\hat{u}_S(\mathbf{x}, \omega) = \sum_{i=1}^{S} \hat{Y}_i(\omega) \hat{U}_i(\mathbf{x}) = \sum_{i=1}^{S} \big(Y_i^q(\omega) + i Y_i^p(\omega)\big)(U_i^q(\mathbf{x}) + i U_i^p(\mathbf{x})) \tag{5.42}$$

with:

- $\hat{\mathbf{U}} = (\hat{U}_1, ..., \hat{U}_S) \in St(S, H^1(D, \mathbb{C}))$,

- $\hat{\mathbf{Y}} = \hat{Y}_1, ..., \hat{Y}_S \in [L^2(\Omega, \mathbb{C})]^S$ *linearly independent random variables.*

Definition 5.6.4. *We define complex manifold of dimension S the collection of all complex S rank random fields:*

$$
\begin{aligned}
\mathcal{M}_S^{\mathbb{C}} &= \big\{\hat{u}_S = \sum_{i=1}^{S} \hat{U}_i \hat{Y}_i \,|\, span(\hat{U}_1, ..., \hat{U}_S) \in \mathcal{G}(S, H^1(D, \mathbb{C})), \; span(\hat{Y}_1, ..., \hat{Y}_S) \in \mathcal{G}(S, L^2(\Omega))\big\} \\
&= \big\{\hat{u}_S = \hat{\mathbf{U}}\hat{\mathbf{Y}}, \, \hat{\mathbf{U}} \in St(S, H^1(D, \mathbb{C})), \, \hat{\mathbf{Y}} = (\hat{Y}_1, .., \hat{Y}_S) \text{ linearly independent}\big\}
\end{aligned}
$$

Observe that $\mathcal{M}_S^{\mathbb{C}}$ is the complex version of the manifold $\mathcal{M}_S$, introduced in Section 5.2 to describe the DO approximation of real parabolic equations. Hence, $\mathcal{M}_S^{\mathbb{C}}$, as well as $\mathcal{M}_S$, can be equipped with a differential manifold structure by means of the same standard tools of differential geometry, recalled in Section 5.2. Complex manifolds of fixed rank have been already used in literature e.g. for the approximation of deterministic Schrödinger equations, see [67, 39]. Let us define the following map:

$$
\begin{aligned}
\pi : (St(S, H^1(D, \mathbb{C})), B(S, L^2(\Omega, \mathbb{C})) &\to \mathcal{M}_S^{\mathbb{C}} \\
(\hat{\mathbf{U}}, \hat{\mathbf{Y}}) &\to \sum_{i=1}^{S} \hat{U}_i \hat{Y}_i = \hat{u}_S
\end{aligned}
$$

where we denote by $B(S, L^2(\Omega, \mathbb{C}))$ the set of all S frames of linearly independent random variables in $L^2(\Omega, \mathbb{C})$. This map is surjective, i.e. $\mathcal{M}_S^{\mathbb{C}}$ is the image of $(St(S, H^1(D, \mathbb{C})), B(S, L^2(\Omega, \mathbb{C}))$ by π, but clearly non injective. The triple $\big(St(S, H^1(D, \mathbb{C})) \times B(S, L^2(\Omega, \mathbb{C})), \mathcal{M}_S^{\mathbb{C}}, \pi\big)$ defines a fiber bundle with fibers given by the group of the unitary matrices $U(S, \mathbb{C}^S) = \{\hat{\mathbf{W}} \in \mathbb{C}^{S \times S} :$

$\hat{\mathbf{W}}^*\hat{\mathbf{W}} = \hat{\mathbf{W}}\hat{\mathbf{W}}^* = \mathbb{I}\}$ and $\mathcal{M}_S^{\mathbb{C}}$ is isomorphic to the quotient space $(St(S, H^1(D,\mathbb{C})/U(S,\mathbb{C}^S)) \times B(S, L^2(\Omega,\mathbb{C}))$. The uniqueness of the representation of $\hat{u}_S \in \mathcal{M}_S^{\mathbb{C}}$ in terms of bases $(\hat{\mathbf{U}}, \hat{\mathbf{Y}}) \in (St(S, H^1(D,\mathbb{C})), B(S, L^2(\Omega,\mathbb{C}))$ is recovered in terms of unique decomposition in the tangent space, by imposing the following Gauge constraints [40, 90]:

$$\langle \delta \hat{U}_i, \hat{U}_j \rangle_h = \langle \delta U_i^q, U_j^q \rangle + \langle \delta U_i^p, U_j^p \rangle + i(\langle \delta U_i^p, U_j^q \rangle - \langle \delta U_i^q, U_j^p \rangle) = 0, \quad \forall i, j = 1, ..., S \quad (5.43)$$

for any $\delta \hat{\mathbf{U}} = (\delta \hat{U}_1, ..., \delta \hat{U}_S) \in \mathcal{T}_{\hat{u}_S}\mathcal{M}_S^{\mathbb{C}}$ and $\hat{\mathbf{U}} = (\hat{U}_1, ..., \hat{U}_S) \in \mathcal{M}_S^{\mathbb{C}}$. This leads to the following parametrization of the tangent space to $\mathcal{M}_S^{\mathbb{C}}$ at $\hat{u}_S = \sum_{i=1}^{S} \hat{U}_i \hat{Y}_i$:

$$\mathcal{T}_{\hat{u}_S}\mathcal{M}_S^{\mathbb{C}} = \left\{ \hat{\delta v} = \sum_{i=1}^{S} \left(\delta \hat{U}_i \hat{Y}_i + \hat{U}_i \delta \hat{Y}_i \right) \quad \text{with } \delta \hat{Y}_i \in L^2(\Omega,\mathbb{C}) \text{ and } \delta \hat{U}_i \in H^1(D,\mathbb{C}), \\ \text{s.t. } \langle \delta \hat{U}_i, \hat{U}_j \rangle_h = 0, \, \forall i, j = 1, ..., S \right\} \quad (5.44)$$

Remark 5.6.3. *$\mathcal{T}_{\hat{u}_S}\mathcal{M}_S^{\mathbb{C}}$ is a complex linear space, hence $\delta \hat{v}$ belongs to $\mathcal{T}_{\hat{u}_S}\mathcal{M}_S^{\mathbb{C}}$ if and only if $i\delta \hat{v}$ belongs to $\mathcal{T}_{\hat{u}_S}\mathcal{M}_S^{\mathbb{C}}$.*

The complex Hilbert space $H^1(D,\mathbb{C}) \otimes L^2(\Omega,\mathbb{C})$, equipped with the usual hermitian L^2 product, can be identified with the real space $[H^1(D,\mathbb{R}) \otimes L^2(\Omega,\mathbb{R})]^2$, equipped with the complex structure associated to $\mathcal{J}_2$. Namely the following map is bijective

$$H^1(D,\mathbb{C}) \otimes L^2(\Omega,\mathbb{C}) \quad \rightarrow [H^1(D,\mathbb{R}) \otimes L^2(\Omega,\mathbb{R})]^2$$
$$\hat{u} = u^q + i u^p \quad \rightarrow (u^q, u^p)^T =: \mathbf{u}$$

and for all $\hat{u}, \hat{v} \in H^1(D,\mathbb{C}) \otimes L^2(\Omega,\mathbb{C})$ we have:

$$\mathbb{E}[\langle \hat{u}, \hat{v} \rangle_h] = \mathbb{E}[\langle \mathbf{u}, \mathbf{v} \rangle] - i\mathbb{E}[\langle \mathbf{u}, \mathcal{J}_2 \mathbf{v} \rangle], \quad (5.45)$$

where $\langle \cdot, \cdot \rangle$ is the standard L^2 product in the real space. Observe that the imaginary part of the Hermitian product (5.45) coincides with the canonical symplectic form of $[H^1(D,\mathbb{R}) \otimes L^2(\Omega,\mathbb{R})]^2$ defined in (5.21) with changed sign:

$$\Im(\mathbb{E}[\langle \hat{u}, \hat{v} \rangle_h]) = -\vartheta(\mathbf{u}, \mathbf{v}) = -\mathbb{E}[\langle \mathbf{u}, \mathcal{J}_2 \mathbf{v} \rangle] \quad (5.46)$$

Similarly $[L^2(\Omega,\mathbb{C})]^S$ can be identified with $[L^2(\Omega,\mathbb{R})]^{2S}$, i.e:

$$[L^2(\Omega,\mathbb{C})]^S \quad \rightarrow [L^2(\Omega,\mathbb{R})]^{2S}$$
$$\hat{\mathbf{Z}} = (Z_1^q + iZ_1^p, ..., Z_S^q + iZ_S^p) \quad \rightarrow (Z_1^q, ..., Z_S^q, Z_1^p, ..., Z_S^p)^T =: (\mathbf{Z}^q, \mathbf{Z}^p)^T = \mathbf{Z}$$

and

$$\mathbb{E}[\hat{\mathbf{Z}}^*\hat{\mathbf{Y}}] = \mathbb{E}[\mathbf{Y}^T\mathbf{Z}] - i\mathbb{E}[\mathbf{Y}^T\mathbf{J}_{2S}\mathbf{Z}]$$
$$= \mathbb{E}[\mathbf{Y}^{\mathbf{q}T}\mathbf{Z}^{\mathbf{q}}] + \mathbb{E}[\mathbf{Y}^{\mathbf{p}T}\mathbf{Z}^{\mathbf{p}}] + i(\mathbb{E}[\mathbf{Y}^{\mathbf{p}T}\mathbf{Z}^{\mathbf{q}}] - \mathbb{E}[\mathbf{Y}^{\mathbf{q}T}\mathbf{Z}^{\mathbf{p}}]).$$

Let $GL([L^2(\Omega,\mathbb{C})]^S, H^1(D,\mathbb{C}) \otimes L^2(\Omega,\mathbb{C}))$ be the set of all bounded linear maps from $[L^2(\Omega,\mathbb{C})]^S$ to $H^1(D,\mathbb{C}) \otimes L^2(\Omega,\mathbb{C})$. We denote by $\hat{\phi}_{\hat{\mathbf{A}}}$ the map of $GL([L^2(\Omega,\mathbb{C})]^S, H^1(D,\mathbb{C}) \otimes L^2(\Omega,\mathbb{C}))$ which

can be represented as:

$$\hat{\phi}_{\hat{\mathbf{A}}} : \begin{array}{ll} [L^2(\Omega,\mathbb{C})]^S & \to H^1(D,\mathbb{C}) \otimes L^2(\Omega,\mathbb{C}) \\ \hat{\mathbf{Z}} & \mapsto \hat{\phi}_{\hat{\mathbf{A}}}(\hat{\mathbf{Z}}) = \hat{\mathbf{A}}\hat{\mathbf{Z}} \end{array} \qquad (5.47)$$

for any (row) vector of deterministic complex functions $\hat{\mathbf{A}} \in [H^1(D,\mathbb{C})]^S$. Let $[L^2(\Omega,\mathbb{C})]^S$ and $H^1(D,\mathbb{C}) \otimes L^2(\Omega,\mathbb{C})$ be identified with $[L^2(\Omega,\mathbb{R})]^{2S}$ and $[H^1(D,\mathbb{R}) \otimes L^2(\Omega,\mathbb{R})]^2$ respectively, and let ϕ be the function $\hat{\phi}_{\hat{\mathbf{A}}}$ in real setting. Then, ϕ must satisfies:

$$\phi : \begin{array}{ll} [L^2(\Omega,\mathbb{R})]^{2S} & \to [H^1(D,\mathbb{R}) \otimes L^2(\Omega,\mathbb{R})]^2 \\ \mathbf{Z} & \mapsto \phi(\mathbf{Z}) = \mathbf{u} \iff \hat{u} = \hat{\mathbf{A}}\hat{\mathbf{Z}} \end{array} \qquad (5.48)$$

where $\mathbf{Z}$ and $\mathbf{u}$ are the realification of $\hat{\mathbf{Z}}$ and $\hat{\phi}(\hat{\mathbf{Z}})$ respectively. The map ϕ is linear and can be written in terms of a matrix of functions $\mathbf{A} \in [H^1(D,\mathbb{R}) \times H^1(D,\mathbb{R})]^{2S}$ such that $\phi(\mathbf{Z}) = \mathbf{AZ} \iff \hat{u} = \hat{\mathbf{A}}\hat{\mathbf{Z}}$. Precisely for any $\hat{\mathbf{A}} = (\mathbf{A}^q + i\mathbf{A}^p) \in [H^1(D,\mathbb{C})]^S$, the map $\hat{\phi}_{\hat{\mathbf{A}}}$ is identified in real setting with $\phi_{\mathbf{A}} : [L^2(\Omega,\mathbb{R})]^{2S} \to [H^1(D,\mathbb{R}) \otimes L^2(\Omega,\mathbb{R})]^2$ where $\mathbf{A}$ is given:

$$\mathbf{A} := \begin{pmatrix} \mathbf{A}^q & -\mathbf{A}^p \\ \mathbf{A}^p & \mathbf{A}^q \end{pmatrix} \qquad (5.49)$$

The proof is an exercise of linear algebra [113, 114]. We say that $\mathbf{A}$ is the real matrix representation of $\hat{\mathbf{A}}$ and write $\hat{\mathbf{A}} \sim \mathbf{A}$. This motivates the real identification of row-vector complex functions which will be used in the following. Observe that in this setting the complex conjugate simply corresponds to the transpose: $\hat{\mathbf{A}}^* \sim \mathbf{A}^T$ and the hermitian product $\langle A_i^q, B_j^q \rangle + \langle A_i^p, B_j^p \rangle + i(\langle A_i^p, B_j^q \rangle - \langle A_i^q, B_j^p \rangle)$ can be computed by matrix multiplication as:

$$\begin{aligned} \langle \hat{A}_i, \hat{B}_j \rangle_h \quad &\sim \left\langle \begin{pmatrix} A_i^q & -A_i^p \\ A_i^p & A_i^q \end{pmatrix} \begin{pmatrix} B_i^q & B_i^p \\ -B_i^p & B_i^q \end{pmatrix} \right\rangle \\ &= \begin{pmatrix} \langle A_i^q, B_j^q \rangle + \langle A_i^p, B_j^p \rangle & \langle A_i^q, B_j^p \rangle - \langle A_i^p, B_j^q \rangle \\ \langle A_i^p, B_j^q \rangle - \langle A_i^q, B_j^p \rangle & \langle A_i^q, B_j^q \rangle + \langle A_i^p, B_j^p \rangle \end{pmatrix} \end{aligned}$$

where the last matrix is indeed the real matrix representation of $\langle \hat{A}_i, \hat{B}_j \rangle_h$. Moreover the real multiplication by $\mathscr{J}_2$ corresponds to the complex multiplication with the imaginary unit i. Namely if $\mathbf{A}$ is the real matrix representation of $\hat{\mathbf{A}}$, then $\mathscr{J}_2\mathbf{A}$ is the real matrix representation of $i\hat{\mathbf{A}}$. The same procedure in finite dimension leads to representing a complex matrix by a real matrix of double dimension, i.e. $\hat{\mathbf{A}} = \mathbf{A}^q + i\mathbf{A}^p \in \mathbb{C}^{S \times S}$ is represented by $\mathbf{A} \in \mathbb{R}^{2S \times 2S}$, written as in (5.49), with $\mathbf{A}^q$ and $\mathbf{A}^p$ real matrices in $\mathbb{R}^{S \times S}$.

Lemma 5.6.1. *The manifold $\mathscr{M}_S^{\mathbb{C}}$ of all S rank complex random fields is isomorphic to the manifold $\mathscr{M}_S^{sym}$ in Definition 5.6.2.*

Proof. The proof is based on the real representation of complex valued functions introduced before. Let $\hat{\mathbf{U}} = (\hat{U}_1, ..., \hat{U}_S) \in St(S, H^1(D,\mathbb{C}))$ and U_i^q, U_i^p denote respectively the real and imaginary part of $\hat{U}_i$ for any $i = 1, ..., S$. The orthonormality condition $\langle \hat{U}_i, \hat{U}_j \rangle_h = \delta_{ij}$ is written

component-wise as:

$$\langle U_i^q, U_j^q \rangle + \langle U_i^p, U_j^p \rangle = \delta_{ij}, \text{ and } \langle U_i^p, U_j^q \rangle - \langle U_i^q, U_j^p \rangle = 0, \qquad \forall i, j = 1, ..., S. \qquad (5.50)$$

Let $\mathbf{U}$ be the real matrix representation of $\hat{\mathbf{U}}$ as defined in (5.49). This is written as:

$$\hat{\mathbf{U}} \sim \mathbf{U} = \begin{pmatrix} \mathbf{Q} & -\mathbf{P} \\ \mathbf{P} & \mathbf{Q} \end{pmatrix} \quad \text{with} \quad \begin{array}{ll} \mathbf{Q} = (Q_1, ..., Q_S) & : Q_i = \hat{U}_i^q \in H^1(D, \mathbb{R}), \quad \forall i = 1, ..., S, \\ \mathbf{P} = (P_1, ..., P_S) & : P_i = \hat{U}_i^p \in H^1(D, \mathbb{R}), \quad \forall i = 1, ..., S. \end{array}$$

Observe that condition (5.50) coincides with condition (5.29). Thus, form Proposition 5.6.1 (point b) we have that $\hat{\mathbf{U}} \in St(S, H^1(D, \mathbb{C}))$ if and only if $\mathbf{U} \in U(S, [H^1(D, \mathbb{R})]^2)$. It follows that any element $\hat{\mathbf{U}} \in St(S, H^1(D, \mathbb{C}))$ can be uniquely identified with an element $\mathbf{U} \in U(S, [H^1(D, \mathbb{R})]^2)$. Consider now $\hat{\mathbf{Z}} \in [L^2(\Omega, \mathbb{C})]^S$ and its realification $\mathbf{Z} = (\mathbf{Z}^q, \mathbf{Z}^p)^T \in [L^2(\Omega, \mathbb{R})]^{2S}$. The components $(\hat{Z}_1, ..., \hat{Z}_S)$ of $\hat{\mathbf{Z}}$ are linearly independent if and only if the following matrix

$$\mathbb{E}[\hat{\mathbf{Z}}\hat{\mathbf{Z}}^*] = \mathbb{E}[\mathbf{Z}^q\mathbf{Z}^{qT}] + \mathbb{E}[\mathbf{Z}^p\mathbf{Z}^{pT}] + i(\mathbb{E}[\mathbf{Z}^p\mathbf{Z}^{qT}] - \mathbb{E}[\mathbf{Z}^q\mathbf{Z}^{pT}]) \in \mathbb{C}^{S \times S} \qquad (5.51)$$

has full rank. Observe that the real matrix representation of $\mathbb{E}[\hat{\mathbf{Z}}\hat{\mathbf{Z}}^*]$ is given by:

$$\begin{pmatrix} \mathbb{E}[\mathbf{Z}^q\mathbf{Z}^{qT}] + \mathbb{E}[\mathbf{Z}^p\mathbf{Z}^{pT}] & \mathbb{E}[\mathbf{Z}^q\mathbf{Z}^{pT}] - \mathbb{E}[\mathbf{Z}^p\mathbf{Z}^{qT}] \\ \mathbb{E}[\mathbf{Z}^p\mathbf{Z}^{qT}] - \mathbb{E}[\mathbf{Z}^q\mathbf{Z}^{pT}] & \mathbb{E}[\mathbf{Z}^q\mathbf{Z}^{qT}] + \mathbb{E}[\mathbf{Z}^p\mathbf{Z}^{pT}] \end{pmatrix} = \left(\mathbb{E}[\mathbf{Z}\mathbf{Z}^T] + \mathbf{J}_{2S}^T \mathbb{E}[\mathbf{Z}\mathbf{Z}^T]\mathbf{J}_{2S} \right) \in \mathbb{R}^{2S \times 2S} \quad (5.52)$$

This implies that $(\hat{Z}_1, ..., \hat{Z}_S)$ are linearly independent if and only if $\mathbb{E}[\mathbf{Z}\mathbf{Z}^T] + \mathbb{E}[\mathbf{J}_{2S}\mathbf{Z}\mathbf{Z}^T\mathbf{J}_{2S}^T]$ is full rank. Observe also that $\mathbb{E}[\mathbf{Z}\mathbf{Z}^T] + \mathbb{E}[\mathbf{J}_{2S}\mathbf{Z}\mathbf{Z}^T\mathbf{J}_{2S}^T]$ is the real matrix representation of $\mathbb{E}[\hat{\mathbf{Z}}\hat{\mathbf{Z}}^*]$, hence the two identifications are consistent. It follows that $B(S, L^2(\Omega, \mathbb{C}))$ can be uniquely identified with $B^{sym}(2S, L^2(\Omega, \mathbb{R}))$.

Finally any $\hat{u}_S = \hat{\mathbf{U}}\hat{\mathbf{Y}} \in \mathcal{M}_S^{\mathbb{C}}$, with $\hat{\mathbf{U}} \in St(S, H^1(D, \mathbb{C}))$ and $\hat{\mathbf{Y}} \in B(S, L^2(\Omega, \mathbb{C}))$, can be uniquely represented in real setting as $\mathbf{u}_S = \mathbf{U}\mathbf{Y} \in \mathcal{M}_S^{sym}$ where $\mathbf{U} \in U(S, [H^1(D, \mathbb{R})]^2)$ and $\mathbf{Y} \in B^{sym}(2S, L^2(\Omega, \mathbb{R}))$ are the real representations of $\hat{\mathbf{U}}$ and $\hat{\mathbf{Y}}$ respectively. $\qquad \square$

We now rewrite Lemma 5.6.1 in real setting to recover a unique representation of S-rank random fields $\mathbf{u}_S \in \mathcal{M}_S^{sym}$ in terms of the bases in $(\mathbf{U}, \mathbf{Y}) \in (U(S, [H^1(D)]^2), B(S, L^2(\Omega, \mathbb{C})))$.

Proposition 5.6.5. *In real setting, the orthogonal condition* (5.43) *is reinterpreted as:*

$$\langle \delta\mathbf{U}_i, \mathbf{U}_j^+ \rangle = \langle \delta\mathbf{U}_i, \mathbf{U}_j \rangle = 0 \qquad \forall i, j = 1, ..., 2S \qquad (5.53)$$

We mention that condition (5.53) can be directly derived, without making use of the isomorphism with $\mathcal{M}_S^{\mathbb{C}}$, by quotienting $U(S, [H^1(D)]^2)$ by $U(S, \mathbb{R}^{2S})$. This is perfectly consistent with the construction discussed before, being $U(S, \mathbb{R}^{2S})$ isomorphic to $\mathcal{O}(S, \mathbb{C})$. Condition (5.53) can be seen as a symplectic orthogonality condition: we ask that $\delta\mathbf{U}$ belongs to the symplectic orthogonal complement to $\mathbf{U}$ at each time:

$$\mathcal{P}_{\mathbf{U}}^{sym}[\delta\mathbf{U}] = \mathcal{P}_{\mathbf{U}}[\delta\mathbf{U}] = \mathbf{0}$$

Observe that condition (5.53) preserves the orthogonal-symplectic structure of the basis, namely if $\mathbf{U}(t)$ is the integrable curve passing by $\mathbf{U}(0) \in U(S, [H^1(D)]^2)$, of a vector field which satisfies (5.53), then $\mathbf{U}(t) \in U(S, [H^1(D)]^2)$ at any time.

Proposition 5.6.6. *Let* $\mathbf{U}(t)$ *be a smooth curve in* $[H^1(D)]^{2S}$ *such that:*

- $\mathbf{U}(0) \in U(S, [H^1(D)]^2)$,

- $\langle \dot{\mathbf{U}}_i, \mathbf{U}_j^+ \rangle = \langle \dot{\mathbf{U}}_i, \mathbf{U}_j \rangle = 0, \; \forall i, j = 1, ..., 2S, \; and \; \forall t \in [0, T],$

then $\mathbf{U}(t) \in U(S, [H^1(D)]^2)$ *for all* t.

Proof. We start by showing that the orthogonality is preserved:

$$\frac{d}{dt}\langle \mathbf{U}_j(t), \mathbf{U}_i(t) \rangle = \langle \dot{\mathbf{U}}_j(t), \mathbf{U}_i(t) \rangle + \langle \mathbf{U}_j(t), \dot{\mathbf{U}}_i(t) \rangle = 0$$

implies that $\langle \mathbf{U}_j(t), \mathbf{U}_i(t) \rangle = \langle \mathbf{U}_j(0), \mathbf{U}_i(0) \rangle = \delta_{ij}, \; \forall i, j = 1, ..., 2S$ and $\forall t \in [0, T]$. Similarly for the symplecticity:

$$\begin{aligned}
\frac{d}{dt}\langle \mathcal{J}_2 \mathbf{U}(t), \mathbf{U}(t) \rangle &= \langle \mathcal{J}_2 \dot{\mathbf{U}}(t), \mathbf{U}(t) \rangle + \langle \mathcal{J}_2 \mathbf{U}(t), \dot{\mathbf{U}}(t) \rangle \\
&= \langle \mathbf{U}(t), \mathscr{P}_{\mathbf{U}(t)}[\mathcal{J}_2 \dot{\mathbf{U}}(t)] \rangle + \langle \mathcal{J}_2 \mathbf{U}(t), \mathscr{P}_{\mathcal{J}_2 \mathbf{U}(t)}[\dot{\mathbf{U}}(t)] \rangle \\
&= \langle \mathbf{U}(t), \mathcal{J}_2 \mathscr{P}_{\mathbf{U}(t)}[\dot{\mathbf{U}}(t)] \rangle + \langle \mathcal{J}_2 \mathbf{U}(t), \mathscr{P}_{\mathbf{U}(t)}[\dot{\mathbf{U}}(t)] \rangle = \mathbf{0}
\end{aligned}$$

which implies $\langle \mathcal{J}_2 \mathbf{U}(t), \mathbf{U}(t) \rangle = \langle \mathcal{J}_2 \mathbf{U}(0), \mathbf{U}(0) \rangle = \mathcal{J}_2$. $\qquad\square$

The dynamic condition (5.53) induces a bijection between $(U(S, [H^1(D)]^2)/U(S, \mathbb{R}^{2S})) \times B^{sym}(2S, L^2(\Omega, \mathbb{R})))$ and $\mathcal{M}_S^{sym}$ which allows to equip $\mathcal{M}_S^{sym}$ with a differential manifold structure. In particular, for any $\mathbf{u}_S \in \mathcal{M}_S^{sym}$, the tangent space to $\mathcal{M}_S^{sym}$ at $\mathbf{u}_S = \mathbf{U}\mathbf{Y}$ is parametrized as follows:

Lemma 5.6.2. *For any* $\mathbf{u}_S = \mathbf{U}\mathbf{Y} \in \mathcal{M}_S^{sym}$, *the tangent space to* $\mathcal{M}_S^{sym}$ *at* $\mathbf{u}_S$ *is the subspace of* $[H^1(D)]^2 \otimes L^2(\Omega)$ *given by:*

$$\begin{aligned}
\mathcal{T}_{\mathbf{u}_S}\mathcal{M}_S^{sym} &= \Big\{ \delta\mathbf{u}_S = (\delta\mathbf{U})\mathbf{Y} + \mathbf{U}\delta\mathbf{Y} \in [H^1(D)]^2 \otimes L^2(\Omega): \quad \delta\mathbf{Y} \in [L^2(\Omega, \mathbb{R})]^{2S}, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \delta\mathbf{U} \in \mathcal{U}^{sym\perp}: \; \mathcal{J}_2^T(\delta\mathbf{U})\mathbf{J}_{2S} = \delta\mathbf{U} \Big\} \\
&= \Big\{ \delta\mathbf{u}_S = \sum_{i=1}^{2S}(\delta\mathbf{U}_i Y_i + \mathbf{U}_i \delta Y_i): \quad \delta Y_i \in L^2(\Omega, \mathbb{R}) \; and \; \delta\mathbf{U}_i \in [H^1(D)]^2, \\
&\qquad\qquad\qquad\qquad s.t. \; \mathcal{J}_2^T \delta\mathbf{U}\mathbf{J}_{2S} = \delta\mathbf{U}, \langle \delta\mathbf{U}_i, \mathbf{U}_j \rangle = 0, \forall i, j = 1, ..., 2S \Big\}
\end{aligned}$$

$$(5.54)$$

The following property holds for any $\mathbf{u}_S = \mathbf{U}\mathbf{Y} \in \mathcal{M}_S^{sym}$:

Proposition 5.6.7. $\mathbf{v} \in \mathcal{T}_{\mathbf{u}_S}\mathcal{M}_S^{sym}$ *if and only* $\mathcal{J}_2\mathbf{v} \in \mathcal{T}_{\mathbf{u}_S}\mathcal{M}_S^{sym}$

Proposition 5.6.7 follows directly from the diffeomorphism between $\mathcal{M}_S^{\mathbb{C}}$ and $\mathcal{M}_S^{sym}$, see Remark 5.6.3 for the same result in complex setting. We emphasize that this property does not apply to arbitrary symplectic manifolds, and in particular, does not hold when the space of symplectic deterministic bases is not restricted to $U(S, [H^1(D)]^2)$. Observe that Proposition 5.6.7 implies that the symplectic form defined in (5.21) is not degenerate in $\mathcal{M}_S^{sym}$. Indeed for any $\mathbf{v} \in \mathcal{T}_{\mathbf{u}_S} \mathcal{M}_S^{sym}$ such that $\mathbf{v} \neq \mathbf{0}$, $\vartheta(\mathscr{J}_2 \mathbf{v}, \mathbf{v}) = \mathbb{E}[\langle \mathscr{J}_2 \mathbf{u}, \mathscr{J}_2 \mathbf{v} \rangle] = \|\mathbf{v}\|_{[L^2(D)]^2 \otimes L^2(\Omega)}^2 > 0$.

Lemma 5.6.3. *Let* $\mathbf{u}_S \in \mathcal{M}_S^{sym}$ *be written as* $\mathbf{u}_S = \mathbf{U}\mathbf{Y}$. *For any* $\mathbf{v} = (\delta \mathbf{U})\mathbf{Y} + \mathbf{U}\delta\mathbf{Y} \in \mathcal{T}_{\mathbf{u}_S} \mathcal{M}_S^{sym}$, $\delta \mathbf{U}$ *and* $\delta\mathbf{Y}$ *are uniquely characterized as:*

$$
\begin{aligned}
\delta\mathbf{Y} &= \langle \mathscr{P}_{\mathbf{U}}[\mathbf{v}], \mathbf{U} \rangle \\
\delta\mathbf{U}(\mathbf{C} + \mathbf{J}_{2S} \mathbf{C} \mathbf{J}_{2S}^T) &= \mathscr{P}_{\mathbf{U}}^{\perp, sym} \left[\mathbb{E}[\mathbf{v}\mathbf{Y}^T] + \mathscr{J}_2 \mathbb{E}[\mathbf{v}\mathbf{Y}^T \mathbf{J}_{2S}^T] \right]
\end{aligned}
\tag{5.55}
$$

Proof. Let $\tilde{\mathbf{v}} \in [H^1(D)]^2 \otimes L^2(\Omega)$ and $\mathbf{v} = (\delta\mathbf{U})\mathbf{Y} + \mathbf{U}\delta\mathbf{Y}$ be the projection of $\tilde{\mathbf{v}}$ in the tangent space $\mathcal{T}_{\mathbf{u}_S} \mathcal{M}_S^{sym}$, that is:

$$
\mathbb{E}[\langle \tilde{\mathbf{v}}, \mathbf{w} \rangle] = \mathbb{E}[\langle \mathbf{v}, \mathbf{w} \rangle] \qquad \forall \mathbf{w} \in \mathcal{T}_{\mathbf{u}_S} \mathcal{M}_S^{sym}
\tag{5.56}
$$

According to (5.54) this can be written as:

$$
\mathbb{E}[\langle \tilde{\mathbf{v}}, \mathbf{W}\mathbf{Y} + \mathbf{U}\mathbf{Z} \rangle] = \mathbb{E}[\langle (\delta\mathbf{U})\mathbf{Y} + \mathbf{U}\delta\mathbf{Y}, \mathbf{W}\mathbf{Y} + \mathbf{U}\mathbf{Z} \rangle]
\tag{5.57}
$$

for any $\mathbf{Z} \in [L^2(\Omega, \mathbb{R})]^{2S}$ and $\mathbf{W} \in \mathcal{U}^{sym\perp}$ which satisfies $\mathscr{J}_2^T \mathbf{W} \mathbf{J}_{2S} = \mathbf{W}$. We need to verify that $\delta\mathbf{U}$ and $\delta\mathbf{Y}$ are uniquely characterized only in terms of $\tilde{\mathbf{v}}$, $\mathbf{U}$ and $\mathbf{Y}$.

- By testing against $\mathbf{U}\mathbf{Z}$ (i.e. setting $\mathbf{W} = 0$), we easily recover the characterization of $\delta\mathbf{Y}$:

$$
\begin{aligned}
\mathbb{E}[\langle \tilde{\mathbf{v}}, \mathbf{U}\mathbf{Z} \rangle] &= \mathbb{E}[\langle \mathbf{U}\delta\mathbf{Y}, \mathbf{U}\mathbf{Z} \rangle], \\
\Rightarrow \quad \mathbb{E}[\langle \tilde{\mathbf{v}}, \mathbf{U} \rangle \mathbf{Z}] &= \mathbb{E}[\mathbf{Z}^T \delta\mathbf{Y}] \qquad \forall \mathbf{Z} \in [L^2(\Omega, \mathbb{R})]^{2S},
\end{aligned}
\tag{5.58}
$$

  which leads to:

$$
\delta\mathbf{Y} = \langle \mathscr{P}_{\mathbf{U}}[\tilde{\mathbf{v}}], \mathbf{U} \rangle.
$$

- We now want to test against $\mathbf{V}\mathbf{Y}$ for $\mathbf{V}$ arbitrary in $\mathcal{U}^{sym\perp}$ and satisfying $\mathscr{J}_2^T \mathbf{V}\mathbf{J}_{2S} = \mathbf{V}$. The last condition can be replaced by setting $\mathbf{V} = \frac{1}{2}\left(\mathscr{J}_2^T \mathbf{W} \mathbf{J}_{2S} + \mathbf{W} \right)$ with arbitrary $\mathbf{W} \in \mathcal{U}^{sym\perp}$. Thus we have:

$$
\mathbb{E}[\langle \tilde{\mathbf{v}}, \mathscr{J}_2^T \mathbf{V}\mathbf{J}_{2S}\mathbf{Y} + \mathbf{V}\mathbf{Y} \rangle] = \mathbb{E}[\langle (\delta\mathbf{U})\mathbf{Y}, \mathscr{J}_2^T \mathbf{V}\mathbf{J}_{2S}\mathbf{Y} + \mathbf{V}\mathbf{Y} \rangle] \qquad \forall \mathbf{V} \in \mathcal{U}^{sym\perp}
\tag{5.59}
$$

  The left hand side can be rewritten as

$$
\mathbb{E}[\langle \tilde{\mathbf{v}}, \mathscr{J}_2^T \mathbf{V}\mathbf{J}_{2S}\mathbf{Y} + \mathbf{V}\mathbf{Y} \rangle] = \langle \mathbb{E}[\mathscr{J}_2 \tilde{\mathbf{v}}\mathbf{Y}^T \mathbf{J}_{2S}^T], \mathbf{V} \rangle + \langle \mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T], \mathbf{V} \rangle,
$$

while for the right hand side we have:

$$
\begin{aligned}
\mathbb{E}[\langle (\delta\mathbf{U})\mathbf{Y}, \mathscr{J}_2^T\mathbf{V}\mathbf{J}_{2S}\mathbf{Y} + \mathbf{V}\mathbf{Y}\rangle] \quad &= \mathbb{E}[\langle (\delta\mathbf{U})\mathbf{Y}, \mathscr{J}_2^T\mathbf{V}\mathbf{J}_{2S}\mathbf{Y}\rangle] + \mathbb{E}[\langle (\delta\mathbf{U})\mathbf{Y}, \mathbf{V}\mathbf{Y}\rangle] \\
&= \mathbb{E}[\langle \mathscr{J}_2(\delta\mathbf{U})\mathbf{Y}, \mathbf{V}\mathbf{J}_{2S}\mathbf{Y}\rangle] + \mathbb{E}[\langle (\delta\mathbf{U})\mathbf{Y}, \mathbf{V}\mathbf{Y}\rangle] \\
&= \mathbb{E}[\langle (\delta\mathbf{U})\mathbf{J}_{2S}\mathbf{Y}, \mathbf{V}\mathbf{J}_{2S}\mathbf{Y}\rangle] + \mathbb{E}[\langle (\delta\mathbf{U})\mathbf{Y}, \mathbf{V}\mathbf{Y}\rangle] \\
&= \langle \delta\mathbf{U}, \mathbf{V}\rangle\mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T + \langle \delta\mathbf{U}, \mathbf{V}\rangle\mathbf{C}
\end{aligned}
\tag{5.60}
$$

where we used the fact that $\mathscr{J}_2^T\delta\mathbf{U}\mathbf{J}_{2S} = \delta\mathbf{U}$. By combining the two parts we get:

$$
\langle \mathbb{E}[\mathscr{J}_2\tilde{\mathbf{v}}\mathbf{Y}^T\mathbf{J}_{2S}^T], \mathbf{V}\rangle + \langle \mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T], \mathbf{V}\rangle = \langle \delta\mathbf{U}, \mathbf{V}\rangle\mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T + \langle \delta\mathbf{U}, \mathbf{V}\rangle\mathbf{C}
\tag{5.61}
$$

for any $\mathbf{V} \in \mathscr{U}^{sym\perp}$. By using Preposition 5.6.4 we finally obtain:

$$
\delta\mathbf{U}(\mathbf{C} + \mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T) = \mathscr{P}_{\mathbf{U}}^{\perp,sym}\Big[\mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T] + \mathscr{J}_2\mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T\mathbf{J}_{2S}^T]\Big].
\tag{5.62}
$$

Observe that $\delta\mathbf{U}$ is completely characterized, thanks to the full rank assumption on $\mathbf{C} + \mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T$.

We finally verify that condition $\mathscr{J}_2(\delta\mathbf{U})\mathbf{J}_{2S}^T = \delta\mathbf{U}$ is actually satisfied. We observe that $\mathscr{J}_2(\delta\mathbf{U})\mathbf{J}_{2S}^T = \delta\mathbf{U}$ applies if and only if

$$
\delta\mathbf{U}(\mathbf{C} + \mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T)\mathbf{J}_{2S}^T = \mathscr{J}_2(\delta\mathbf{U})\mathbf{J}_{2S}^T(\mathbf{C} + \mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T)\mathbf{J}_{2S}^T = -\mathscr{J}_2(\delta\mathbf{U})(\mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T + \mathbf{C}).
$$

Then, from (5.62) and Proposition 5.6.4 follows that:

$$
\begin{aligned}
\delta\mathbf{U}(\mathbf{C} + \mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T)\mathbf{J}_{2S}^T \quad &= \mathscr{P}_{\mathbf{U}}^{\perp,sym}\Big[\mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T] + \mathscr{J}_2\mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T\mathbf{J}_{2S}^T]\Big]\mathbf{J}_{2S}^T \\
&= \mathscr{P}_{\mathbf{U}}^{\perp,sym}\Big[\mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T\mathbf{J}_{2S}^T] + \mathscr{J}_2^T\mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T]\Big] \\
&= \mathscr{J}_2^T\mathscr{P}_{\mathbf{U}}^{\perp,sym}\Big[\mathscr{J}_2\mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T\mathbf{J}_{2S}^T] + \mathbb{E}[\tilde{\mathbf{v}}\mathbf{Y}^T]\Big] \\
&= \mathscr{J}_2^T\delta\mathbf{U}(\mathbf{C} + \mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T) \\
&= -\mathscr{J}_2\delta\mathbf{U}(\mathbf{C} + \mathbf{J}_{2S}\mathbf{C}\mathbf{J}_{2S}^T)
\end{aligned}
$$

which concludes the proof. $\qquad\square$

5.6.2 DLR Variational Principle in complex and real setting

Our goal is to find a dynamical low rank approximation $\mathbf{u}_S \in \mathscr{M}_S^{sym}$ of problem (5.3), which is written as:

$$
\mathbf{u}_S(\mathbf{x}, t, \omega) = \sum_{i=1}^{2S} \mathbf{U}_i(\mathbf{x}, t)\, Y_i(t, \omega)
\tag{5.63}
$$

To do so we exploit the diffeomorphism between $\mathscr{M}_S^{sym}$ and $\mathscr{M}_S^{\mathbb{C}}$.
We start by considering problem (5.41). In complex setting, since this is a first order PDE we can apply the DO approximation described in Section 5.2. The DO variational principle for

problem (5.41) reads as follows:

Complex DLR Variational Principle 1. *At each $t \in (0, T]$, find $\hat{u}_S(t) \in \mathcal{M}_S^{\mathbb{C}}$ such that:*

$$\mathbb{E}\left[\left\langle i\dot{\hat{u}}_S - \partial_{\hat{u}_S^*} H_\omega(\hat{u}_S, \hat{u}_S^*, \cdot), \hat{v}\right\rangle_h\right] = 0, \qquad \forall \hat{v} \in \mathcal{T}_{\hat{u}_S(t)} \mathcal{M}_S^{\mathbb{C}}. \tag{5.64}$$

with initial condition $u_{0,S}$ given by a suitable S rank approximation of u_0 by e.g. a truncated Karhunen-Loève expansion.

Since $\mathcal{T}_{\hat{u}_S(t)} \mathcal{M}_S^{\mathbb{C}}$ is a complex linear space (which means that $\hat{v} \in \mathcal{T}_{\hat{u}_S(t)} \mathcal{M}_S^{\mathbb{C}}$ if and only if $i\hat{v} \in \mathcal{T}_{\hat{u}_S(t)} \mathcal{M}_S^{\mathbb{C}}$), we get the same conditions if we take only the real part or the imaginary part of (5.64). Following the discussion of Section 5.6.1, and in particular by means of (5.46) and Lemma 5.6.1, we can recast problem (5.64) in the real setting as follows:

Symplectic DLR Variational Principle. *At each $t \in (0, T]$, find $\mathbf{u}_S(t) \in \mathcal{M}_S^{sym}$ such that:*

$$\mathbb{E}\left[\left\langle \mathcal{J}_2 \dot{\mathbf{u}}_S + \nabla H_\omega(\mathbf{u}_S, \cdot), \mathbf{v}\right\rangle\right] = 0, \qquad \forall \mathbf{v} \in \mathcal{T}_{\mathbf{u}_S(t)} \mathcal{M}_S^{sym}, \tag{5.65}$$

with initial conditions given by the symplectic projection of the initial data into $\mathcal{M}_S^{sym}$.

The term $\mathbb{E}\left[\left\langle \nabla H_\omega(\mathbf{u}_S, \cdot), \mathbf{v}\right\rangle\right]$ in (5.65) is interpreted as $\frac{d}{dt}_{|t=0} \mathbb{E}[H_\omega(\boldsymbol{\gamma}_S(t))]$, i.e. the directional derivative along a curve $\boldsymbol{\gamma}_S(t) \in \mathcal{M}_S^{sym}$ with $\boldsymbol{\gamma}_S(0) = \mathbf{u}_S$ and $\dot{\boldsymbol{\gamma}}_S(0) = \mathbf{v}$.

Observe that the variational principle (5.65) corresponds to a symplectic projection of the governing equation onto the (time-dependent) tangent space to the manifold along the trajectory of the approximate solution. We call symplectic dynamical low rank (or symplectic DO) approximation of problem (5.19) the solution to (5.65). This belongs to $\mathcal{M}_S^{sym}$ at any t and is written as:

$$\mathbf{u}_S(\mathbf{x}, t, \omega) = \begin{pmatrix} q_S(\mathbf{x}, t, \omega) \\ p_S(\mathbf{x}, t, \omega) \end{pmatrix} = \sum_{i=1}^{2S} \mathbf{U}_i(\mathbf{x}, t) Y_i(\omega, t) = \begin{pmatrix} \sum_{i=1}^{S} Q_i Y_i - \sum_{i=1}^{S} P_i Y_{S+i} \\ \sum_{i=1}^{S} P_i Y_i + \sum_{i=1}^{S} Q_i Y_{S+i} \end{pmatrix}, \tag{5.66}$$

with $\mathbf{U}(t) \in U(S, [H^1(D)]^2)$, $\mathbf{Y}(t) \in \mathrm{B}^{sym}(2S, L^2(\Omega))$. The peculiarity of the symplectic dynamical low rank approximation is the conservation of energy:

Lemma 5.6.4. *Assuming that problem (5.65) admits a smooth solution $\mathbf{u}_S$, for all $t \in [0, T]$, the expected value of the Hamiltonian is conserved along the approximate solution.*

Proof. Equation (5.65) can be rewritten as

$$\vartheta(\dot{\mathbf{u}}_S, \mathbf{v}) = -\mathbb{E}\left[\left\langle \nabla H_\omega(\mathbf{u}_S, \cdot), \mathbf{v}\right\rangle\right].$$

where ϑ is the symplectic form defined in (5.21). Then, taking $\mathbf{v} = \dot{\mathbf{u}}_S$ we get:

$$0 = \vartheta(\dot{\mathbf{u}}_S, \dot{\mathbf{u}}_S) = -\mathbb{E}\left[\left\langle \nabla H_\omega(\mathbf{u}_S, \cdot), \dot{\mathbf{u}}_S\right\rangle\right] = \frac{d}{dt}_{|t=0} \mathbb{E}[H_\omega(\boldsymbol{u}_S(t))],$$

which implies $\mathbb{E}[H_\omega(\boldsymbol{u}_S(t))] = \mathbb{E}[H_\omega(\boldsymbol{u}_S(0))]$ for all $t \in [0, T]$. $\qquad\square$

Similarly to the Symplectic Order reduction, the energy that is conserved by the approximate solution, i.e. $H(\boldsymbol{u}_S(t)) = \mathbb{E}[H_\omega(\boldsymbol{u}_S(t))]$, is not necessary equal to the energy of the exact solution $H(\boldsymbol{u}(t)) = \mathbb{E}[H_\omega(\boldsymbol{u}(t))]$. However, such discrepancy is constant in time and depends only the approximation of the initial data:

$$|\mathbb{E}[H_\omega(\boldsymbol{u}(t)) - H_\omega(\boldsymbol{u}_S(t))]| = |\mathbb{E}[H_\omega(\boldsymbol{u}(0))] - \mathbb{E}[H_\omega(\boldsymbol{u}_S(0))]|.$$

Moreover, thanks to the analogy with the complex DO, the Symplectic Dynamical Low-rank approximation has the same approximation properties as the standard DO approach. In particular, if the differential operator is linear and deterministic, i.e. $\nabla H = L$ with L deterministic, linear and self-adjoint, the following holds:

Proposition 5.6.8. *The symplectic dynamical low-rank approximation of linear deterministic Hamiltonian systems with random initial condition coincides with the exact solution, provided that the initial condition belongs to $\mathcal{M}_S^{sym}$.*

More generally, the symplectic dynamical low-rank approximation of linear deterministic Hamiltonian systems is optimal in L^2-sense provided that there is no crossing between the omitted and not omitted singular values of the exact solution. As discussed in Chapter 3 (see also [96]), this condition is an intrinsic limitation of dynamical low rank methods, and generally can not be avoided without data-driven adaptivity strategies or closure models: when such crossings occur, the neglected modes, which become dominant in the exact solution, can not be tracked by the reduced system, which evolves only the modes that were dominant at initial time.

By using the parametrization of the tangent space in (5.54) we finally derive the symplectic DO reduced system. The variational problem (5.65) is rewritten in terms of dynamic equations for $(\mathbf{Y}, \mathbf{U})$ as follows:

Proposition 5.6.9. *Let $(\mathbf{U}(t), \mathbf{Y}(t)) \in U(S, [H^1(D)]^2) \times \mathrm{B}^{sym}(2S, L^2(\Omega))$ be a solution of the following system:*

$$\begin{cases} \dot{\mathbf{Y}} = \langle \mathcal{J}_2 \nabla_{\mathbf{u}_S} H_\omega(\mathbf{u}_S), \mathbf{U}^+ \rangle = \mathbf{J}_{2S} \nabla_{\mathbf{Y}} \tilde{H}_\omega(\mathbf{Y}) & \text{(5.67a)} \\[2mm] \begin{aligned} \dot{\mathbf{U}}(\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S}) &= \mathcal{P}_{\mathbf{U}^+}^\perp \left[\nabla H(\mathbf{u}_S) \mathbf{Y}^T \mathbf{J}_{2S} + \mathcal{J}_2 \nabla H(\mathbf{u}_S) \mathbf{Y}^T \right] \\ &= \mathcal{P}_{\mathbf{U}^+}^\perp \left[\mathbb{E}[\nabla H_\omega(\mathbf{u}_S) \mathbf{Y}^T \mathbf{J}_{2S}] + \mathbb{E}[\mathcal{J}_2 \nabla H_\omega(\mathbf{u}_S) \mathbf{Y}^T] \right] \end{aligned} & \text{(5.67b)} \end{cases}$$

with initial conditions given by the complex SVD. Then $\mathbf{u}_S(t) = \mathbf{U}(t)\mathbf{Y}(t) \in \mathcal{M}_S^{sym}$ satisfies the DO variational principle (5.65) at any $t \in [0, T]$.

Proof. The equations in (5.67a)-(5.67b) can be simply obtained by replacing $\tilde{\mathbf{v}}$ with $\mathcal{J}_2 \nabla H(\mathbf{u}_S)$ in the proof of Lemma 5.6.3. $\qquad\square$

Observe that system (5.67a)-(5.67b) consists of $2S$ random ODEs coupled to $2S$ deterministic PDEs. However, exploiting the unitary structure of $\mathbf{U}$ (5.28), we actually need to solve only S PDEs to completely characterize the deterministic basis at each time. Indeed, the dynamic condition (5.53) preserves at continuous level this unitary structure (5.28), provided that $\mathbf{U} \in U(S, [H^1(D)]^2)$ at initial time. This can be directly verified by looking at the set of equations for $\dot{\mathbf{U}}$ (5.67b). First of all, let $\mathbf{A} = \mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S}$ or $\mathbf{A} = (\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S})^{-1}$, in both cases, it holds: $\mathbf{J}_{2S} \mathbf{A} \mathbf{J}_{2S}^T = \mathbf{J}_{2S}^T \mathbf{A} \mathbf{J}_{2S} = \mathbf{A}$. The analogous property is satisfied by the term on the right hand side of (5.67b):

$$\mathscr{J}_2\Big(\mathscr{P}_{\mathbf{U}^+}^\perp \big[\mathbb{E}[\nabla H_\omega(\mathbf{u}_S)\mathbf{Y}^T\mathbf{J}_{2S}] + \mathbb{E}[\mathscr{J}_2\nabla H_\omega(\mathbf{u}_S)\mathbf{Y}^T]\big]\Big)\mathbf{J}_{2S}^T$$
$$= \Big(\mathscr{P}_{\mathbf{U}^+}^\perp \big[\mathbb{E}[\mathscr{J}_2\nabla H_\omega(\mathbf{u}_S)\mathbf{Y}^T] - \mathbb{E}[\nabla H_\omega(\mathbf{u}_S)\mathbf{Y}^T\mathbf{J}_{2S}^T]\big]\Big)$$

where we use Proposition 5.6.4 and the properties of the Poisson matrix. This implies that the same property is necessarily satisfied by $\dot{\mathbf{U}}$, i.e. $\mathscr{J}_2^T\dot{\mathbf{U}}\mathbf{J}_{2S} = \mathscr{J}_2\dot{\mathbf{U}}\mathbf{J}_{2S}^T = \dot{\mathbf{U}}$ and the structure (5.28) is preserved by the dynamic system. On the other hand at discrete level the time discretization scheme has to be carefully chosen to preserve the unitary structure of $\mathbf{U}$.

5.6.3   Isolating the mean

In our context of partial differential equations with random parameters, since we are usually interested in computing the statistics of the solution, it may be worth approximating separately the mean of the solution, as proposed by [116] and adopted in [117],[96],[30] for the DO approximation of parabolic equations. For this aim we re-define S rank random field as follows.

Definition 5.6.5. *We call S rank random field (in the isolated mean format) any function that can be exactly expressed as* $\mathbf{u}_S = \bar{\mathbf{u}}_S + \mathbf{U}\mathbf{Y}$, *where:*

- $\bar{\mathbf{u}}_S = \mathbb{E}[\mathbf{u}_S] \in [H^1(D)]^2 \otimes L^2(D)$.

- $\mathbf{U} \in U(S, [H^1(D)]^2)$,

- $\mathbf{Y} = (Y_1, ..., Y_{2S}) \in \mathrm{B}^{sym}(2S, L^2(\Omega))$ *such that* $\mathbb{E}[Y_i] = 0$ *for any* $i = 1, ..., S$.

We define $\mathring{\mathcal{M}}_S^{sym} \subset (H^1(D) \times L^2(D)) \otimes L_0^2(\Omega)$ *the manifold of all symplectic S rank random fields with zero mean.*

In this setting, the symplectic Low Rank approximation of problem (5.3) is sought in $([H^1(D)]^2 \otimes L^2(D)) \times \mathring{\mathcal{M}}_S^{sym}$ and satisfies:

$$\begin{cases} \dot{\mathbf{Y}} = \mathbf{J}_{2S}\nabla_{\mathbf{Y}}\tilde{H}_\omega^\circ(\mathbf{Y}) \\ \dot{\bar{\mathbf{u}}}_S = \mathbb{E}[\mathscr{J}_2\nabla H_\omega(\mathbf{u}_S)] \\ \dot{\mathbf{U}}(\mathbf{C} + \mathbf{J}_{2S}^T\mathbf{C}\mathbf{J}_{2S}) = \mathscr{P}_{\mathbf{U}^+}^\perp\big[\mathbb{E}[\nabla H_\omega^\circ(\mathbf{u}_S)\mathbf{Y}^T\mathbf{J}_{2S}] + \mathbb{E}[\mathscr{J}_2\nabla H_\omega^\circ(\mathbf{u}_S)\mathbf{Y}^T]\big] \end{cases} \tag{5.68}$$

where $H_\omega^\circ(\cdot) = H_\omega(\cdot) - \mathbb{E}[H_\omega]$ and $\tilde{H}_\omega^\circ = \tilde{H}_\omega \circ \mathbf{U}$.

5.7 Numerical tests

5.7.1 Linear Deterministic Hamiltonian: validation 1

For the validation of the Symplectic DO method we consider the following straightforward problem in the one dimensional domain $D = (0, 2\pi)$:

$$\begin{cases} \ddot{q}(x, t, \omega) = \Delta q(x, t, \omega) & x \in (0, 2\pi), \omega \in \Omega, t \in (0, T] \\ q(0, t, \omega) = q(2\pi, t, \omega) = 0 & \omega \in \Omega, t \in (0, T] \\ q(x, 0, \omega) = Z(\omega) \frac{1}{\sqrt{\pi}} \sin(x) & x \in (0, 2\pi), \omega \in \Omega \\ \dot{q}(x, 0, \omega) = 0 & x \in (0, 2\pi), \omega \in \Omega \end{cases} \quad (5.69)$$

where Z is a uniformly distributed random variable in $[-1, 1]$. The analytical solution, given by:

$$q(x, t, \omega) = Z(\omega) \cos(t) \frac{1}{\sqrt{\pi}} \sin(x), \qquad p(x, t, \omega) = -Z(\omega) \sin(t) \frac{1}{\sqrt{\pi}} \sin(x)$$

is clearly a $1-$rank symplectic function, namely $\mathbf{u} = (q, p)$ belongs to $\mathcal{M}_1^{sym}$ and can be written as $\mathbf{u} = \mathbf{UY}$ with:

$$\mathbf{U} = \frac{1}{\sqrt{\pi}} \begin{bmatrix} \sin(x) & 0 \\ 0 & \sin(x) \end{bmatrix} \in U(1, [H_0^1(D)]^2), \quad \mathbf{Y} = \begin{bmatrix} Z(\omega) \cos(t) \\ -Z(\omega) \sin(t) \end{bmatrix} \in B^{sym}(2, L^2(\Omega)).$$

In particular, this means that the rank of the exact solution, which is equal to 1 at $t = 0$, remains constant in time. The same generally applies to any solution of linear deterministic Hamiltonian systems with finite rank initial condition. We start by rewriting problem (5.69) in Hamiltonian form:

$$\begin{cases} \dot{\mathbf{u}}(x, t, \omega) = \mathcal{J}_2 \mathbf{L} \mathbf{u}(x, t, \omega) \\ \mathbf{u}(x, 0, \omega) = (Z(\omega) \frac{1}{\sqrt{\pi}} \sin(x), 0) & \text{with} \quad \mathbf{L} = \begin{bmatrix} -\Delta & 0 \\ 0 & \mathbb{I} \end{bmatrix}. \\ u_1(0, t, \omega) = u_1(2\pi, t, \omega) = 0 \end{cases} \quad (5.70)$$

Then by following (5.67a)-(5.67b), one can easily derive the reduced Symplectic DO system, which is given by:

$$\begin{cases} \dot{\mathbf{Y}}(t, \omega) = < \mathcal{J}_2 \mathbf{L} \mathbf{U}(\cdot, t), \mathbf{U}(\cdot, t) > \mathbf{Y}(t, \omega) & \omega \in \Omega, t \in (0, T] \\ \dot{\mathbf{U}}(t)(\mathbf{C}(t) + \mathbf{J}_2^T \mathbf{C}(t) \mathbf{J}_2) = \mathcal{P}_{\mathbf{U}(t)}^\perp [\mathcal{J}_2 \mathbf{L} \mathbf{U}(t) \mathbf{C}(t) + \mathbf{L} \mathbf{U}(t) \mathbf{C}(t) \mathbf{J}_2] & x \in (0, 2\pi), t \in (0, T] \end{cases} \quad (5.71)$$

with initial conditions:

$$\mathbf{U}(0) = \begin{bmatrix} Q(0) & -P(0) \\ P(0) & Q(0) \end{bmatrix} = \frac{1}{\sqrt{\pi}} \begin{bmatrix} \sin(x) & 0 \\ 0 & \sin(x) \end{bmatrix}, \quad \mathbf{Y}(0) = \begin{bmatrix} Z(\omega) \\ 0 \end{bmatrix} \quad (5.72)$$

and completed with homogeneous Dirichlet boundary conditions: $Q(0, t) = Q(2\pi, t) = P(0, t) = P(2\pi, t) = 0$ for all $t \in [0, T]$. After observing that $\mathbf{U}(0)$ is an eigenfunction of $\mathbf{L}$ with eigenvalue
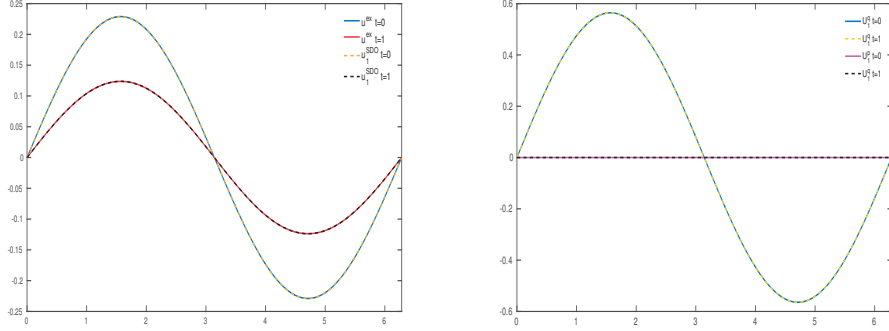
Figure 5.1 – Left: the exact solution (solid line) and the symplectic DO approximate solution with $S = 1$ (dotted line) for $Z = 0.4058$ at $t = 0$ and $t = 1$: the two solutions coincide. Right: the deterministic modes of the symplectic DO approximate solution with $S = 1$ at $t = 0$ and $t = 1$: the modes are constant in time. Discretization parameters: number of Gauss-Legendre collocation points $N_y = 7$, spatial discretization $h = 0.01$, time-step $\Delta t = 0.01$.

equal to 1, i.e. $\mathbf{LU}(0) = \mathbf{U}(0)$, we claim that the Symplectic DO system (5.71) recovers the exact solution of problem (5.69) and keep the deterministic modes constant in time. Namely we want to show that the exact solution, written as $\mathbf{u}(t) = \mathbf{U}(t)\mathbf{Y}(t)$ with $\mathbf{U}(t) = \mathbf{U}(0)$ and $\mathbf{Y}(t) = (Z(\omega)\cos(t), -Z(\omega)\sin(t))^T$, satisfies (5.71). To verify this, we start by assuming that $\mathbf{C} + \mathbf{J}_2^T\mathbf{CJ}_2$ has full rank, with $\mathbf{C}$ denoting the moments matrix, i.e. $\mathbb{E}[\mathbf{YY}^T]$. Under this assumption, one can easily see that equations (5.71) are automatically satisfied by $\mathbf{U}(t) = \mathbf{U}(0)$, by observing that

$$
\begin{aligned}
0 = \dot{\mathbf{U}}(\mathbf{C} + \mathbf{J}_2^T\mathbf{CJ}_2) \quad &= \mathscr{P}_{\mathbf{U}}^{\perp}\left[\mathscr{J}_2\mathbf{UC} - \mathbf{UCJ}_2^T\right] \\
&= \mathscr{P}_{\mathbf{U}}^{\perp}[\mathbf{U}](\mathbf{J}_2\mathbf{C} - \mathbf{CJ}_2^T) \\
&= 0
\end{aligned}
$$

since $\mathscr{P}_{\mathbf{U}}^{\perp}[\mathbf{U}]$ is clearly equal to zero. Thus, the Symplectic DO system, which is reduced to the Hamiltonian system for the evolution of the coefficients $\mathbf{Y}$, degenerates to the proper symplectic decomposition proposed in [107]. Specifically we have $\dot{\mathbf{Y}} = \mathbf{J}_2\mathbf{Y}$ with initial condition $\mathbf{Y}(0)$, which admits a unique solution given by $\mathbf{Y}(t) = (Z(\omega)\cos(t), -Z(\omega)\sin(t))^T$. We finally verify that the assumption on the rank of $\mathbf{C} + \mathbf{J}_2^T\mathbf{CJ}_2$ is actually fulfilled, by observing that $\mathbf{C} + \mathbf{J}_2^T\mathbf{CJ}_2 = \begin{pmatrix} \mathbb{E}[Z^2] & 0 \\ 0 & \mathbb{E}[Z^2] \end{pmatrix}$ at any time. This allows us to conclude that the Symplectic DO method recovers the exact solution by keeping the deterministic basis constant in time. The numerical results perfectly agree with the previous analysis, with the only care in choosing a symplectic time discretization scheme, see Figure 5.1.
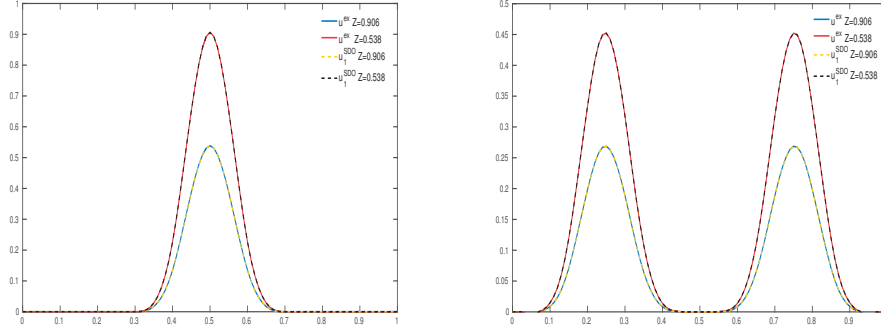
Figure 5.2 – The solution for two different realizations of Z,i.e. $Z = 0.906$ and $Z = 0.538$, at $t = 0$ on the left and $t = 0.8$ on the right. The symplectic DO solution coincides with the reference solution computed with the Stochastic collocation method. Discretization parameters: number of Gauss-Legendre collocation points $N_y = 5$, spatial discretization $h = 0.01$, time-step $\Delta t = 0.01$.

5.7.2 Linear Deterministic Hamiltonian: validation 2

Next, we consider again a linear wave equation but with a more general initial condition:

$$\begin{cases} \ddot{q}(x, t, \omega) = 0.1\Delta q(x, t, \omega) & x \in (0, 1), \omega \in \Omega, t \in (0, T] \\ q(0, t, \omega) = q(1, t, \omega) = 0 & \omega \in \Omega, t \in (0, T] \\ q(x, 0, \omega) = Z(\omega)h(10 \times |x - 0.5|) & x \in (0, 1), \omega \in \Omega \\ \dot{q}(x, 0, \omega) = 0 & x \in (0, 1), \omega \in \Omega \end{cases} \tag{5.73}$$

with:

$$h(s) = \begin{cases} 1 - 1.5s^2 + 0.75s^3 & 0 \geq s \leq 1 \\ 0.25(2 - s)^3 & 1 < s \leq 2 \\ 0 & s > 2 \end{cases}$$

Since the Hamiltonian is linear and deterministic, the exact solution, which at time $t = 0$ is a symplectic 1-rank function, has rank which is constant in time and can be written as $\mathbf{u}(x, t, \omega) = \mathbf{Z}(\omega)(q(x, t), p(x, t))$. By observing that $\mathscr{J}_2\mathbf{Lu}$ and $\mathbf{Lu}$ belong to the tangent space $\mathscr{T}_{u(t)}\mathscr{M}_1^{sym}$ at any time, we claim that the Symplectic DO method recovers again the exact solution. In particular, the Symplectic DO approximate solution, which is initialized as:

$$\mathbf{U}_0 = \begin{bmatrix} \frac{h(10 \times |x-0.5|)}{\|h(10 \times |x-0.5|)\|} & 0 \\ 0 & \frac{h(10 \times |x-0.5|)}{\|h(10 \times |x-0.5|)\|} \end{bmatrix} \quad \mathbf{Y}_0 = \begin{bmatrix} \|h(10 \times |x - 0.5|)\|Z \\ 0 \end{bmatrix},$$

is expected to evolve as $(\mathbf{U}(t), Z\mathbf{X}(t))$, where $\mathbf{X}(t)$ is a rescaling factor, and satisfies $\mathbf{U}(t)\mathbf{Y}(t) = \mathbf{u}$ at any time. The numerical results validate the exactness of the Symplectic DO method for the problem under consideration, up to the numerical discretization error in time and space. The validation is done by comparing the Symplectic DO approximate solution to the reference
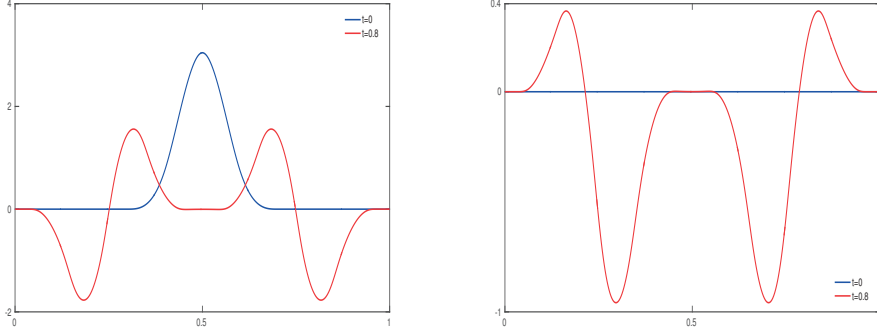
Figure 5.3 – The deterministic modes Q (left) and P (right) of the symplectic DO approximate solution at time $t = 0$ and $t = 0.8$. We observe that both the modes evolve in time by following the variability spread of the solutions. Discretization parameters: number of Gauss-Legendre collocation points $N_y = 5$, spatial discretization $h = 0.01$, timestep $\Delta t = 0.01$.

solution computed with the Stochastic Collocation method with Gauss-Legendre points ([5]). Figure 5.2 shows the solution for two different realizations of Z and at two different times $t = 0$ and $t = 0.6$: we see that the DO solution and the reference solution coincide. Contrary to the previous example (in which the deterministic basis remains fixed in time), Figure 5.3 shows that in this case, the deterministic modes evolve in time by following the wave propagation. In particular, we observe that the mode $\mathbf{P}_1$, initialized to zero, will be automatically activated by the method, which means that the approximation will not be restricted to the diagonal structure of $\mathbf{U}_0$, used for the initialization. This shows the potential of the Symplectic DO method with respect to a reduced order method with fixed (in time) bases.

5.7.3 Wave equation with random wave speed

We now consider a linear wave equation with random speed and random initial data, in the 2-dimensional physical domain $D = (0,1)^2$, with boundary $\partial D = \bar\Gamma_N \cup \bar\Gamma_D$, $\bar\Gamma_N = \{(x,y) \in \mathbb{R}^2, x \in (0,1), y = 1\}$, $\Gamma_D = \partial D \backslash \Gamma_N$. The problem reads as follows:

$$
\begin{cases}
\ddot{q}(\mathbf{x},t,\omega) = c^2(\omega)\Delta q(\mathbf{x},t,\omega) & \mathbf{x} \in (0,1)^2, \ t \in (0,T], \ \omega \in \Omega \\
q(\mathbf{x},t,\omega) = 0 & \mathbf{x} \in \Gamma_D, \ t \in (0,T], \ \omega \in \Omega \\
\partial_n q(\mathbf{x},t,\omega) = 0 & \mathbf{x} \in \Gamma_N, \ t \in (0,T], \ \omega \in \Omega \\
q(\mathbf{x},0,\omega) = \alpha(\omega) q_0(\mathbf{x}) & \mathbf{x} \in (0,1)^2, \ \omega \in \Omega \\
\dot{q}(\mathbf{x},0,\omega) = 0 & \mathbf{x} \in (0,1)^2, \ \omega \in \Omega
\end{cases}
\tag{5.74}
$$

with:

$$
q_0(\mathbf{x}) = \begin{cases}
e^{\frac{-\|\mathbf{x}-0.5\|^2}{2(0.1)^2}} & \|\mathbf{x}-0.5\|^2 < 0.8 \\
0 & \|\mathbf{x}-0.5\|^2 \geq 0.8
\end{cases}
\tag{5.75}
$$

Here the randomness arises form both the diffusion coefficient and the initial data. We assume that the uncertainty in the initial condition is independent from the randomness of the wave speed. The stochastic space is parametrized in terms of 2 independent random variables Z_1, Z_2, affecting respectively the initial position $q(0)$ and the diffusion coefficient. Specifically we assume $\alpha = (Z_1 + 0.1)^2$ and $c^2 = 0.1 + 0.05 Z_2$ with Z_1, Z_2 linearly independent and uniformly distributed in $[-1, 1]$. The goal here is to test the symplectic DO method on a problem in which the probability distribution (and consequently the rank) of the exact solution changes over time. We start by rewriting problem (5.74) in Hamiltonian form:

$$\begin{cases} \dot{\mathbf{u}}(\mathbf{x}, t, \omega) = \mathscr{J}_2 \mathbf{L}(\omega) \mathbf{u}(\mathbf{x}, t, \omega) \\ \mathbf{u}(\mathbf{x}, 0, \omega) = \big((Z_1(\omega) + 0.1)^2 q_0(\mathbf{x}), 0 \big) \end{cases} \quad \text{with} \quad \mathbf{L} = \begin{bmatrix} -c^2(\omega)\Delta & 0 \\ 0 & \mathbb{I} \end{bmatrix}, \tag{5.76}$$

Observe that the Hamiltonian explicitly depends on the random variable c^2:

$$H_\omega(q, p, \omega) = \frac{1}{2} \int_D (|p|^2 - c(\omega)^2 |\nabla q|^2).$$

We look for an approximate solution $\mathbf{u}_S \in \mathscr{M}_S^{sym}$ written as:

$$\mathbf{u}_S(\mathbf{x}, t, \omega) = \begin{pmatrix} q_S(\mathbf{x}, t, \omega) \\ p_S(\mathbf{x}, t, \omega) \end{pmatrix} = \sum_{i=1}^{2S} \mathbf{U}_i(\mathbf{x}, t) Y_i(\omega, t) = \begin{pmatrix} \sum_{i=1}^{2S} Q_i(\mathbf{x}, t) Y_i(t, \omega) - \sum_{i=1}^{2S} P_i(\mathbf{x}, t) Y_i(t, \omega) \\ \sum_{i=1}^{2S} P_i(\mathbf{x}, t) Y_i(t, \omega) + \sum_{i=1}^{2S} Q_i(\mathbf{x}, t) Y_i(t, \omega) \end{pmatrix}, \tag{5.77}$$

which satisfies

$$\begin{cases} \dot{\mathbf{Y}} = <\mathscr{J}_2 \mathbf{L} \mathbf{U}, \mathbf{U}> \mathbf{Y} \tag{5.78a} \\ \dot{\mathbf{U}}(\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S}) = \mathscr{P}_{\mathbf{U}}^\perp \big[\mathscr{J}_2 \mathbb{E}[\mathbf{L}\mathbf{U}\mathbf{Y}\mathbf{Y}^T] - \mathbb{E}[\mathbf{L}\mathbf{U}\mathbf{Y}\mathbf{Y}^T]\mathbf{J}_{2S}^T \big] \tag{5.78b} \end{cases}$$

at any time and for some $S \geq 1$. Despite the initial condition is a 1-rank function, the rank of the exact solution is expected to increase in time. Indeed, even if the governing equation is linear, the parameters-to-solution maps is non-linear, due the randomness which affects the differential operator. Thus it is reasonable to assume that the Symplectic DO approximation needs $S > 1$ modes to achieve good levels of accuracy.

We look for a Symplectic DO approximate solution in $\mathscr{M}_S^{sym}$ for $S > 1$, and for the initialization of the modes we adopt the same strategy used in [96, 95]; namely the deterministic modes are initialized randomly and the redundant stochastic coefficients are set to zero. Precisely, after setting $\tilde{Q}_1 = q_0$ and $\tilde{Y}_1 = (Z_1 + 0.1)^2$, we initialize $\tilde{Q}_2, ..., \tilde{Q}_S$ randomly with associated stochastic coefficients $\tilde{Y}_2, ..., \tilde{Y}_S$ equal to zero. Then, in order to get a symplectic orthogonal basis, we factorize $\tilde{\mathbf{Q}}$, by using the (real) QR factorization, in $\tilde{\mathbf{Q}} = \mathbf{Q}\mathbf{R}$ and we initialize:

$$\mathbf{U} = \begin{bmatrix} \mathbf{Q} & 0 \\ 0 & \mathbf{Q} \end{bmatrix}, \quad \mathbf{Y}_i = \sum_{j=1}^S \mathbf{R}_{ij} \tilde{Y}_j \text{ and } \tilde{Y}_{S+i} = 0, \forall i = 1, ..., S.$$

Roughly speaking, we use a number of modes larger then what needed to approximate the initial data (although the approximate solution thus constructed has deficient rank), but we evolve in time only the "active" modes (possibly after a suitable rotation of the basis), i.e. those corresponding to non vanishing singular values. The problem of dealing with approximate solutions with deficient rank is however an issue which generally affects the dynamically low rank approximation with fixed rank, at initial and successive time. To deal with it, we implemented two alternative strategies: the first one simply consists in multiplying both sides of (5.78b) by the pseudo inverse of $(\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S})$; the second is based on the complex diagonalization of $(\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S})$. Detailing more the second strategy, let $\tilde{\mathbf{C}}$ denotes the sum $(\mathbf{C} + \mathbf{J}_{2S}^T \mathbf{C} \mathbf{J}_{2S})$. Observe that $\tilde{\mathbf{C}}$ satisfies $\tilde{\mathbf{C}} = \mathbf{J}_{2S}^T \tilde{\mathbf{C}} \mathbf{J}_{2S}$, so it can be written as: $\tilde{\mathbf{C}} = \begin{bmatrix} \tilde{\mathbf{C}}_1 & -\tilde{\mathbf{C}}_2 \\ \tilde{\mathbf{C}}_2 & \tilde{\mathbf{C}}_1 \end{bmatrix}$, with $\tilde{\mathbf{C}}_1, \tilde{\mathbf{C}}_2 \in \mathbb{R}^{S \times S}$. This means that $\tilde{\mathbf{C}}$ can be identified by the complex hermitian matrix $\hat{\mathbf{C}} = \tilde{\mathbf{C}}_1 + i\tilde{\mathbf{C}}_2 \in \mathbb{C}^{S \times S}$. Let $\hat{\mathbf{D}}, \hat{\mathbf{V}}$ be respectively the (complex) eigenvalues and eigenvectors of $\hat{\mathbf{C}}$, and $\mathbf{V}$ the real matrix representation of $\hat{\mathbf{V}}$, i.e. $\mathbf{V} = \begin{bmatrix} \mathrm{Re}(\hat{\mathbf{V}}) & -\mathrm{Im}(\hat{\mathbf{V}}) \\ \mathrm{Im}(\hat{\mathbf{V}}) & \mathrm{Re}(\hat{\mathbf{V}}) \end{bmatrix}$. We define $\tilde{\mathbf{U}}_i = (\tilde{U}_i^Q, \tilde{U}_i^P)^T = \sum_{j=1}^{2S} \mathbf{U}_j \mathbf{V}_{ji}$ and we rewrite equations (5.78b) with respect to the rotated basis $\tilde{\mathbf{U}}$. Observe that the complex diagonalization guarantees that the rotated basis $\tilde{\mathbf{U}}$ belongs to $U(S, [H^1(D)]^2)$, since the product of symplectic orthogonal matrices is as well symplectic orthogonal. Then we actually solve only the equations corresponding to not vanish eigenvalues, i.e. the equations in $\dot{\mathbf{U}}_i$ for which $\mathbf{D}_{ii}$ (which is real) is larger then a prescribed tolerance, for any $i = 1, ..., S$. Denoting by r the rank of $\mathbf{D}$, the remaining modes $\tilde{\mathbf{U}}_{r+1}, ..., \tilde{\mathbf{U}}_S$ are kept constant to the previous time iteration. Finally, by exploiting the unitary structure in (5.28), we reconstruct the complete basis as:

$$\tilde{\mathbf{U}} = \begin{bmatrix} \tilde{\mathbf{U}}^Q & -\tilde{\mathbf{U}}^P \\ \tilde{\mathbf{U}}^P & \tilde{\mathbf{U}}^Q \end{bmatrix} \tag{5.79}$$

and we get the updated modes in the original coordinates by multiplying by $\mathbf{V}^T$. Despite the two strategies lead to comparable numerical results, the technique based on the complex diagonalization, has the computational advantage of solving the minimum number of equations required. In practice, in the results reported here, the rank r is computed with respect to a threshold ϵ that is weighted by the largest eigenvalue of $\mathbf{D}$ at each time, specifically we set threshold equal to $\epsilon = 10^{-15} \max_{i=1,...,S} D_{ii}^n$ at any $t^n = n\Delta t$.

5.7.4 Numerical Discretization

The implementation of all numerical tests in this Chapter has been developed within the open source Finite Element library FEniCs [3]. The Finite Element method is used for the discretization in the physical space, namely for solving (5.78b) and for computing the $L^2(D)$-projection in (5.78a). Specifically we use $P1$ finite elements on a uniform triangular grid of equal edges $h = 0.04$. For what concerns the discretization of the random modes, we parametrize the stochastic space in terms of a uniformly distributed random vector $\boldsymbol{\eta}$, in accordance with

the distribution of the input random data. Thus the stochastic space $(\Omega, \mathscr{A}, \mathscr{P})$ is replaced by $(\Lambda, \mathscr{B}(\Lambda), f(\boldsymbol{\eta}) d\boldsymbol{\eta})$ where here $\Lambda = [-1, 1]^2$, $\mathscr{B}(\Lambda)$ and $f = \frac{1}{4}$ denote respectively the domain, the Borel σ-algebra and the density function of $\boldsymbol{\eta}$. Then, equations (5.78a) are solved with the Stochastic Collocation method on Gauss-Legendre collocation points with tensorized Gaussian grid [5]. The corresponding quadrature formula is used to compute the covariance matrix and any expected value in (5.78b). However, the use of *sparse* stochastic collocation grids is recommended for problems in higher dimensional stochastic spaces. For details see e.g. [135, 5, 101].

The time discretization scheme has to be carefully chosen in order to preserve the symplectic structure of the problem. For a complete review of symplectic schemes we refer to [56] and references therein. Moreover, since numerical symplectic schemes do not necessarily preserve the orthogonal structure (5.28) at the discrete level, especially for approximate solutions with deficient rank, special attention has been paid to preserve both the orthogonal and symplectic structure of the deterministic modes. We propose two possible time discretization strategies, described hereafter, both finalized to preserve the symplecticity of the flow and guarantee the orthogonality of the deterministic basis. Based on the linear reversibility of wave equations, which states that the time reversed solution of a wave equation is also solution to the same wave equation, we look for a numerical scheme which, when applied to a reversible differential equation, produces a reversible numerical flow, in order to get a consistent long-time behavior. Based on the link between reversibility and symmetric schemes [56], we propose two possible symplectic time discretization methods based respectively on a symmetric splitting and on the implicit midpoint rule (which is a symmetric scheme). The two procedures can be summarized as follows:

- Strang splitting in $\mathbf{U}, \mathbf{Y}$ combined with the symplectic Euler scheme. Starting from $\mathbf{u}_S^n = \mathbf{U}^n \mathbf{Y}^n$ at $t = t^n$:

    – we compute $\mathbf{Y}^{n+1/2} \approx \mathbf{Y}(t^n + \frac{\Delta t}{2})$ by solving system (5.78a) discretized in time with the Symplectic Euler scheme for half time step;

    – we compute $\mathbf{U}^{n+1} \approx \mathbf{U}(t^n + \Delta t)$ by solving system (5.78b) with the Symplectic Euler scheme and the updated coefficients $\mathbf{Y}^{n+1/2}$;

    – we re-orthogonalize $\mathbf{U}^{n+1}$ by using the complex QR factorization;

    – we compute $\mathbf{Y}^{n+1} \approx \mathbf{Y}(t^n + \Delta t)$ by solving system (5.78a) for half time step, with initial values $\mathbf{Y}^{n+1/2}$ and the updated deterministic basis. The equations are discretized by the adjoint Symplectic Euler scheme with respect to the one used in the first half-step.

- Standard Lie-Trotter splitting in $\mathbf{U}, \mathbf{Y}$ combined with the implicit midpoint scheme for the time discretization of both system (5.78a) and system (5.78b). We apply the complex diagonalization strategy to (5.78b) and we denote by $\mathbf{u}_S^n = \mathbf{U}^n \mathbf{Y}^n = \tilde{\mathbf{U}}^n \tilde{\mathbf{Y}}^n$ the approximate solution at time $t^n = n\Delta t$ in standard and rotated bases respectively. Equations (5.78a)-
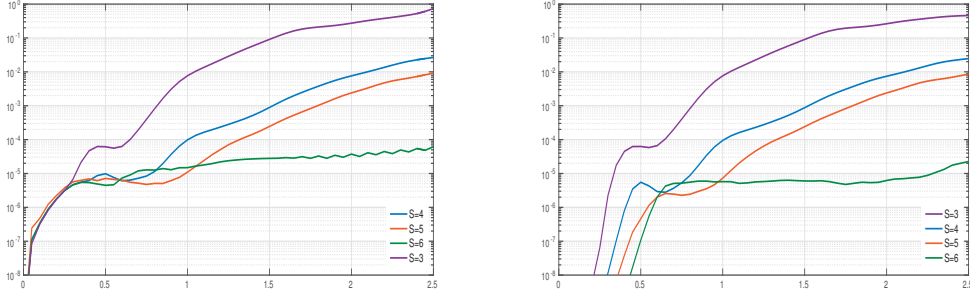
Figure 5.4 – Evolution in time of the approximation error of the Symplectic DO method with different number of modes (S=3,4,5,6). The error is computed in norm $H^1(D) \otimes L^2(\Omega)$ with respect to a reference solution computed with the Stochastic Collocation method. On the left the approximation error with the Strang splitting combined with the symplectic Euler scheme. On the right the Lie-Trotter splitting combined with the implicit midpoint scheme (right). Discretization parameters: stochastic tensor grid with Gauss-Legendre collocation points, number of points: $N_y = 49$, spatial discretization: triangular mesh with edge $h = 0.04$, uniform time-step $\Delta t = 0.001$.

(5.78b) are discretized in time as follows:

$$\frac{1}{\Delta t}\mathbf{Y}^{n+1} - \frac{1}{2} < \mathcal{J}_2\mathbf{LU}^n, \mathbf{U}^n > \mathbf{Y}^{n+1} = \frac{1}{\Delta t}\mathbf{Y}^n + \frac{1}{2} < \mathcal{J}_2\mathbf{LU}^n, \mathbf{U}^n > \mathbf{Y}^n$$

$$
\begin{aligned}
\frac{1}{\Delta t}\tilde{\mathbf{U}}_i^{n+1}\hat{\mathbf{D}}_{ii}^{n+1} - \frac{1}{2}\mathscr{P}_{\mathbf{U}^n}^{\perp}\big[\,\mathcal{J}_2\mathbb{E}[\mathbf{L}\tilde{\mathbf{U}}_i^{n+1}\tilde{\mathbf{Y}}_i^{n+1}\tilde{\mathbf{Y}}_i^{n+1}] + \mathbb{E}[\mathbf{L}\tilde{\mathbf{U}}_i^{n+1}\tilde{\mathbf{Y}}_i^{n+1}\tilde{\mathbf{Y}}_{S+i}^{n+1}]\big] \quad &= \frac{1}{\Delta t}\tilde{\mathbf{U}}_i^n\hat{\mathbf{D}}_{ii}^{n+1}\\
&+ \frac{1}{2}\mathscr{P}_{\mathbf{U}^n}^{\perp}\big[\,\mathcal{J}_2\mathbb{E}[\mathbf{L}\tilde{\mathbf{U}}_i^n\tilde{\mathbf{Y}}_i^n\tilde{\mathbf{Y}}_i^{n+1}] + \mathbb{E}[\mathbf{L}\tilde{\mathbf{U}}_i^n\tilde{\mathbf{Y}}_i^{n+1}\tilde{\mathbf{Y}}_{S+i}^{n+1}]\big]\\
&+ \mathscr{P}_{\mathbf{U}^n}^{\perp}\big[\,\mathcal{J}_2\mathbb{E}[\sum_{\substack{j=1\\j\neq i}}^{2S}\mathbf{L}\tilde{\mathbf{U}}_j^n\tilde{\mathbf{Y}}_j^{n+1}\tilde{\mathbf{Y}}_i^{n+1}] + \mathbb{E}[\sum_{\substack{k=1\\k\neq S+i}}^{2S}\sum_{j=1}^{S}\mathbf{L}\tilde{\mathbf{U}}_j^n\tilde{\mathbf{Y}}_j^{n+1}\tilde{\mathbf{Y}}_k^{n+1}(\mathbf{J}_{2S})_{ki}]\big]
\end{aligned}
$$

$$(5.80)$$

Concerning the re-orthogonalization of the deterministic modes in the second strategy, we recall that the midpoint rule has the convenient property of conserving quadratic invariants and in particular the implicit midpoint scheme is a unitary integrator [41]. We numerically observe that the implicit midpoint scheme helps in preserving the symplectic orthogonal structure of the deterministic basis, thus reducing the number of the (computationally expensive) re-orthogonalizations. However, we emphasize that the midpoint scheme proposed here does not preserve exactly the unitary structure of $\mathbf{U}$ and reorthogonalization is still needed for approximate solutions with deficient rank. In particular, the unitary structure is slightly compromised by the explicit treatment of the coupling terms in (5.80). However for the problem under consideration, this scheme allows us to apply a complex QR re-orthogonalization only in the very first time steps, when the solution has deficient rank, and then around every 100 iterations (one possibility is to apply the complex QR decomposition only when the error in the orthonormalization of $\mathbf{U}$ is larger then a prescribed tolerance). Figure 5.4 shows the approximation error of the Symplectic DO approximate solution, implemented with the Strang splitting combined with Symplectic Euler method on the left, and with the Lie-Trotter splitting

combined with the midpoint scheme on the right, with different numbers of modes. The error is computed in norm $H^1(D) \otimes L^2(\Omega)$ with respect to a reference solution computed with the Stochastic Collocation method on Gauss-Legendre points using 7 points in each direction and the same discretization parameters in time and space, i.e. a triangular mesh with edge $h = 0.04$ and uniform time-step $\Delta t = 0.001$. We observe that good level of accuracy can be reached with only a few modes and in particular, for $S = 6$ the magnitude of error tends to stay constant in time and lower than 10^{-4}. Despite the two strategies lead to comparable numerical results, we point out that the second one is generally computationally more efficient since a smaller number of re-orthogonalizations is required. We conclude by reporting here same qualitative results to show the effectiveness of the Symplectic DO method in reproducing the exact follow of the solutions. In Figure 5.5 and Figure 5.6 we compare the exact and the approximate solution with $S = 5$, evaluated in $\alpha = 1$, $c^2 = 0.121$ and $\alpha = 0.4$, $c^2 = 0.063$ respectively, at different times. One can see that, even if the two realizations (i.e. for $\alpha = 1$, $c^2 = 0.121$ and $\alpha = 0.4$, $c^2 = 0.063$) are quite different at fixed time, the Symplectic DO is able to effectively reproduce both of them.

5.8 Conclusion

In this Chapter, we developed a dynamical low-rank technique for the approximation of wave equations with random parameters, which combines the DLR approach with the use of symplectic deterministic (dynamic) bases. The governing equation is rewritten in the Hamiltonian form in a suitable symplectic space, and the approximate solution is sought in the set of all random fields which can be expanded, in separable form, over a symplectic-orthogonal deterministic basis of dimension $2S$. After deriving the proper conditions on the stochastic coefficients to equip this set, denoted by $\mathcal{M}_S^{sym}$, with a manifold structure, we formulated the Symplectic DLR variational principle as the symplectic projection of the Hamiltonian system onto the tangent space to $\mathcal{M}_S^{sym}$ along the approximate trajectory. We showed that this coincides with rewriting the governing Hamiltonian system in complex variables and looking for a DLR approximation in the manifold $\mathcal{M}_S^{\mathbb{C}}$ of all the complex-valued random fields with rank S. We used the analogy between the complex manifold $\mathcal{M}_S^{\mathbb{C}}$ and its real representation $\mathcal{M}_S^{sym}$ to determine a suitable parametrization of the tangent space to $\mathcal{M}_S^{sym}$ (in real form). After deriving the associated orthogonal constraints on the dynamics of the deterministic modes, we recovered the reduced dynamical system which, in the real framework, consists of a set of equations for the constrained dynamics of the deterministic modes, coupled with a reduced order Hamiltonian system for the evolution of the stochastic coefficients. The Symplectic DO shares with the symplectic order reduction the use of symplectic deterministic bases, and, as the "classic" DO approximation, allows both the stochastic and the deterministic modes to evolve in time. As a result, the approximate solution preserves the (approximated) mean Hamiltonian energy and continuously adapts in time to the structure of the solution.
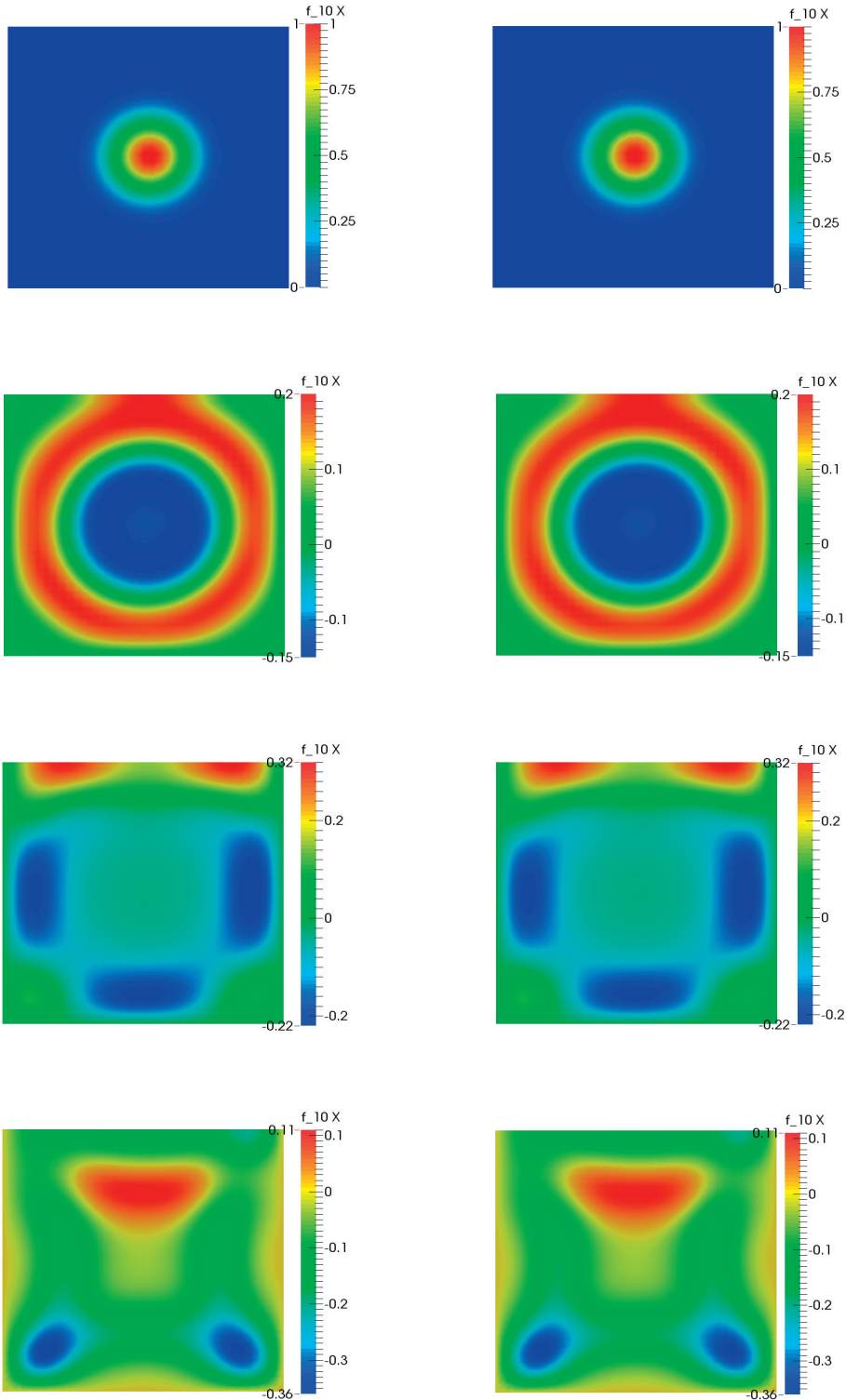
Figure 5.5 – Reference solution (left) and Symplectic DO approximate solution with $S = 5$ (right) for $\alpha = 1$ and $c^2 = 0.121$ at $t = 0$, $t = 1$, $t = 1.5$, and $t = 2$. Discretization parameters: stochastic tensor grid with Gauss-Legendre collocation points, number of points: $N_y = 49$, spatial discretization: $P1$ finite elements over a triangular mesh with edge $h = 0.04$, uniform time-step $\Delta t = 0.001$ (with implicit midpoint scheme).
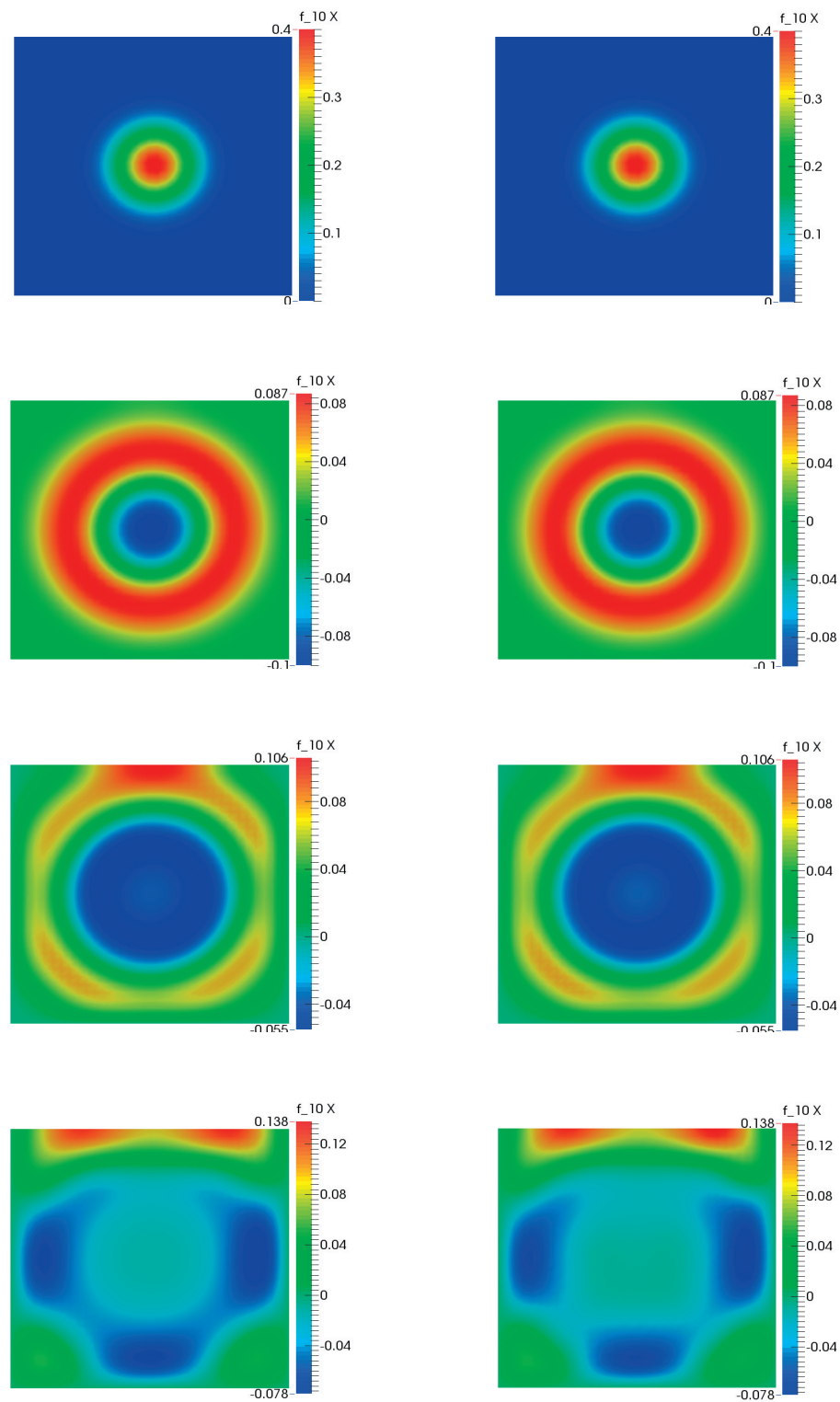
Figure 5.6 – Reference solution (left) and Symplectic DO approximate solution with $S = 5$ (right) for tree different values of $\alpha = 0.4$ and $c^2 = 0.063$ at $t = 0$, $t = 1$, $t = 1.5$, and $t = 2$. Discretization parameters: stochastic tensor grid with Gauss-Legendre collocation points, number of points: $N_y = 49$, spatial discretization: triangular mesh with edge $h = 0.04$, uniform time-step $\Delta t = 0.001$ (with implicit midpoint scheme).

175

6 Conclusions and perspectives

In this thesis, we developed and analyzed dynamical low-rank techniques for the approximation of time evolving PDEs with random data.

The starting point has been to establish a link between the DO method, proposed in [116, 117] for the approximation of PDEs with random initial data, and the MCTDH method used for the approximation of deterministic Schrödinger equations, or its discrete analogue Dynamical Low Rank approximation of evolution matrix or tensor equations [39, 66]. After defining the approximation manifold as the collection of all the S rank random fields, i.e. the set of all functions which can be expressed as a sum of S linearly independent deterministic modes combined with S linearly independent stochastic modes, we formalized the Dynamical Low Rank variational principle for parabolic equations with random parameters, which consists in a Galerkin projection of the governing equations onto the tangent space to the approximation manifold along the approximate trajectory. In this setting the DO reduced system is rederived by means of one of the (many) possible parametrizations of the tangent space. The variational approach proposed here, besides giving a unified formulation for the DO method and other dynamical low-rank techniques such as the DyBO and the DDO method, allows for a suitable mathematical analysis of such approaches. There were, indeed, no previous works in literature addressing this issue in the context of PDEs with random data. In light of the theoretical results developed for the MCTDH method and the Dynamical Low-Rank approximation, we investigated the properties of the manifold $\mathcal{M}_S$ for a linear parabolic equation with random parameters. Specifically, we exploited the curvature bounds of $\mathcal{M}_S$ to derive a theoretical bound for the approximation error of the $S-$terms DO solution in terms of the corresponding S-terms best approximation, i.e. the truncated $S-$terms Karhunen-Loève expansion at each time instant. The bound is applicable for full rank DLR approximate solutions on the largest time interval in which the best $S-$terms approximation is continuously differentiable in time. We observed by means of simple analytical examples that the regularity assumption on the Karhunen-Loève decomposition is actually a necessary condition to maintain an effective control on the DO approximation error. As confirmed by the numerical results, the DO approximation error is properly bounded in terms of best approximation error as long as

the eigenvalues of Karhunen-Loève expansion included in the S rank approximation, do not cross the ones which have been initially omitted. This provides an indication of the effectiveness of the DO method for different types of problems. More challenging appears instead the extension of the error analysis to non linear problems, as for instance Navier-Stokes equations, for which the analysis is complicated by the difficulty in obtaining bounds on the time derivative in the stochastic space. Moreover, the possibility of extending the error analysis to approximate solutions with deficient rank remains an open issue. The problem of the rank-deficiency is as well the main obstacle in the analysis of the existence and uniqueness of the approximate solution. While our theoretical analysis requires the full rank condition, we overcome this problem in the numerical scheme by using a strategy based on the pseudo-inverse operator: we initialize the solution with a large number of modes (although the initial solution thus constructed might be rank deficient), and only the "active" modes (possibly after a suitable rotation of the basis), i.e. those corresponding to non-vanishing singular values, are actually evolved in time. This allows the rank of the approximate solution to adapt in time without losing computational efficiency.

However, it looks promising for future developments to investigate formulations of DLR techniques with an adaptive choice of the number of modes, based on suitable a-posteriori error estimators. The strategy to reduce the rank is quite straightforward. By analyzing the eigenvalues and eigenvectors of the correlation matrix of the non orthogonalized modes, one can easily drop the components corresponding to the eigenvalues which are below a given threshold. On the other hand, the technique to increase the rank is not obvious. Preliminary results have been obtained for the Dual DO formulation of parabolic diffusion equations with an adaptive strategy based on:

- an a posteriori error estimator to determine if and at which time the rank needs to be increased,

- the random initialization of the new stochastic modes, combined with a suitable time splitting, to exploit the instantaneous time-adaptivity of the DLR method.

The effectiveness of the random initialization has been already observed in this thesis for the DLR approximation with fixed number of modes (possibly rank deficient). Indeed the same strategy has been applied in Chapter 4 and Chapter 5 for the initialization of the modes at $t = 0$, with remarkable results. On the other hand, in the context of rank adaptivity, the advantage in terms of computational effort is evident: it allows to reduce the computational cost due to the initialization of the new modes, which can be done for instance by sampling the orthogonal component to the tangent space, as proposed in [117] or by power-type or Arnoldi iterations similarly to what proposed in [124] in the context of Proper Orthogonal Decomposition. The a posteriori error estimator is instead computed in the same spirit as what typically done in the Reduced Basis method [52, 55]. This technique is based on two ingredients: an estimate of the dual norm of the residual:

$$\epsilon(t^n) \cong \sup_{\omega \in \Omega} \left(\sup_{u \in \mathcal{H}} \frac{r(u, t^n, \omega)}{\|u\|_{\mathcal{H}}} \right)$$

where $r(u, t^n, \omega)$ is the residual of the equation, obtained by plugging the Dual DO approximate solution in the governing equation evaluated in ω, and a lower bound α_{min} for the coercivity constant. Thus, the error can be bounded as:

$$\| u(t) - u^{DLR}(t) \|^2_{\mathcal{H} \otimes L^2(\Omega)} \lesssim \left(\frac{\Delta t}{\alpha_{min}} \sum_{l=1}^{n} (\epsilon(t^l))^2 \right)$$

The algorithm under investigation can be summarized as follows:

Algorithm 1

1: $\bar{u}_{tmp} \leftarrow$ solve equation for the mean
2: **if** $S > 0$ **then**
3:     $\mathbf{U}_{tmp} \leftarrow$ solve equations for the determinitic modes
4:     $\mathbf{Y}_{tmp} \leftarrow$ solve equations for the stochastic modes with updated $\mathbf{U}_{tmp}$
5: $\Delta \leftarrow$ compute the residual
6: **if** $\Delta < \epsilon$ **then**
7:     $\bar{u}^{n+1} \leftarrow \bar{u}_{tmp}$
8:     $\mathbf{U}^{n+1} \leftarrow \mathbf{U}_{tmp}$
9:     $\mathbf{Y}^{n+1} \leftarrow \mathbf{Y}_{tmp}$
10: **else**
11:     $S \leftarrow S + 1$
12:     $Z \leftarrow$ random initialization , $\mathbf{Y}^n \leftarrow [\mathbf{Y}^n, Z]$
13:     $V \leftarrow 0, \ \mathbf{U}^n \leftarrow [\mathbf{U}^n, V]$
14:     QR-orthogonalization of $\mathbf{Y}^n$
15:     $\bar{u}^{n+1} \leftarrow$ solve equation for the mean
16:     $\mathbf{U}^{n+1} \leftarrow$ solve equations for the determinitic modes
17:     $\mathbf{Y}^{n+1} \leftarrow$ solve equations for the stochastic modes with updated $\mathbf{U}^{n+1}$

The effectiveness of the proposed algorithm has been tested on the following toy-problem:

$$\begin{cases} \dfrac{\partial u(x, t, \omega)}{\partial t} - \dfrac{\partial^2 u(x, t, \omega)}{\partial x^2} = f(x, t, \omega) & x \in (0, \pi), \ t \in (0, T], \ \omega \in \Omega \\ u(0, t; \omega) = u(2\pi, t; \omega) = 0 & t \in (0, T], \ \omega \in \Omega \\ u(x, 0; \omega) = \sin(x) & x \in (0, 2\pi), \ \omega \in \Omega \end{cases} \quad (6.1)$$

where

$$\begin{aligned} f(x, t, \omega) = \ & Z_1(\omega) x(x - \pi) + Z_2(\omega)(|\cos(x)| - 1)(t \geq 0.01) + Z_1(\omega) \sin(4x)(t \geq 0.015) \\ & + (Z_2^3(\omega) + Z_3(\omega)) \sin(2x)(t \geq 0.02) \end{aligned}$$

with Z_1, Z_2, Z_3 independent and uniformly distributed random variables in $[-1, 1]$. The exact solution is deterministic at the initial time and affected by randomness as time evolves for the effect of the forcing term. We start the approximation with $S = 0$. Figure 6.1 (left) shows that the algorithm is able to increase the number of modes when needed, namely at $t = 0$ to include the randomness coming from the term $Z_1(\omega) x(x - \pi)$, then at $t = 0.01$ and $t = 0.02$ when the terms
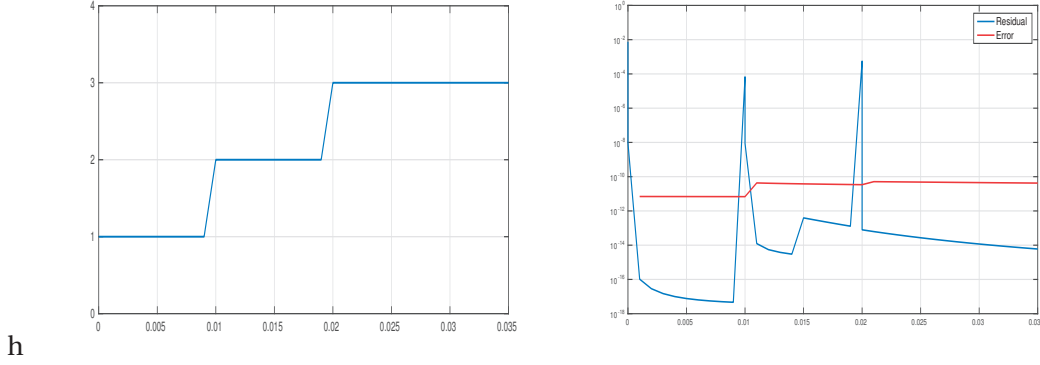
h

Figure 6.1 – Left: The number of modes. Right: evolution in time of the residual norm (blue) and the approximation error in norm $L^2(\Omega) \otimes H^1(D)$ of the Dual DO approximate solution with rank adaptive strategy, compared to a reference solution computed with Stochastic Collocation method. Discretization parameters: space steps $h = 0.01$, time step $\Delta t = 0.001$, stochastic tensor grid with Gauss-Legendre collocation points with 5 points in each directions.

$Z_2(\omega)(|\cos(x)| - 1)$ and $Z_3(\omega)\sin(2x)$ are activated respectively. Moreover, we observe that, as expected, no adding occurs at $t = 0.015$. This means that the term $Z_1(\omega)\sin(4x)$ belongs to the tangent space to $u_S(t)$ at $t = 0.015$, which confirms that the first added stochastic mode has been properly adapted by the algorithm. The analysis of the error in Figure 6.1 (right) supports the claim on the effectiveness of the technique. We observe that the approximation error of the Dual DO solution with rank adaptive strategy, which is, however, smaller than 10^{-10}, is due to the time discretization scheme and converges to zero as the time step decreases.

In the same direction, further adaptive strategies to drive a mesh adaptation during the time evolution or an adaptive choice of the time discretization step can be investigated to improve the performance of the DLR approximation.

The second achievement has been the extension of the DRL approach to parabolic equations, and, in particular, Navier Stokes equations, with random Dirichlet boundary conditions. We have proposed a convenient strategy to strongly impose to the DLR approximate solution a suitable (low rank) approximation of the Dirichlet boundary conditions. This resulted in an efficient dynamical low-rank approximation which accurately takes into account the randomness arising from the boundary data at the price of a slightly reduced flexibility in the evolution of the random modes. In particular, we showed that the set of S rank random fields, constrained to satisfy an approximation of the boundary datum of the exact solution, can be equipped with a structure of a differential manifold, allowing for a parametrization of its tangent space in terms of dynamical constraints on the stochastic coefficients. To do so we proposed a Dual DO formulation in which the stochastic modes are kept orthonormal. Under the assumption that the boundary datum g can be properly approximated by a linear combination g_M of $M < S$ terms written in the separable form, we fixed M stochastic modes in the approximate solution equal to those in the decomposition of g_M. This allowed us to identify the proper boundary

conditions for each (time dependent) deterministic mode and guarantee that the boundary constraint is fulfilled at each time. The same formulation is also used to conveniently include the incompressibility constraint when dealing with incompressible Navier-Stokes equations with random parameters. Indeed we were able to effortlessly impose the solenoidal constraint in each deterministic mode, facilitated by the fact that in the Dual DO formulation no numerical orthonormalization or dynamic constraint are required in the deterministic modes. In conclusion, Navier Stokes equations with random parameters, including random Dirichlet boundary conditions, has been reduced to S coupled deterministic PDEs of Navier-Stokes type and a system of $S - M$ stochastic ODEs. The Dual DO method has been tested on two fluid dynamics problems: the classical benchmark of a laminar flow around a cylinder with random inflow velocity, and a biomedical application simulating blood flows in a realistic carotid artery reconstructed from MRI data, where the inflow boundary conditions are taken as random due to the uncertainty and large errors in Doppler measurements of the inflow velocity. The results highlight the remarkable potentials of the Dual DO method for this type of applications.

The third achievement has been the development of a dynamical low-rank technique for the approximation of wave equations with random parameters. We proposed the Symplectic DO method, which combines the DLR approach, devised for parabolic equations, to the use of symplectic deterministic bases, as proposed in the Symplectic Order reduction of parametric Hamiltonian systems. The governing equation is rewritten in the Hamiltonian form in a suitable symplectic space and the approximate solution is expanded over a set of $2S$ (time dependent) symplectic-orthogonal deterministic modes, with (time dependent) stochastic coefficients. We derived the proper conditions to embed the set $\mathcal{M}_S^{sym}$, i.e. the collection of all S rank random fields which can be exactly expanded over a symplectic -orthonormal basis of dimension $2S$, with a manifold structure and we formulated the Symplectic DLR variational principle as the symplectic projection of the Hamiltonian system onto the tangent space to $\mathcal{M}_S^{sym}$ along the approximate trajectory. We showed that this corresponds to rewrite the governing Hamiltonian system in complex variables and looking for a DLR approximation in the manifold $\mathcal{M}_S^{\mathbb{C}}$ of all the complex-valued random fields with rank S. We used the analogy between the complex manifold $\mathcal{M}_S^{\mathbb{C}}$ and its real representation $\mathcal{M}_S^{sym}$ to determine a suitable parametrization of the tangent space in the real framework. After deriving the associated orthogonal constraints on the dynamics of the deterministic modes, we recover the reduced dynamical system which consists of a set of equations for the constrained dynamics of the deterministic modes, coupled with a reduced order Hamiltonian system for the evolution of the stochastic coefficients. The Symplectic DO shares with the symplectic order reduction the use of symplectic deterministic bases, and, as the "classic" DO approximation, allows both the stochastic and the deterministic modes to evolve in time. As a result, the approximate solution preserves the (approximated) mean Hamiltonian energy and continuously adapts in time to the structure of the solution.

The variational formulation of the Symplectic DO method sets a basis for extending to wave equations with random data the error analysis derived for the DLR approximation of lin-

ear parabolic equations. Promising seems also the analysis of the well-posedness of the Symplectic DLR problem, as strategies that exploit the conservation of energy may be employed here to prove existence and uniqueness of the approximate solution. Envisaged future investigations concern also the generalization of the Symplectic DLR approach to dynamical low-rank approximations with arbitrary (not necessarily orthonormal) symplectic bases. Indeed, the assumption of orthonormality of the deterministic modes is used here to exploit the complex representation and parametrize the tangent space but is not strictly needed to define the Symplectic DLR variational principle. Interesting is then the study of some possible characterizations of the approximation manifold when this condition is removed. This approach could potentially lead to remarkable improvements in the effectiveness of the approximation.

Bibliography

[1] P-A. Absil, R. Mahony, and R. Sepulchre. Riemannian geometry of Grassmann manifolds with a view on algorithmic computation. *Acta Applicandae Mathematica*, 80(2):199–220, 2004.

[2] P-A Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[3] M. Alnæs, J. Hake, R. Kirby, H. Langtangen, A. Logg, and G. Wells. The fenics manual. *FEniCS Project, version October 31st*, 2011.

[4] A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. *Journal of Non-Newtonian Fluid Mechanics*, 139(3):153–176, 2006.

[5] I. Babuska, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM review*, 52(2):317–355, 2010.

[6] I. Babuska, R. Tempone, and G. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM Journal on Numerical Analysis*, 42(2):800–825, 2004.

[7] C. Bardos, I. Catto, N. Mauser, and S. Trabelsi. Setting and analysis of the multi-configuration time-dependent Hartree–Fock equations. *Archive for Rational Mechanics and Analysis*, 198(1):273–330, 2010.

[8] C. Bardos, I. Catto, N.J. Mauser, and S. Trabelsi. Global-in-time existence of solutions to the multiconfiguration time-dependent hartree–fock equations: A sufficient condition. *Applied Mathematics Letters*, 22(2):147 – 152, 2009.

[9] M. Baumann and U. Helmke. Singular value decomposition of time-varying matrices. *Future Gener. Comput. Syst.*, 19(3):353–361, April 2003.

[10] U. Baur, P. Benner, B. Haasdonk, C. Himpe, and I. M. M. Ohlberger. Comparison of methods for parametric model order reduction of instationary problems. *preprint*, 2015.

Bibliography

[11] J. Beck, R. Tempone, F. Nobile, and L. Tamellini. On the optimal polynomial approximation of stochastic PDEs by galerkin and collocation methods. *Mathematical Models and Methods in Applied Sciences*, 22(09):1250023, 2012.

[12] M. H. Beck, A. Jäckle, G.A. Worth, and H-D. Meyer. The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets. *Physics reports*, 324(1):1–105, 2000.

[13] G. Berkooz, P. Holmes, and J. L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual review of fluid mechanics*, 25(1):539–575, 1993.

[14] M. Billaud-Friess and A. Nouy. Dynamical model reduction method for solving parameter-dependent dynamical systems. *arXiv preprint arXiv:1604.05706*, 2016.

[15] R. Botnar, G. Rappitsch, M. B. Scheidegger, D. Liepsch, K. Perktold, and P. Boesiger. Hemodynamics in the carotid artery bifurcation:: a comparison between numerical simulations and in vitro MRI measurements. *Journal of Biomechanics*, 33(2):137 – 144, 2000.

[16] N. Bourbaki. *Variétés différentielles et analytiques: fascicule de résultats*. Springer Science & Business Media, 2007.

[17] S. Boyaval, C. Le Bris, T. Lelievre, Y. Maday, N. Nguyen, and A. Patera. Reduced basis techniques for stochastic problems. *Archives of Computational methods in Engineering*, 17(4):435–454, 2010.

[18] A. Buffa, Y. Maday, A. T. Patera, C. Prud'homme, and G. Turinici. A priori convergence of the greedy algorithm for the parametrized reduced basis method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(3):595–603, 2012.

[19] A. Bunse-Gerstner, R. Byers, V. Mehrmann, and N. K. Nichols. Numerical computation of an analytic singular value decomposition of a matrix valued function. *Numerische Mathematik*, 60(1):1–39, 1991.

[20] J. Burkardt, M.D. Gunzburger, and C. Webster. Reduced order modeling of some nonlinear stochastic partial differential equations. *Int. J. Numer. Anal. Model.*, 4(3-4):368–391, 2007.

[21] R. E. Caflisch. Monte carlo and quasi-monte carlo methods. *Acta numerica*, 7:1–49, 1998.

[22] R. Cameron and W. Martin. The orthogonal development of non-linear functionals in series of fourier-hermite functionals. *Annals of Mathematics*, pages 385–392, 1947.

[23] K. Carlberg and C. Farhat. A low-cost, goal-oriented compact proper orthogonal decomposition basis for model reduction of static systems. *International Journal for Numerical Methods in Engineering*, 86(3):381–402, 2011.

[24] K. Carlberg, R. Tuminaro, and P. Boggs. Preserving Lagrangian structure in nonlinear model reduction with application to structural dynamics. *SIAM Journal on Scientific Computing*, 37(2):B153–B184, 2015.

[25] I. Chavel. *Riemannian geometry: a modern introduction*, volume 98. Cambridge university press, 2006.

[26] P. Chen, A. Quarteroni, and G. Rozza. A weighted reduced basis method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 51(6):3163–3185, 2013.

[27] P. Chen, A. Quarteroni, and G. Rozza. Comparison between reduced basis and stochastic collocation methods for elliptic problems. *Journal of Scientific Computing*, 59(1):187–216, 2014.

[28] P. Chen, A. Quarteroni, and G. Rozza. Reduced order methods for uncertainty quantification problems. *ETH Zurich, SAM technical report*, 3, 2015.

[29] P. Chen and C. Schwab. Model order reduction methods in computational uncertainty quantification. 2015.

[30] M. Cheng, T. Y. Hou, and Z. Zhang. A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations II: Adaptivity and generalizations. *J. Comput. Phys.*, 242:753–776, June 2013.

[31] M. Cheng, T.Y. Hou, and Z. Zhang. A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations I: Derivation and algorithms. *Journal of Computational Physics*, 242:843 – 868, 2013.

[32] F. Chinesta and P. Ladevèze. *Separated Representations and PGD-Based Model Reduction.* Springer, 2014.

[33] F. Chinesta, P. Ladeveze, and E. Cueto. A short review on model order reduction based on proper generalized decomposition. *Archives of Computational Methods in Engineering*, 18:395–404, 2011.

[34] M. Choi, T. P. Sapsis, and G. E. Karniadakis. A convergence study for SPDEs using combined polynomial chaos and dynamically-orthogonal schemes. *Journal of Computational Physics*, 245:281–301, 2013.

[35] M. Choi, T.P. Sapsis, and G.E. Karniadakis. On the equivalence of dynamically orthogonal and bi-orthogonal methods: Theory and numerical simulations. *Journal of Computational Physics*, 270:1 – 20, 2014.

[36] M. Choi, T.P. Sapsis, and G.E. Karniadakis. A robust bi-orthogonal/dynamically-orthogonal method using the covariance pseudo-inverse for the stochastic Navier-Stokes equations. *preprint*, 2016.

Bibliography

[37] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Computing and Visualization in Science*, 14(1):3, 2011.

[38] A. Cohen, R. Devore, and C. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs. *Analysis and Applications*, 9(01):11–47, 2011.

[39] D. Conte and C. Lubich. An error analysis of the multi-configuration time-dependent Hartree method of quantum dynamics. *Mathematical modelling and numerical analysis*, 44(4):759–780, 6 2010.

[40] L. Dieci and T. Eirola. On smooth decompositions of matrices. *SIAM Journal on Matrix Analysis and Applications*, 20(3):800–819, 1999.

[41] L. Dieci, R. Russell, and E. S. Van Vleck. Unitary integrators and applications to continuous orthonormalization techniques. *SIAM Journal on Numerical Analysis*, 31(1):261–281, 1994.

[42] P. AM Dirac. *Note on exchange phenomena in the Thomas atom.* Mathematical Proceedings of the Cambridge Philosophical Society, Cambridge Univ Press, 1930.

[43] J. L. Eftang, D. J. Knezevic, and A. T. Patera. An hp certified reduced basis method for parametrized parabolic partial differential equations. *Mathematical and Computer Modelling of Dynamical Systems*, 17(4):395–422, 2011.

[44] H. C. Elman and Q. Liao. Reduced basis collocation methods for partial differential equations with random coefficients. *SIAM/ASA Journal on Uncertainty Quantification*, 1(1):192–217, 2013.

[45] L.C. Evans. *Partial Differential Equations (Graduate Studies in Mathematics vol 19)(Providence, RI: American Mathematical Society).* Oxford University Press, 1998.

[46] G. Fishman. *Monte Carlo: concepts, algorithms, and applications.* Springer Science Business Media, 2013.

[47] J.I. Frenkel. *Wave mechanics: advanced general theory.* Clarendon Press Oxford, 1934.

[48] M. Gerritsma, J-B. van der Steen, P. Vos, and G. Karniadakis. Time-dependent generalized polynomial chaos. *J. Comput. Phys.*, 229(22):8333–8363, 2010.

[49] R.G. Ghanem and P.D. Spanos. *Stochastic Finite Elements: a Spectral Approach.* Springer–Verlag, New York, 1991.

[50] R. W. Gill. Measurement of blood flow by ultrasound: accuracy and sources of error. *Ultrasound in medicine & biology*, 11(4):625–641, 1985.

[51] I. G. Graham, F. Y. Kuo, D. Nuyens, R. Scheichl, and I. H. Sloan. Quasi-monte carlo methods for elliptic PDEs with random coefficients and applications. *Journal of Computational Physics*, 230(10):3668–3694, 2011.

186

[52] M. A. Grepl and A. T. Patera. A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *ESAIM: M2AN*, 39(1):157–181, 2005.

[53] B. Haasdonk. Convergence rates of the pod–greedy method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47(3):859–873, 2013.

[54] B. Haasdonk. Convergence rates of the pod–greedy method. *ESAIM: M2AN*, 47(3):859–873, 2013.

[55] B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 42(2):277–302, 2008.

[56] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, volume 31. Springer Science & Business Media, 2006.

[57] P. Harms and A. Mennucci. Geodesics in infinite dimensional Stiefel and Grassmann manifolds. *Comptes Rendus Mathematique*, 350(15-16):773–776, 2012.

[58] J.S. Hesthaven, G. Rozza, and B. Stamm. *Certified reduced basis methods for parametrized partial differential equations.* SpringerBriefs in Mathematics. Springer, Cham; BCAM Basque Center for Applied Mathematics, Bilbao, 2016.

[59] V. Heuveline and M. Schick. A local time–dependent generalized polynomial chaos method for stochastic dynamical systems. EMCL Preprint Series 04, EMCL, Karlsruhe Institute of Technology, 2011.

[60] P. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamical systems and symmetry.* Cambridge university press, 1998.

[61] L. Iapichino. *Reduced basis methods for the solution of parametrized PDEs in repetitive and complex networks with application to CFD.* PhD thesis, n: 5529, École Polytechnique Fédérale de Lausanne, 2012.

[62] I. Jolliffe. *Principal component analysis.* Wiley Online Library, 2002.

[63] S. Kobayashi and K. Nomizu. *Foundations of differential geometry.* New York, 1963.

[64] S. Kobayashi and K. Nomizu. *Foundations of differential geometry*, volume 1. New York, 1963.

[65] O. Koch, W. Kreuzer, and A. Scrinzi. Approximation of the time-dependent electronic Schrödinger equation by MCTDHF. *Appl. Math. Comput.*, 173(2):960–976, February 2006.

Bibliography

[66] O. Koch and C. Lubich. Dynamical low-rank approximation. *SIAM J. Matrix Anal. Appl.,* 29(2):434–454, 2007.

[67] O. Koch and C. Lubich. Regularity of the multi-configuration time-dependent Hartree approximation in quantum molecular dynamics. *ESAIM: Mathematical Modelling and Numerical Analysis,* 41(2):315–331, 2007.

[68] O. Koch and C. Lubich. Dynamical tensor approximation. *SIAM J. Matrix Anal. Appl.,* 31(5):2360–2375, 2010.

[69] D. Kressner, M. Steinlechner, and B. Vandereycken. Low-rank tensor completion by riemannian optimization. *BIT Numerical Mathematics,* 54(2):447–468, 2014.

[70] A. Kriegl and P. Michor. *The convenient setting of global analysis,* volume 53. American Mathematical Soc., 1997.

[71] Dieci L. and Eirola T. On smooth decompositions of matrices. *SIAM Journal on Matrix Analysis and Applications,* 20(3):800–819, 1999.

[72] P. Ladeveze. New algorithms: mechanical framework and development. *Compte rendu de l'académie des Sci,* 300(2):41–44, 1985.

[73] S. Lall, P. Krysl, and J. Marsden. Structure-preserving model reduction for mechanical systems. *Physica D: Nonlinear Phenomena,* 184(1):304–318, 2003.

[74] S. Lang. *Differential Manifolds.* Springer New York, 2012.

[75] O. P. Le Maître and O. M. Knio. *Spectral methods for uncertainty quantification.* Scientific Computation. Springer, New York, 2010. With applications to computational fluid dynamics.

[76] J. Lee. Smooth manifolds. In *Introduction to Smooth Manifolds,* pages 1–29. Springer, 2003.

[77] P. Libermann and C.M. Marle. *Symplectic geometry and analytical mechanics,* volume 35. Springer Science & Business Media, 2012.

[78] M. Loève. Probability theory, vol. ii. *Graduate texts in mathematics,* 46:0–387, 1978.

[79] G.J. Lord, C.E. Powell, and T. Shardlow. *An introduction to computational stochastic PDEs.* Cambridge Texts in Applied Mathematics. Cambridge University Press, New York, 2014.

[80] C. Lubich. *From quantum to classical molecular dynamics: reduced models and numerical analysis.* European Mathematical Society, 2008.

[81] C. Lubich, I.V. Oeledets, and B. Vandereycken. Time integration of tensor trains. *SIAM J. Numer. Anal.,* 53(2):917–941, 2015.

[82] C. Lubich and I. V. Oseledets. A projector-splitting integrator for dynamical low-rank approximation. *BIT Numerical Mathematics*, 54(1):171–188, 2014.

[83] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken. Dynamical approximation by hierarchical Tucker and tensor-train tensors. *SIAM Journal on Matrix Analysis and Applications*, 34(2):470–494, 2013.

[84] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken. Dynamical approximation of hierarchical Tucker and tensor-train tensors. *SIAM J. Matrix Anal. Appl.*, 34(2):470–494, 2013.

[85] X. Ma and G. E. Karniadakis. A low-dimensional model for simulating three-dimensional cylinder flow. *Journal of Fluid Mechanics*, 458:181–190, 2002.

[86] B. Maboudi and J. S. Hesthaven. Structure preserving model reduction of parametric Hamiltonian systems. *Siam Journal on Scientific Computing*, 2016.

[87] J. E Marsden and T. Ratiu. Introduction to mechanics and symmetry. *Physics Today*, 48(12):65, 1995.

[88] J. E Marsden and T. Ratiu. *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*, volume 17. Springer Science & Business Media, 2013.

[89] H. G. Matthies and A. Keese. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Computer methods in applied mechanics and engineering*, 194(12):1295–1331, 2005.

[90] N.J. Mauser and S. Trabelsi. L2 analysis of the multi-configuration time-dependent hartree–fock equations. *Mathematical Models and Methods in Applied Sciences*, 20(11):2053–2073, 2010.

[91] D. McDuff and D. Salamon. *Introduction to symplectic topology*. Oxford University Press, 1998.

[92] H-D. Meyer, U. Manthe, and L. S. Cederbaum. The multi-configurational time-dependent Hartree approach. *Chemical Physics Letters*, 165(1):73–78, 1990.

[93] B. Mishra, G. Meyer, S. Bonnabel, and R. Sepulchre. Fixed-rank matrix factorizations and Riemannian low-rank optimization. *Computational Statistics*, 29(3-4):591–621, 2014.

[94] M. Motamed, F. Nobile, and R. Tempone. Analysis and computation of the elastic wave equation with random coefficients. *Computers & Mathematics with Applications*, 70(10):2454–2473, 2015.

[95] E. Musharbash and F. Nobile. Dual Dynamically Orthogonal approximation of incompressible Navier Stokes equations with random boundary conditions. Technical report, EPFL, 2017.

Bibliography

[96] E. Musharbash, F. Nobile, and T. Zhou. Error analysis of the Dynamically Orthogonal approximation of time dependent random PDEs. *SIAM Journal on Scientific Computing*, 37(2):A776–A810, 2015.

[97] N.C. Nguyen, G. Rozza, and A. T. Patera. Reduced basis approximation and a posteriori error estimation for the time-dependent viscous burgers equation. *Calcolo*, 46(3):157–185, 2009.

[98] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*. SIAM, 1992.

[99] B. R. Noack, K. Afanasiev, M. Morzynski, G. Tadmor, and F. Thiele. A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *Journal of Fluid Mechanics*, 497:335–363, 2003.

[100] F. Nobile and R. Tempone. Analysis and implementation issues for the numerical approximation of parabolic equations with random coefficients. *International journal for numerical methods in engineering*, 80(6-7):979–1006, 2009.

[101] F. Nobile, R. Tempone, and C. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.

[102] A. Nouy. A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 196(45):4521–4537, 2007.

[103] A. Nouy. Generalized spectral decomposition method for solving stochastic finite element equations: invariant subspace problem and dedicated algorithms. *Computer Methods in Applied Mechanics and Engineering*, 197(51):4718–4736, 2008.

[104] A. Nouy. A priori model reduction through proper generalized decomposition for solving time-dependent partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 199(23):1603–1626, 2010.

[105] A. Nouy. Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems. *Archives of Computational Methods in Engineering*, 17(4):403–434, 2010.

[106] A. Nouy and O. P. Le Maître. Generalized spectral decomposition for stochastic nonlinear problems. *J. Comput. Phys.*, 228(1):202–235, 2009.

[107] L. Peng and K. Mohseni. Symplectic model reduction of Hamiltonian systems. *SIAM Journal on Scientific Computing*, 38(1):A1–A27, 2016.

[108] K. Perktold and G. Rappitsch. Computer simulation of local blood flow and vessel mechanics in a compliant carotid artery bifurcation model. *Journal of Biomechanics*, 28(7):845 – 856, 1995.

[109] S. Prajna. Pod model reduction with stability guarantee. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 5, pages 5254–5258. IEEE, 2003.

[110] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations*, volume 92 of *Unitext*. Springer, 2016.

[111] A. Quarteroni, A. Veneziani, and C. Vergara. Geometric multiscale modeling of the cardiovascular system, between theory and practice. *Computer Methods in Applied Mechanics and Engineering*, 302:193–252, 2016.

[112] F Riesz and B. Nagy. Functional analysis. Dover. *New York*, 1990.

[113] S. Roman. *Advanced linear algebra*, volume 3. Springer, 2005.

[114] Z. Ruan. On real operator spaces. *Acta Mathematica Sinica*, 19(3):485–496, 2003.

[115] S. Salsa. *Partial differential equations in action: from modelling to theory*, volume 99. Springer, 2016.

[116] T.P. Sapsis and P.F.J. Lermusiaux. Dynamically Orthogonal field equations for continuous stochastic dynamical systems. *Phys. D*, 238(23-24):2347–2360, 2009.

[117] T.P. Sapsis and P.F.J. Lermusiaux. Dynamical criteria for the evolution of the stochastic dimensionality in flows with uncertainty. *Phys. D*, 241(1):60–76, 2012.

[118] C. Schwab and C. Gittelson. Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. *Acta Numerica*, 20:291–467, 2011.

[119] C. Schwab and R. A. Todor. Karhunen–loève approximation of random fields by generalized fast multipole methods. *Journal of Computational Physics*, 217(1):100–122, 2006.

[120] R. C. Smith. *Uncertainty quantification*, volume 12 of *Computational Science & Engineering*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2014. Theory, implementation, and applications.

[121] C. Soize and R. Ghanem. Physical systems with random uncertainties: chaos representations with arbitrary probability measure. *SIAM Journal on Scientific Computing*, 26(2):395–410, 2004.

[122] N. Steenrod. *The topology of fibre bundles*, volume 14. Princeton University Press, 1951.

[123] T. J. Sullivan. *Introduction to uncertainty quantification*, volume 63 of *Texts in Applied Mathematics*. Springer, Cham, 2015.

[124] L. Tamellini, O. Le Maitre, and A. Nouy. Model reduction based on proper generalized decomposition for the stochastic steady incompressible navier–stokes equations. *SIAM Journal on Scientific Computing*, 36(3):A1089–A1117, 2014.

[125] L.J.P. Timmermans, P.D. Minev, and F.N. Van De Vosse. An approximate projection scheme for incompressible flow using spectral elements. *International journal for numerical methods in fluids*, 22(7):673–688, 1996.

[126] S. Trabelsi. Solutions of the multiconfiguration time-dependent hartree–fock equations with coulomb interactions. *Comptes Rendus Mathematique*, 345(3):145 – 150, 2007.

[127] M.P. Ueckermann, P.F.J. Lermusiaux, and T.P. Sapsis. Numerical schemes for dynamically orthogonal equations of stochastic fluid and ocean flows. *Journal of Computational Physics*, 233:272 – 294, 2013.

[128] X. Wan and G.E. Karniadakis. Long-term behavior of polynomial chaos in stochastic flow simulations. *Comput. Methods Appl. Mech. Engrg.*, 195(41-43):5582–5596, 2006.

[129] A. Weinstein. Symplectic manifolds and their Lagrangian submanifolds. *Advances in mathematics*, 6(3):329–346, 1971.

[130] F. M. White. Fluid mechanics. *Boston: McGraw-Hill Book Company, 5th ed.*, 2003.

[131] N. Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.

[132] K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. *AIAA Journal*, 40(11):2323–2330, 2002.

[133] D. Xiu. Efficient collocational approach for parametric uncertainty analysis. *Commun. Comput. Phys*, 2(2):293–309, 2007.

[134] D. Xiu. *Numerical methods for stochastic computations*. Princeton University Press, Princeton, NJ, 2010. A spectral method approach.

[135] D. Xiu and J. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM Journal on Scientific Computing*, 27(3):1118–1139, 2005.

[136] D. Xiu and G. E. Karniadakis. The wiener–askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2):619–644, 2002.

[137] D. Xiu and G. E. Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal of computational physics*, 187(1):137–167, 2003.

[138] J. Zanghellini, M. Kitzler, C. Fabian, T. Brabec, and A. Scrinzi. An MCTDHF approach to multielectron dynamics in laser fields. *Laser Physics*, 13(8):1064–1068, 2003.

[139] T. Zhou and T. Tang. Galerkin methods for stochastic hyperbolic problems using bi-orthogonal polynomials. *Journal of Scientific Computing*, 51(2):274–292, 2012.

Curriculum Vitae

NAME: Eleonora Musharbash
DATE OF BIRTH: July 28th, 1986
NATIONALITY: Italian
E-mail: eleonora.musharbash@epfl.ch

Education

- May 2013 - present: **PhD student** and teaching assistant in the Department of **Applied Mathematics at EPFL**, Lausanne, Switzerland.
  Focus: reduced order methods for PDEs with random parameters.
  Supervisor: Fabio Nobile (Associate Professor at EPFL, Lausanne)

- Set. 2010 - Dec. 2012: **MSc. Mathematical Engineering**, Politecnico di Milano, Italy. Major: scientific computing. Dissertation: "Dynamical low rank approximation of time dependent PDEs with random parameters". Final Mark: 110/110. Supervisor: Prof. Fabio Nobile, Co-Supervisor: Dr. Tao Zhou, (Assistant Professor, Institute of Computational Mathematics, Chinese Academy of Sciences; formerly postdoc at EPFL.)

- Oct. 2009: Curriculum Integrations: Engineering Mathematics (single courses), Politecnico di Milano, Italy.

- Nov. 2008: **Bachelor Degree in Applied Mathematics** (cum laude), Università degli Studi di Perugia, Italy. Dissertation on Probability Measures

Further Experiences

- Feb. - Jul. 2008: Erasmus program, Sheffield University, Sheffield, United Kingdom.

- Jul. - Dec 2011: International Exchange program, Pontificia Universitade Catolica, Rio de Janeiro, Brasil .

- May - Oct 2012: Exchange program and Thesis Project at EPFL, Lausanne, Switzerland.

Publications

- E. Musharbash, F. Nobile and T. Zhou. Error Analysis of the Dynamically Orthogonal Approximation of Time Dependent Random PDEs, in Siam Journal on Scientific Computing, vol. 37, num. 2, p. A776-A810, 2015.

Bibliography

- E. Musharbash, F. Nobile. Dynamically Orthogonal approximation of incompressible Navier Stokes equations with random boundary conditions, submitted for publication.

Contributed Talks in Conferences/Workshop

- SIMAI 2016 (Milan, Italy)

- MORCIP 2016 (Lausanne, Switzerland)

- Swiss Numerics 2016 (Fribourg, Switzerland)

- SIAM-UQ 2016 (Lausanne, Switzerland)

- workshop COST EU-MORNET WG1 ROM 2016 (Trieste, Italy)

- ICIAM 2015 (Beijing, China)

- Swiss Numerics 2014 (Geneva, Switzerland)

- SIAM UQ 2014 (Savannah, USA)