

# Entropy in the Coil-to-Globule Transition of Macromolecules: Insights from Simple Models

THÈSE N° 5256 (2012)

PRÉSENTÉE LE 13 JANVIER 2012  
À LA FACULTÉ DES SCIENCES DE BASE  
LABORATOIRE DE BIOPHYSIQUE STATISTIQUE  
PROGRAMME DOCTORAL EN PHYSIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Carlo MAFFI

acceptée sur proposition du jury:

Prof. M. Q. Tran, président du jury  
Prof. P. De Los Rios, directeur de thèse  
Prof. G. Dietler, rapporteur  
Prof. F. Seno, rapporteur  
Prof. C. Vanderzande, rapporteur



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

Suisse  
2012



# Acknowledgments

I acknowledge my supervisor, Prof. Paolo De Los Rios. Since our first contact, he has always shown an authentic and contagious passion while discussing about physics, even if strained to the limit during the tense period preceding his promotion; I am happy to have been present during this so important step of his career! His support and incentive have been fundamental for the progress and the conclusion of this work.

Many thanks to Lucio for many exchanges of ideas and for support in occasion of presentations or travels; to Vishal, who helped me with useful comments on the thesis; and to Salvo, not only for his Modica chocolate! The assistance activity gave me always big motivation: I acknowledge all students, and particularly Joana, a very brilliant student I worked together for a long while.

Paolo Torrielli, while he was working as a crazy on many other projects, he was always there for discussing some difficult calculation or helping me to understand which direction to take after the doctorate: thank you very much! Thanks to people in the Open Space, where the atmosphere was always so nice! With many of them I shared particularly amusing and happy times. Thanks to Vittoria Rezzonico (Vicky) for her precious and expert and nice support in part of my work.

One huge and special thank goes to my friend Biagio, who sustained me at work and in life during the whole duration of my doctorate. I want to mention in particular the support I received from the maturity and the pleasantness of Andrea Wulzer.

And I dedicate an immensely tender thought to Andrea.



# Abstract

The coil to globule transition is a fundamental phenomenon in the physics of macro-molecules by reason of the multiplicity of arrangements of their conformation. Such conformational freedom is the main source of entropy in the molecule and is the main opponent to the transition towards the compact state, since a system tends to the state of maximum entropy.

This phenomenon is captured by very simple models, such as the ensemble of Interacting Self avoiding Walks on the lattice. This model shows that the coil to globule transition belongs to the universality class of continuous transition called  $\Theta$  point.

Starting from a critical inspection of the definition of the interacting walks model, we introduce a refinement aiming to represent more precisely the entropy sourced from the local fluctuations of the molecule around its equilibrium conformations; this contribution is absent in the standard model which includes only the entropy generated by the multiplicity of the global conformations.

Through the study of the vibrational properties of the model and an exhaustive numerical simulation of the phase transition we show that the new model presents a transition which is different in character and belongs to a distinct universality class. In particular, in 3 dimensions the model shows a discontinuous transition; thanks to this, the model presents a common framework underlying the physics of both homo-polymers and proteins.

Considering the results obtained in 2 and 3 dimensions we identify the convexity of the vibrational entropy as the responsible of the new class of transitions. These results and a revision of experimental measures of viscosity brings into question the standard description of the homo-polymer transition.

## Key words:

coil-globule transition,  $\Theta$  point, conformational entropy, self avoiding walk, tetrahedral lattice, square lattice, Monte Carlo simulation, cross over exponent, polymer, intrinsic viscosity, thermal blob, Hessian matrix, elastic network



# Riassunto

La transizione dallo stato disordinato allo stato compatto è un fenomeno fondamentale nella fisica delle macro-molecole, in virtù della loro possibilità di riarrangiare la propria conformazione in innumerevoli modi. Questa libertà conformazionale è la principale sorgente di entropia nella molecola e la principale opponente alla transizione verso lo stato compatto, in base al principio che i sistemi tendono a massimizzare la propria entropia.

Tale fenomeno può essere catturato da modelli molto semplici, quale l'insieme dei cammini auto evitanti e interagenti sul reticolo. Da questo modello si deduce che la transizione dalla fase aperta alla fase compatta appartiene alla classe di universalità di transizioni continue denominata punto  $\Theta$ .

Attraverso un riesame della definizione del modello dei cammini interagenti abbiamo introdotto un miglioramento atto a rappresentare più precisamente l'entropia generata dalle fluttuazioni locali della molecola attorno alle sue conformazioni di equilibrio; tale contributo è assente nel modello classico in cui solo l'entropia proveniente dalla molteplicità delle conformazioni globali è rappresentato.

Attraverso uno studio delle proprietà vibrazionali del modello e un ampio studio numerico della transizione di fase dimostriamo che il nuovo modello presenta una transizione di carattere diverso, appartenente a una diversa classe di universalità. In particolare in 3 dimensioni il modello presenta una transizione discontinua che permette di riunire in un unico quadro la fenomenologia di homo-polimeri e proteine.

Considerando i risultati ottenuti in 2 e 3 dimensioni individuiamo nella convessità dell'entropia vibrazionale il responsabile della nuova fenomenologia di transizione. I risultati di questo modello e una rivisitazione di misure sperimentali di viscosità rimettono in discussione la classica descrizione della transizione di omopolimeri.

## Parole chiave:

collasso di polimeri, punto  $\Theta$ , entropia conformazionale, cammino auto evitante, reticolo tetraedrico, reticolo quadrato, simulazione Monte Carlo, esponente di cross over, polimero, viscosità intrinseca, blob termico, matrice





# Contents

<b>Acknowledgments</b>	<b>3</b>
<b>Abstract (English/Italian)</b>	<b>5</b>
<b>List of figures</b>	<b>11</b>
<b>List of tables</b>	<b>14</b>
<b>1 Introduction</b>	<b>17</b>
1.1 Entropy and disorder . . . . .	18
1.2 Entropy and conformational freedom . . . . .	20
1.3 Coil to globule transition . . . . .	23
1.4 The $\Theta$ point . . . . .	25
1.5 Variations on the ISAW . . . . .	28
1.5.1 Model with stiffness . . . . .	28
1.5.2 Bond fluctuation model . . . . .	29
1.5.3 HP model . . . . .	29
1.6 Studies on conformational entropy . . . . .	30
1.7 Summary . . . . .	31
<b>2 A lattice model with vibrations</b>	<b>33</b>
2.1 Lattice models and elastic networks . . . . .	33
2.2 The partition function . . . . .	37
2.2.1 Relation between B-factor and vibrational entropy . .	40
2.3 The observables and the assumption of <i>small fluctuations</i> . .	42
2.4 An alternative approach: Energy Landscapes . . . . .	44
2.5 The Hessian matrix . . . . .	47
2.5.1 2 dimensions . . . . .	47
2.5.2 3 dimensions . . . . .	48
2.6 Comparison with off lattice force fields . . . . .	49
2.6.1 Small fluctuations . . . . .	52
2.7 Methods . . . . .	54
2.7.1 Units of measurement . . . . .	54
2.7.2 Simulation technique . . . . .	54

2.7.3	Data analysis . . . . .	55
2.7.4	Multiple Histogram Method . . . . .	56
2.7.5	Low noise differentiator . . . . .	57
<b>3</b>	<b>Vibrational spectrum of linear chains</b>	<b>59</b>
3.1	Linear chains and solids . . . . .	59
3.2	Coils . . . . .	62
3.2.1	2 dimensions . . . . .	62
3.2.2	3 dimensions . . . . .	64
3.3	Globules . . . . .	64
3.4	Spectrum of coil and compact ensemble . . . . .	66
3.5	Fractons . . . . .	67
<b>4</b>	<b>The coil to globule transition</b>	<b>69</b>
4.1	Choice of the spring constants in 3 dimensions . . . . .	70
4.2	3 dimensions: high and low temperatures . . . . .	72
4.3	3 dimensions: first order phase transition . . . . .	74
4.3.1	Van't Hoff test . . . . .	78
4.4	3 dimensions: metric properties . . . . .	80
4.4.1	High temperature . . . . .	81
4.4.2	$\Theta$ point . . . . .	82
4.5	2 dimensions: a more abrupt second order phase transition . . . . .	83
4.6	Convexity of the entropy . . . . .	87
4.6.1	2 dimensions . . . . .	90
4.6.2	3 dimensions . . . . .	92
4.6.3	Vibrational microcanonical entropy . . . . .	94
4.7	Conclusions . . . . .	94
<b>5</b>	<b>Other results</b>	<b>97</b>
5.1	Different spring constants . . . . .	97
5.1.1	The case $k_{co} = 0$ . . . . .	97
5.1.2	The case $k_{co} > 0$ . . . . .	99
5.2	Entropic exponent in 2 dimensions . . . . .	101
5.2.1	The connective constant $\mu < 1$ . . . . .	102
5.2.2	The exponent $\gamma < 0$ . . . . .	103
5.2.3	Normalization of data . . . . .	103
5.2.4	Conclusion . . . . .	104
5.3	Topological properties: preliminary results . . . . .	104
5.4	Chain adsorbed on a wall: preliminary results in $2d$ . . . . .	107
5.4.1	The model . . . . .	107
5.4.2	Vibrations . . . . .	108
5.4.3	Results . . . . .	108

<b>6</b>	<b>Viscosity experiments</b>	<b>111</b>
6.1	Thermal Blobs . . . . .	111
6.1.1	Flory Argument for $\phi = \frac{1}{2}$ . . . . .	113
6.1.2	Our result . . . . .	113
6.2	Intrinsic Viscosity . . . . .	113
6.3	Results . . . . .	115
6.3.1	Poly( <i>p</i> -methylstyrene): pPMS . . . . .	115
6.3.2	Poly(methyl methacrylate) in acetonitrile . . . . .	118
6.3.3	Poly(methyl acrylate) in toluene: pMA . . . . .	118
6.3.4	Poly(2-cinnamoyloxyethyl methacrylate): pMMA . . . . .	118
6.4	Conclusions . . . . .	119
<b>7</b>	<b>Conclusions</b>	<b>121</b>
<b>A</b>	<b>Hessian matrix of 2d walk</b>	<b>123</b>
A.1	bond . . . . .	123
A.2	angle . . . . .	124
A.3	contact . . . . .	124
<b>B</b>	<b>Hessian matrix of 3d walk</b>	<b>127</b>
B.1	bond . . . . .	128
B.2	bend . . . . .	128
B.3	trans . . . . .	129
B.4	gauche . . . . .	129
<b>C</b>	<b>Angular contribution to the vibrational spectrum</b>	<b>131</b>



# List of Figures

1.1	Images of DNA. . . . .	22
1.2	Coil to globule transition of polystyrene. . . . .	23
1.3	ISAWs on the lattice. . . . .	26
2.1	Self Avoiding Walks on the tetrahedral lattice. . . . .	35
2.2	Configurations, B-factor and Vibrational entropy. . . . .	41
2.3	Correlation between the B factor and the vibrational entropy. . . . .	42
2.4	Approximation of the energy landscape. . . . .	46
2.5	The Lanczos low noise differentiator. . . . .	57
3.1	Frequency spectrum of a linear chain. . . . .	60
3.2	Low frequency spectrum of a 3-dimensional solid. . . . .	61
3.3	Frequency spectrum of coil polymers. . . . .	63
3.4	Frequency spectrum of compact polymers. . . . .	65
3.5	Spectrum and density of frequencies of polymers in 3 dimensions. . . . .	66
4.1	Results in three dimensions. . . . .	71
4.2	Scaling properties of the transition in 3 dimensions. . . . .	76
4.3	Gaussian fit of double peaks. . . . .	77
4.4	The Van't Hoff ratio. . . . .	78
4.5	Radius of gyration in 3 dimensions. . . . .	80
4.6	Scaling of the model in 2 dimensions. . . . .	84
4.7	Metric properties of the 2 dimensions model. . . . .	85
4.8	Microcanonical entropy and canonical energy in 3 dimensions. . . . .	89
4.9	Microcanonical entropy and canonical energy in 2 dimensions. . . . .	91
4.10	Vibrational microcanonical entropy. . . . .	93
5.1	Vibrational entropy and spring constants. . . . .	98
5.2	Cross-over exponent at varying spring constants. . . . .	100
5.3	Partition function with respect to the size in 2 dimensions. . . . .	102
5.4	Helices in 2 dimensions. . . . .	105
5.5	Average contact length distribution . . . . .	106

6.1	Intrinsic Viscosity and thermal blob size. . . . .	116
6.2	Thermal blob size of Poly( <i>p</i> -methylstyrene) in diethyl succinate.	117
6.3	pMMA and pMA: reduced temperature and thermal blob mass	118
6.4	Thermal blob size of Poly(2-cinnamoyloxyethyl methacrylate) as a function of the reduced temperature. . . . .	119
C.1	3d polymer normal modes . . . . .	132

# List of Tables

1.1	End to end distances of a DNA molecule. . . . .	22
1.2	Critical exponents of the $\Theta$ point. . . . .	28
2.1	Table of symbols . . . . .	37
2.2	Observables of a SAW of $N$ monomers. . . . .	43
2.3	Interactions of VIBR. . . . .	50
2.4	Interaction parameters of the United Atom Potential . . . . .	51
2.5	Interaction parameters of Coarse Grained protein potential . . . . .	52
2.6	Some terms of the interaction parameters of Coarse Grained protein potential . . . . .	53
2.7	Average fluctuations of contacts and angles . . . . .	54
4.1	Typical spring constants in the VIBR partition function. . . . .	72
4.2	Spring constants in the VIBR partition function in 2 dimen- sions. . . . .	83
6.1	MHS exponent of pPMS in diethyl succinate . . . . .	117
6.2	Results of viscosity experiments . . . . .	120
A.1	Bond Hessian. . . . .	124
A.2	Right angle Hessian. . . . .	125
A.3	Flat angle Hessian. . . . .	125





# Chapter 1

## Introduction

Entropy is an important physics concept that has entered common language. Its etymology is the Greek word, *entropia*, meaning literally *a turning toward, conversion*. Curiously, its original meaning does not reflect entirely the one we attribute to it nowadays: the entropy is typically considered as the *measure of disorder* of a system. This common interpretation has foundation on the *scientific* usage of the word, in which context the concept of *disorder* demands to be precisely defined: we try to fulfill this request in the first section of this Introduction.

More specifically, if the system in question is a molecule, its disorder consists of the relentless internal activity, which is the effect of the interatomic interactions between the elements into play: the molecule itself, other molecules in close proximity and the solvent around them. To put order in a molecule means, for example, to confine it or to give it some shape; this action requires energy because it goes against the tendency of being in motion under the influence of the other interactions. In a similar way, the compression of a volume of gas requires energy: it is not a matter of lacking of space, but of hindering the molecular motion. Instead of running after the many molecular paths, physicists elaborate approximate, but effective, methods to deal with complex systems. Sometimes, very simple models give a correct interpretation of the experiments: the Ideal Gas Theory is a well known example; another case, closer to the subject treated in this Thesis, the stretching of DNA, is described in section 1.2. In this work we study the coil to globule transition. The concerned systems are polymers, sequences of many small molecules one after another, which are complex under many aspects: they present the complications of a one dimensional chain immersed in space<sup>1</sup>; at the same time their behavior is enriched by the variety of its components and interactions between them. This second aspect is particularly relevant in proteins because the composing units can be of 20 different

---

<sup>1</sup>Every one has experimented this kind of complexity when trying to unwinding the cable of his music player headphones.

types (the amino-acids), and the number of possible compositions becomes huge<sup>2</sup>. The formulas presented in section 1.3 are just the most concise depiction of the forces competing in the macromolecular folding. When trying to be more detailed, the solution of even the simplest physics model of homopolymer (a polymer whose units are identical) resorts to sophisticated calculations relying on Classical Field Theory, as we try to vaguely report in section 1.4.

Fortunately, today we have an increasing power of the computational means, and the models can be numerically investigated without resorting to complicated theories and strong approximations. The only approximations concern the model itself and how much reality one wants to implement; of course, some approximation is also imposed by the finiteness of the computers power. Without reaching the complexity of the Molecular Dynamics, several simple models have been elaborated during the years, and many of them are small variations of the mentioned simplest model. Each of these variations stresses a particular aspect of the problem and neglects others. Since this Thesis aims to present a new variation, in section 1.5 we briefly review some of the already existing ones, in order to get the feeling of the work. They are interesting to us because they show a transition with a phenomenology more complex than the original version, consequence of a refinement of the interactions. The modification proposed by us does not deal with the interactions, but with the representation of the disorder of the molecule or, in more technical terms, of its conformational entropy. Before presenting the development and the results of our model, we will mention, in section 1.6, other previous studies on the conformational entropy connected with the physics of molecules, as well as with very different fields, such as the magnetic transitions and the crystallization of metals.

## 1.1 Entropy and disorder

The first definition of *entropy* is due to R. Clausius, early 1850s, and concerns the relation between heat and temperature. Only in the 1870s, when L. Boltzmann began the microscopic interpretation of the thermodynamics, founding statistical physics, the interpretation of the entropy as a measure of the disorder of the system could be introduced.

In the thermodynamic context, entropy is a function of state that varies if any flow of heat occurs within the system. The behavior of this function is constrained by the second principle of the thermodynamics:

- *First principle*: the energy of the universe is constant;
- *Second principle*: the entropy of the universe tends to a maximum.

---

<sup>2</sup>This year, for example, the Protein Data Bank contains more than 75000 structures.

This synthetic formulation of the laws of the thermodynamics is due to Clausius ([25] is an interesting monograph dedicated to this subject); they regulate every day life, as we can realize solving the following elementary exercise.

**Exercise.** *Two identical bodies have temperatures  $T_H$  (hot) and  $T_C$  (cold) respectively,  $T_H > T_C$ . They are put in thermal contacts, so that they can exchange energy, but they are isolated from the external environment: together they form their own universe. How do the temperatures evolve?*

**Solution.** The energy exchanged by each body ( $dU$ ) is proportional to the variation of its temperature ( $dT$ ); because the two bodies are identical, they have the same heat capacity,  $\mathcal{C}$ , and the variation of energy and temperature are related by:  $dU = \mathcal{C}dT$ . The conservation of the energy of the universe,  $dU_H + dU_C = 0$ , implies that the temperature variation of each body is opposite to the temperature variation of the other:

$$dT_H = -dT_C$$

This relation does not state in which direction the heat flows: a situation in which the hot body warms up at the cold body's expense would be compatible with the first law, even if counterintuitive. The principle of maximization of entropy shows the right direction of the flow: the entropy variation of each body is  $dS = \frac{dU}{T} = \frac{\mathcal{C}dT}{T}$ , and for the entire universe (the two bodies):

$$dS_{TOT} = \mathcal{C} \left( \frac{dT_H}{T_H} + \frac{dT_C}{T_C} \right) = \mathcal{C} \left( \frac{dT_H}{T_H} - \frac{dT_H}{T_C} \right) > 0$$

which implies  $dT_H < 0$  (the hot body cools off). The final state is reached when the two bodies reach the same temperature. •

In the context of statistical physics, any thermodynamic quantity, which characterizes the *macro-state* of the system, is interpreted in terms of the state of its microscopic components, the *micro-states*: for example, the pressure of a gas is the effect of the mechanical motion of the molecules. The entropy has a special interpretation: through the maximization principle (second law), it plays a prominent role in the determination of the state of a system, which is denoted by a probability distribution over the ensemble of its micro-states, and in particular the most entropic one. The microscopic components of a system, for example the atoms composing a molecule, behave on smaller space lengths and faster time scales than the scales detectable by human observation. A system at equilibrium looks stable, but its micro components keep changing their state and its macroscopic aspect is the average of those states.

For example, we imagine a system whose states can be effectively enumerated and labeled with an integer counter  $i$ ; the observed energy is the average over all the micro-states, weighted with its probability:

$$\langle E \rangle = \sum_i E_i P_i. \quad (1.1)$$

The success of a statistical physics model relies on:

1. elaborating a satisfactory representation of the micro-states of the system;
2. establishing a good expression for the probability of each micro-state.

The probability of one micro-state depends on the constraints of the system; for example the total energy may be fixed and only micro-states with that energy have non zero probability (microcanonical ensemble); in another example the temperature is fixed, the energy can vary, but the micro-states with very high energy are less probable than micro-states with low energy (canonical ensemble). Usually those constraints can be accomplished by several distributions, but the system chooses the one which maximizes the entropy:

$$S = -k_B \sum_i P_i \ln P_i. \quad (1.2)$$

The choice of the maximum entropy probability distribution turns out to be the choice of the *most probable* probability distribution. This is shown in a non trivial way which involves the use of an ensemble of replicas of the system, as introduced by Gibbs [32], and we don't repeat it here. The definition (1.2) shows eventually the difference between an ordered and a disordered distribution. The lowest value for the entropy is  $S_{min} = 0$  because  $P_i \leq 1$  and this is the entropy of a *mono-disperse* distribution: only one state is possible,  $P_1 = 1$  and  $P_{i>1} = 0$ . An *ordered* system is a system of which we know exactly the microscopic state. On the opposite hand, a system wandering uniformly through all the micro-states,  $P_i = \frac{1}{\Omega}$ , is associated with the maximum entropy, given by the logarithm of the number of micro-states:

$$S = k_B \ln \Omega \quad (1.3)$$

This is an alternative well known expression of the entropy, but it is a special case of the previous one: equation (1.3) is the entropy of a system whose  $\Omega$  micro-states are equiprobable.

## 1.2 Entropy and conformational freedom

The expression (1.3) means that the entropy of a system increases with the number of available micro-states ( $\Omega$ ). In the physics of polymer and macromolecules the amount of micro-states is called the conformational freedom. In this section we illustrate the meaning of *conformational entropy*, basing on an example from the polymer physics: the stretching of a DNA molecule [7].

Ubiquitous in all living beings, the DNA encodes the genetic information for building the constituent bricks of the organism, the proteins. It is a

very long polymer, typically (in human beings) a sequence of millions of nucleotides. This means that one molecule is about 50 mm long, that is about  $10^6$  times its thickness, 2.5 nm. The string is also very stiff: its persistence length (the distance at which the correlation in the direction of the chain is almost lost) is about 50 nm, more than 10 times its diameter, but still thousands of time shorter than the whole molecule. For these characteristics the geometric conformation of a DNA molecule can be approximated with a random coil, a succession of  $N$  steps of equal length,  $b$ , and uncorrelated random direction. One step should be chosen long enough to represent a portion of DNA the ends of which have uncorrelated directions. A good choice is a step length equal to two times the persistence length. The end to end distance of a chain is the length of the vector joining the starting and ending points of the chain: it measures  $R_{min} = 0$  when the ends touch and  $R_{max} = bN$  when the chain is completely stretched. All the other configurations show intermediate values. Four quite short fragments of DNA are shown in Figure 1.1; the possible configurations are infinitely more numerous. The probability distribution of the end to end distance over the possible configurations of a freely jointed chain is the distribution of the sum of  $N$  (many) uncorrelated steps:

$$P(\vec{R}_e) = \frac{e^{-\frac{3}{2} \frac{\vec{R}_e^2}{Nb^2}}}{(2\pi Nb^2/3)^{\frac{3}{2}}}. \quad (1.4)$$

The average end to end distance is the size of the random walk,

$$R_{ave} \equiv \sqrt{\langle \vec{R}_e^2 \rangle} = bN^{1/2}, \quad (1.5)$$

which is much smaller than the length of the completely stretched chain; on the other hand a completely packed chain occupies even less place (see Table 1.1).

The DNA molecule at normal physiological conditions interacts with much smaller solvent molecules, continually receiving kicks from all directions. The DNA molecule is constantly in motion and changing its conformation on very fast time scales, so its size is well represented by the average over all the possible random configurations. The only assumption of randomness gives a quantitatively good approximation of the statistical distribution of the molecule conformation; this result is quite striking because that simple assumption summarizes the whole ensemble of molecular interactions which determine the dynamics of the molecule. More precisely, the free energy of the ensemble of random chains with end to end distance  $\vec{R}$  is

$$F(\vec{R}) = -k_B T \ln P(\vec{R}) = U(\vec{R}) - TS(\vec{R}) = Tk_B \frac{3}{2} \frac{\vec{R}_e^2}{Nb^2} \quad (1.6)$$

The free energy is purely entropic ( $U = 0$ ) because the random chain does not interact with itself. The force for stretching the DNA calculated from

length	dependence	DNA (mm)
$R_{max}$	$bN$	50
$R_{coil}$	$bN^{3/5}$	0.2
$R_{ave}$	$bN^{1/2}$	0.07
$R_{cmp}$	$bN^{1/3}$	0.008

Table 1.1: **End to end distances of a DNA molecule.** The molecule is modeled as a chain of  $N = 500000$  steps of Kuhn length  $b = 100$  nm, equal to twice the persistence length (50 nm) in a worm like chain model.  $R_{max}$  is the maximum elongation of a typical human DNA molecule [7];  $R_{coil}$  is the good solvent average;  $R_{ave}$  is the random walk average;  $R_{cmp}$  is the size of a fully packed chain of impenetrable *beads* with diameter  $b$ .

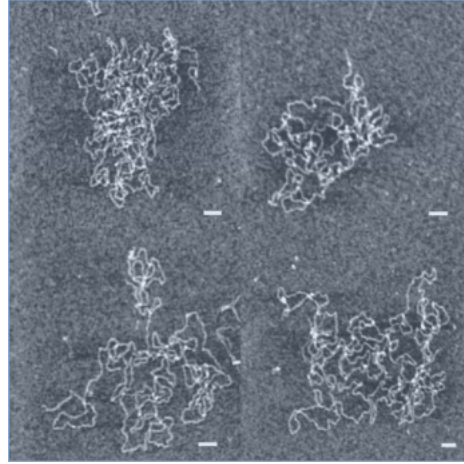


Figure 1.1: **Images of DNA.** Atomic Force Microscopy images of DNA fragments on a surface, from ref. [70]. Each fragment is less than 50000 base pairs long, i.e. about 17000 nm or 340 persistence lengths. The scale bar represents 100 nm, equal to twice the persistence length. The experimental average end to end distance is calculated on this kind of images.

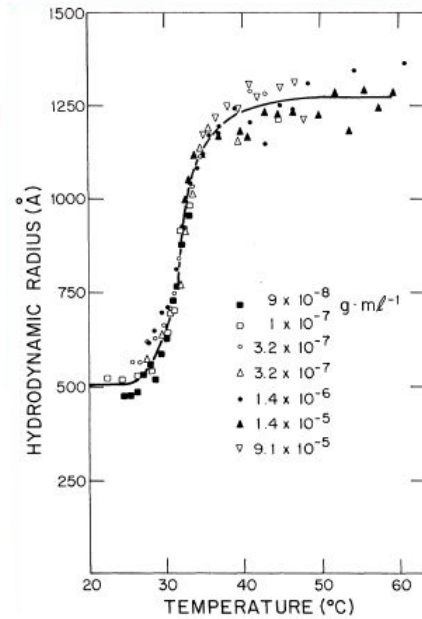


Figure 1.2: **Coil to globule transition of polystyrene.** The temperature dependence of the hydrodynamic radius of polystyrene in cyclohexane is plotted for several concentrations. The length of the molecule is about 260000 monomers ( $2.7 \times 10^7 M_W$ ). From ref. [64]. The hydrodynamic radius, or Stokes radius, is the radius of a sphere with the same diffusion coefficient as the polymer.

equation (1.6) gives results in very good agreement with experiments, at least in the early stages of elongation. This force is entropic in origin, ultimately a consequence of the Brownian impulses acting on the polymer that tend to keep it coiled.

The study in [7] is pursued to the limit of maximal stretching. In this regime the theory presented here is not sufficient any more: other effects, namely the bending energies, has to be added to the model. The study goes even further, describing the experiments of pulling the molecule beyond its maximal elongation: the subject is very interesting, but this is not the place for developing it.

### 1.3 Coil to globule transition

The conformational entropy plays a key role in all situations where the macromolecule is free to explore many different arrangements of its components. The most *natural* aspect of a polymer in a good solvent is that of a disordered coil, similar to the average random walk calculated in the previ-

ous section. In this situation, if the excluded volume of monomers is also taken into account, the average end to end distance results bigger than the random walk. The self avoidance acts as a repulsion between the monomers and makes the molecule to swell: Table 1.1 reports the metric exponent of the coil calculated from the standard Flory theory [30, 58].

Nevertheless, polymers in a compact conformation are observed, depending on the quality of the solvent. In general, if two molecules come closer, the van der Waals attraction manifests, so the system can lower its energy if they stay *in contact* (the excluded volume prevents them to overlap). If the attraction between a monomer and a solvent molecule is stronger than the attraction between two monomers, the polymer maximizes its contacts with the good solvent assuming an open conformation. On the opposite hand, if the monomer-monomer interaction is stronger, the molecule tends to maximize its intrachain contacts and it assumes a globular conformation which excludes the bad solvent. These interactions can be as strong as a hydrogen bond, which is responsible of the formation of secondary structures (a particularly ordered compact conformation) of proteins. As these interactions may develop both between monomers and between monomer and solvent, the overall conformation is the result of the interplay between many factors and no interaction can be excluded from the calculation. In proteins, however, a decisive factor for the folding is believed to be the hydrophobicity of certain amino-acids. An hydrophobic amino-acid tries to avoid contact with the solvent water hiding itself in the bulk of the collapsed protein [20].

In the process of collapse, a favorable negative energetic difference between the compact and coil conformation,  $H_{cmp} - H_{coil} = \Delta H < 0$ , has to balance the unfavorable loss in conformational entropy,  $S_{cmp} - S_{coil} = \Delta S < 0$ . The two contributions are combined in the free energy difference between the compact and coil state, at temperature  $T$  and constant pressure:

$$G_{cmp} - G_{coil} = \Delta G = \Delta H - T\Delta S. \quad (1.7)$$

Since the state of the molecule corresponds to the state of minimum free energy, a variation of the temperature induces a change of the conformation of the molecule when the sign of  $\Delta G$  switches from positive to negative. This phenomenon is known as Protein Folding (from coil to globule) or Denaturation (from globule to coil) or simply coil-to-globule transition, referring to generic polymers. The transition is located at the temperature where the two states have the same free energy,  $\Delta G = 0$ . Supposing that the entropy and energy difference between the two states is constant, the critical temperature is

$$T_c = \frac{\Delta H}{\Delta S}. \quad (1.8)$$

This is a very simple sketch of the phenomenon and does not tell anything about the kind of transition. For example, the molecule may conserve its



compact form while the temperature grows until the critical point, where it would jump directly to the coil state: at the critical temperature the two opposing phases (the coil and the compact) coexist. In an opposite picture, the molecule gradually evolves from one state to the other; at the critical point no hint of the two opposing phases can be seen, the molecule assumes instead an intermediate structure. The first scenario corresponds in physics to a discontinuous transition, also called *all-or-none* or first order. The second scenario corresponds to a continuous or second order transition.

Since the seminal experiments on the native state of proteins [4], protein folding is commonly described as an *all-or-none* phenomenon. This fact is of vital importance in living systems where protein misfolding (collapse in the *wrong* compact state) makes it lose its functionality. Instead, in the generic case of homopolymers, made by a single repeating unit, the phenomenology is less clear: if sometimes signatures of the continuous behavior are seen [54, 76], in many cases the phenomenon is described as a discontinuous process [74, 69]. The experimental probe of one or the other phenomenology is limited by the difficulties common to the single molecule experiments and multiple behaviors can not be excluded.

The first temperature induced coil-globule transition of a single polystyrene chain (a homopolymer) was reported in [64]. The hydrodynamic radius of the molecule is measured by light scattering technique, see Figure 1.2.

## 1.4 The $\Theta$ point

The essential factors playing in the coil to globule transition of the polymer are:

1. randomness in the configuration of the molecule;
2. excluded volume interaction;
3. short range attraction between monomers.

The randomness works in favor of the coil conformation, which has more arrangements at disposal than a globular molecule. The attraction works in favor of the globular conformation, which has more intra-chain contacts. The excluded volume contributes further to the swelling of the chain but it prevents also the compact molecule to collapse into a unique point under the effect of its internal attraction. These ingredients are contained in a basic model, consisting of a Self Avoiding Walk on a lattice with nearest neighbor attractions. The energy of a walk is proportional to the number of intrachain contacts

$$E = -\varepsilon n_{co} \quad (1.9)$$

(Figure 1.3): it is then energetically convenient to assume a compact conformation, but the temperature has to be low enough to face the swelling

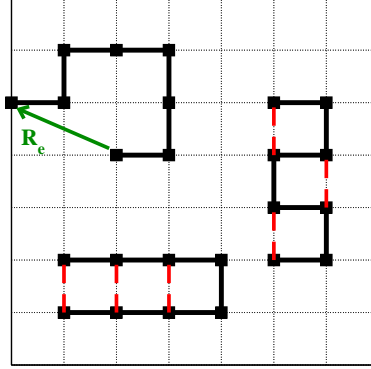


Figure 1.3: **ISAWs on the lattice.** Three Self Avoiding Walks with 8 monomers on a square lattice. Two of them have 3 contacts, one of them has no contacts. One end to end distance vector is drawn

tendency coming from the conformational entropy. In this model, Interacting Self Avoiding Walk (ISAW), the critical temperature at which the coil changes to the compact conformation is called  $\Theta$ . A precise definition of a compact and a coil state is not possible here, so the definition (1.8) of the critical temperature does not make sense. The temperature corresponding to the maximum variation of the energy or the radius is the finite size critical temperature. The nature of the  $\Theta$  transition has been long investigated through several complementary approaches: a huge review of these theoretical efforts is contained in [19]. It is commonly accepted that the  $\Theta$  point is a tricritical phenomenon, which means, in the words of Des Cloizeaux, that it is two times (not three!) critical: once because the length of the chain is infinite, once because the critical temperature corresponds to the balance between the entropic push towards a coiled state and the internal attraction towards a compact conformation. We try to describe these two criticalities on the lattice ISAW model.

A deep link exists between this model and a generalization of the Ising model of the ferromagnetic transition. In the Ising model a magnet is represented as a lattice of 1-component spins ( $S_i = \pm 1$ ) and the critical temperature separates the paramagnetic state (high temperature) from the ferromagnetic state (low temperature). The Ising model has been generalized to  $n$ -components spins. In [17] it is shown that the magnetic susceptibility of the  $(n \rightarrow 0)$ -components model ( $O(0)$ ), close to the critical temperature (of the magnet), can be calculated as the sum of all the SAWs of any length (from 1 to  $\infty$ ), each weighted by a power of the inverse temperature:

$$\chi = \sum_{N=1}^{\infty} Z_N K^N \sim (K_c - K)^{-\gamma} \quad K \rightarrow K_c. \quad (1.10)$$

The coefficient  $Z_N$  is precisely the number of SAWs of length  $N$  and  $K_c$  is the critical inverse temperature of the magnetic model (note that at this stage the temperature of the polymer has not been introduced yet). As it is customary in magnetic models, the critical exponent  $\gamma$  describes the divergence of the susceptibility close to the critical point. It is also customary to approximate the number of SAWs of length  $N$  with the following expression:

$$Z_N \sim \mu^N N^{\gamma-1} \quad (1.11)$$

where  $\mu$  is a constant depending on the lattice, while  $\gamma$  is a universal exponent, depending only on the dimension. Substituting equation (1.11) into the sum in equation (1.10) one finds that the two critical exponents are actually the same. The connectivity constant is related to the critical magnetic temperature,  $\mu^{-1} = K_c$ , which also depends on the type of lattice. Thanks to the identification between SAWs and magnetic models, other critical exponents can be calculated; in particular the metric exponent  $\nu$  is related with the divergence of the correlation length in the magnetic model. Their values are obtained on the basis from the mean field theory if the dimension of the lattice is  $d \geq 4$  (a Landau free energy to the 4th order:  $F_L = am^2 + bm^4$  with varying  $a$ ), while in the interesting cases ( $d = 3, 2$ ) they can be calculated through an  $\epsilon$  expansion technique [71].

This result shows that the ensemble of (non interacting!) SAWs of infinite length has critical exponents corresponding to the critical point of the  $O(0)$  magnetic model. In the Interacting SAW a second parameter, the temperature of the polymer itself, is added and brings an ulterior criticality which is manifested at the  $\Theta$  temperature *and* infinite length. A system with two critical parameters can be described through a mean field theory based on a Landau free energy function of the 6th order,  $F_L = am^2 + bm^4 + cm^6$ , with two varying parameters,  $a$  and  $b$ . The tricritical point corresponds to the meeting point of a line of first order phase transitions with a line of second order phase transitions, which happens at the point  $a = b = 0$ . In this case, the mean field results are valid also at dimension 3, a part from strong logarithmic corrections. In 2 dimensions (where the mean field theory is no more valid) the critical exponents of the  $\Theta$  point have been exactly calculated showing the equivalence with SAWs on a hexagonal lattice with random forbidden hexagons [27]. This equivalence were attested not without some debate: firstly, the difference of this model with the original SAW on the square lattice [55], but also the discrepancy between theoretical and numerical results concerning some entropic exponents [60, 61] were the reasons of the doubt about the identity of the universality class of the two models. The puzzle of the  $\Theta$  and  $\Theta'$  point was then solved with the correct interpretation of the calculated entropic exponents [72]. The theoretical results were confirmed by successive numerical works on very long chains [35, 34].

The extreme complexity of a deep understanding of this subject is discouraging. However, for our concerns, we retain that the coil to globule

Exponent		definition	2 dim	3 dim
metric	$\nu$	$R \propto N^\nu$	4/7	1/2
cross-over	$\phi$	$\mathcal{C} \propto N^{2\phi}$	3/7	1/2
entropic	$\gamma$	$Z \propto \mu^N N^{\gamma-1}$	8/7	1

Table 1.2: **Critical exponents of the  $\Theta$  point.**

transition of an ISAW on the lattice is a continuous transition. The critical temperature is manifested by a (very) slight divergence of the specific heat, in 3 dimensions,  $c = \frac{1}{N} \frac{d\langle E \rangle}{dT}(T_\Theta) \sim (\ln N)^x$ . In 2 dimensions there is no such divergence ( $\phi < 1/2$ ) and the criticality is manifested in higher order derivatives of the energy.

## 1.5 Variations on the ISAW

The Interacting Self Avoiding Walk contains the strictly necessary ingredients for observing a coil to globule transition. Variations on this model are easy to cook up; in this section we present a few of them, which are interesting because they present a more complex behavior. In particular in all of these variants the coil to globule transition splits to a classic continuous transition (melting) and a sharper low temperature first order transition (crystallization). These variants were introduced with explicit reference to the phenomenology observed in proteins, where the folding is notoriously an all-or-none transition, but also in view of the controversies about the homopolymer behavior [15].

### 1.5.1 Model with stiffness

In the lattice model of a semiflexible homopolymer [23], a stiffness energy term proportional to the number of turns is added to the contacts energy. The energy of a chain is then given by

$$E_{stiff} = -\varepsilon n_{co} + \varepsilon_k n_{kink}, \quad (1.12)$$

with both positive constants  $\varepsilon_k$  and  $\varepsilon$ : the more a walk is straight, the lower its stiffness energy and the higher its contact energy is. On the other hand, kinks are unavoidable in the compact conformations which minimize the contacts energy (increasing the number of contacts), but they raise stiffness energy. The competition between these two factors makes the collapse a two stage process: lowering the temperature, the polymer undergoes first the *standard* coil to globule  $\Theta$  transition (second order) and afterwards a discontinuous freezing transition from the *liquid* globule to a crystalline compact state. The crystalline state of minimum energy looks like a cuboid

extended in one direction with an aspect ratio which increases with increasing  $\varepsilon_k/\varepsilon$ . For very high values of the stiffness constant, ( $\varepsilon_k/\varepsilon > 10$ , [24]), the freezing first order transition precedes the  $\Theta$  collapse, and the molecule crystallizes directly from a coil to the crystal phase.

The stiffness may be related to the formation of secondary structure, preliminary to the collapse in the native state of protein, or to a structural stiffness as in the DNA molecule.

### 1.5.2 Bond fluctuation model

In the bond fluctuation model [56], the monomers occupy the sites of a cubic lattice and they are connected by bonds of variable length and direction: two successive monomers are separated by a distance which can vary up to a threshold. Apart from that, the model is similar to the ISAW: the monomers interact through excluded volume (they can not occupy the same site) and a short range attraction (a cut-off distance is defined). Similarly to the semiflexible homopolymer, in this model the collapse happens also in two steps: first the continuous coil to globule transition, then the first order crystallization transition. In the bond fluctuation model, the two transitions coincide in the thermodynamic limit, where the chain collapses directly from the coil to the crystal compact state.

This phenomenology is observed also in an off-lattice model consisting of a freely jointed chain of beads with excluded volume and short range attraction interactions [78]. In this case the collapse of the two transitions (direct freezing from the coil to the ordered compact state) happens if the cut-off distance of the contacts attraction is below a critical value.

### 1.5.3 HP model

The HP model was introduced to study the relationship between the amino acid sequence of a protein and its native structure [46]. It consists of a SAW on the lattice with excluded volume and a binary label (H or P) attributed to each monomer. The model is based on the assumption that the Hydrophobic effect is the leading force in protein folding [20]. Indeed, two hydrophobic amino-acids tend to be buried inside the molecule in order to escape contact with the solvent, which would cause an entropically unfavorable ordering of the solvent molecules. Two polar molecules instead gain energy by being in contact with the solvent. This effect is synthesized in the HP model defining the energy of a walk proportional to the number of H-H contacts,  $E = -\varepsilon n_{HH}$ . The ISAW corresponds to the situation of a chain of only H monomers (homopolymer). An important factor introduced by the sequence is the decrease of number of low energy configurations: in order to model protein folding effectively a sequence should be chosen to have a unique native state, but this is difficult to verify for long sequences. Computational

studies of the thermodynamic of the HP model of long chains in 2 and 3 dimensions [14], [8] reveal that the folding of the HP model takes place in two steps: a continuous coil to globule transition precedes the crystallization to the unique native state. The small number of states close to the minimum energy is a characteristic feature of the model. The analysis of short HP chains [21] reveals that an energy gap between the native and unfolded state is sufficient to create a discontinuous phase transition. In general, the necessary condition is that the logarithm of the density of states is concave upward (plotted against the energy) close to the transition, which is a less rigid condition than the energy gap. Evolution may have chosen HP sequences in order to provide that energy gap.

## 1.6 Studies on conformational entropy

The example of polymer stretching (section 1.2) shows that a simple representation of the conformational freedom (entropy) of the molecule, modeled as a freely jointed chain, is sufficient for the study of the phenomenon. This is not always the case. The coil to globule transition lies on a subtle equilibrium between many forces, ranging from electrostatic interactions to van der Waals and hydrogen bonds, from *intrinsic propensities* to hydrophobic effects [20]. Conformational entropy induces an opposing force to folding, because a compact chain has less available configuration than an open chain. Thermal denaturation is mainly driven by the raising of the entropic part in free energy (see equation 1.7). The sources of conformational entropy splits into two parts [44, 43]: one corresponds to the existence of many structures (non-local entropy), and the other corresponds to the local fluctuations around such structures (local-entropy). If the non-local entropy clearly competes against the folding transition, the local entropy can have ambiguous roles. In some studies [43] it is considered to be irrelevant because more or less independent on the structure; as only differences in entropy may have an effect on the thermodynamics of the system, such a behavior would not affect the physics of the molecule. In other studies local contributions to the entropy are presented as stabilizing factors of the folded state: this is the case, for example, of lysine-arginine substitution in the folded state of thermophilic proteins [10], the first having more accessible rotamers than the second one; an other example shows that some compact structures may be stabilized by the possibility of a better arrangement of the side chains, gaining in mobility, which increases the total entropy of the folded state and contributes to stabilize it [77].

In lattice models, multiplicity of configurations represents the non-local contribution to the entropy well, but the local fluctuations are completely neglected. In this thesis we propose that the local entropy affects the nature of the transition in a considerable way. Our results show that its presence

contributes to enhance the cooperativity of the coil to globule transition. If this contribution was never studied on the coil to globule lattice model, in other fields similar phenomena have been treated.

The change in internal dynamics of proteins when they bind a ligand has been investigated experimentally by nuclear magnetic resonance spectroscopy [31]. The internal dynamics is a proxy for the local conformational entropy, as the protein conserves its global structure. The study shows that conformational entropy contributes significantly to the free energy of the protein-ligand association.

A variation of the Ising model itself has been elaborated in order to describe the effects of the elastic degrees of freedom on the magnetic properties [11]. The elasticity is introduced in the Hamiltonian of the standard Ising model as an additional coupling between nearest neighbors depending on the instantaneous separation vector (as opposed to the equilibrium or rigid lattice separation vector). As a consequence, the exactly soluble magnetoelastic lattice shows that the elastic degrees of freedom induces a first order magnetic transition which takes the place of the usual second order phase transition. In the study of martensitic transitions [50], it has been shown that the vibrational entropy stabilizes the high temperature phase (metastable) with respect to the low (stable) temperature phase because the softer high temperature phase presents low frequency phonons which increases the entropy. The Hamiltonian of the model with vibrational entropy differs from the traditional Hamiltonian for a anharmonic coupling between nearest neighbors: a harmonic coupling would not be enough because the coupling does not change at the transition. As a consequence the already first order transition becomes stronger and the critical temperature lowers.

## 1.7 Summary

This thesis proceeds as follows. In chapter 2 the model is presented in detail. An analytic solution of the model is not reliable at present, but we put a lot of effort in the numerical investigation: the algorithms are presented in the same chapter. It comes out that the conformational entropy of this model can be calculated through normal modes analysis: chapter 3 presents the typical normal mode spectrum of the Self Avoiding Walks. The analysis reveals that the coil state is rich in soft modes, more than the compact state. This observation gives a partial explanation of the coil to globule transition as it appears in this model. Chapter 4 and chapter 5 describe all the results obtained by the simulations and some analytical argument. In particular, in three dimensions the standard continuous transition transforms into a first order transition under the influence of vibrational entropy. We also study the transition in two dimensions and present preliminary results which encourage further studies, in particular on the formation of secondary structure

and the phenomenon of adsorption. Finally, we discover that a characteristic universal exponent of the transition can be experimentally deduced from measures of intrinsic viscosity: we present the theory and results in chapter 6, before the conclusions.



## Chapter 2

# A lattice model with vibrations

In the classic Interacting Self Avoiding Walk the conformational entropy is sourced by the amount of lattice configurations. Representing each Walk as an Elastic Network, we define a more complex model in which both global motion and local vibrations are present. Through some mathematics we keep the model on the lattice and the usual investigation methods (in particular numeric simulations, section 2.7) are applicable. The classic model and the model decorated with vibrations can be interpreted as successive approximations of the potential function of the molecule (section 2.4). The most technical aspects of the definition are contained in section 2.5 and in the Appendix.

### 2.1 Lattice models and elastic networks

In this work we synthesize two different and somewhat opposite approaches to the description of a macromolecule. In one of these, a molecule is modeled as a Self Avoiding Walk on a lattice; this is a global and generic description of the macromolecule in all its possible conformations ranging from the completely stretched to the globular shape and it is independent of the chemical structure of any real molecule. In the second approach, the molecule is depicted as an Elastic Network; this model is based on experimentally detected structures and, despite the highly simplified representation of the interactions, it is able to reproduce the molecular motion of the molecule to the details of the single monomer.

The Self Avoiding Walk on the lattice (SAW) reproduces the variety of shapes of any object constituted by a big number of chained components, which is one essential feature of polymers. The model consists of a sequence of points (*monomers*) linked by identical bonds (*steps*). The position of the monomers is restricted to the sites of a regular lattice. Each lattice site has

the same number of nearest neighbors as any other site (the *lattice coordination number*) and the distance between two nearest neighbors is fixed and equal to the *lattice spacing*. Since two atoms can not occupy the same position, the monomer chain can not visit an already visited point. Nevertheless the chain is allowed to touch itself, which happens whenever two monomers are positioned on nearest neighbor lattice sites but are not consecutive along the chain. In this case the monomers are said to be *in contact*. The number of contacts is an important quality of the SAW: in the coil to globule transition model the energy of a walk is negative and proportional to the number of contacts. The walk on the lattice may seem a very crude approximation, as it limits the directions of an object, living naturally in the continuous space, to a discrete set of possibilities. Apart from rendering calculations much simpler, this choice is motivated by the observation that the chemical bonds have a few number of preferred directions or angles between the atoms, which corresponds to states of minimum energy and are indeed much more represented than the other possible arrangements. For example, the four successive  $\text{CH}_2$  groups composing the butane molecule are found more likely in three dispositions (instead of the continuous possible rotations), the *trans*, *gauche* and *gauche'*, in which the dihedral angle formed by the three bonds is  $\pi$ ,  $\frac{\pi}{3}$  and  $-\frac{\pi}{3}$  respectively (Figure 2.1). In general the alkanes, polymers with a backbone made by Carbon atoms, are a random sequence of *trans* and *gauche* configurations [51]. Also in a protein, the amino-acids favor some combination of successive dihedral angles over the whole rotational availability, as a Ramachandran plot typically shows [75]. On top of that, the covalent bond between carbons has a fixed length, of order the Angstrom. The backbone conformation of a macromolecule resembles a self avoiding zigzag and is indeed more similar to a walk on a lattice than to a continuous line or a freely jointed chain. The lattice approach was primarily developed by De Gennes [17] in the late Seventies; From then on, the SAW has been extensively studied as a prominent model of polymers. The lattice model is able to reproduce the metric exponents, to represent the polymer close to an obstacle or a surface and to sketch the conformational change at the coil to globule transition [71]. In general the lattice model establishes a framework which is shared by all macromolecules, on top of which each specific molecule brings its own peculiarity depending on its intimate structure. In other words this is the context for studying the universal properties of polymers.

The use of self avoiding walks as a model of molecules provided of a very complex chemistry can be better understood considering a concept already introduced in section 1.2: the persistence length. In that section, we have shown that the shape of one DNA molecule can be traced out with a chain of uncorrelated bonds of length bigger than the persistence length of the polymer. The simple sketch works because the persistence length is equal to the length of a section of a chain whose ends are statistically uncorrelated.

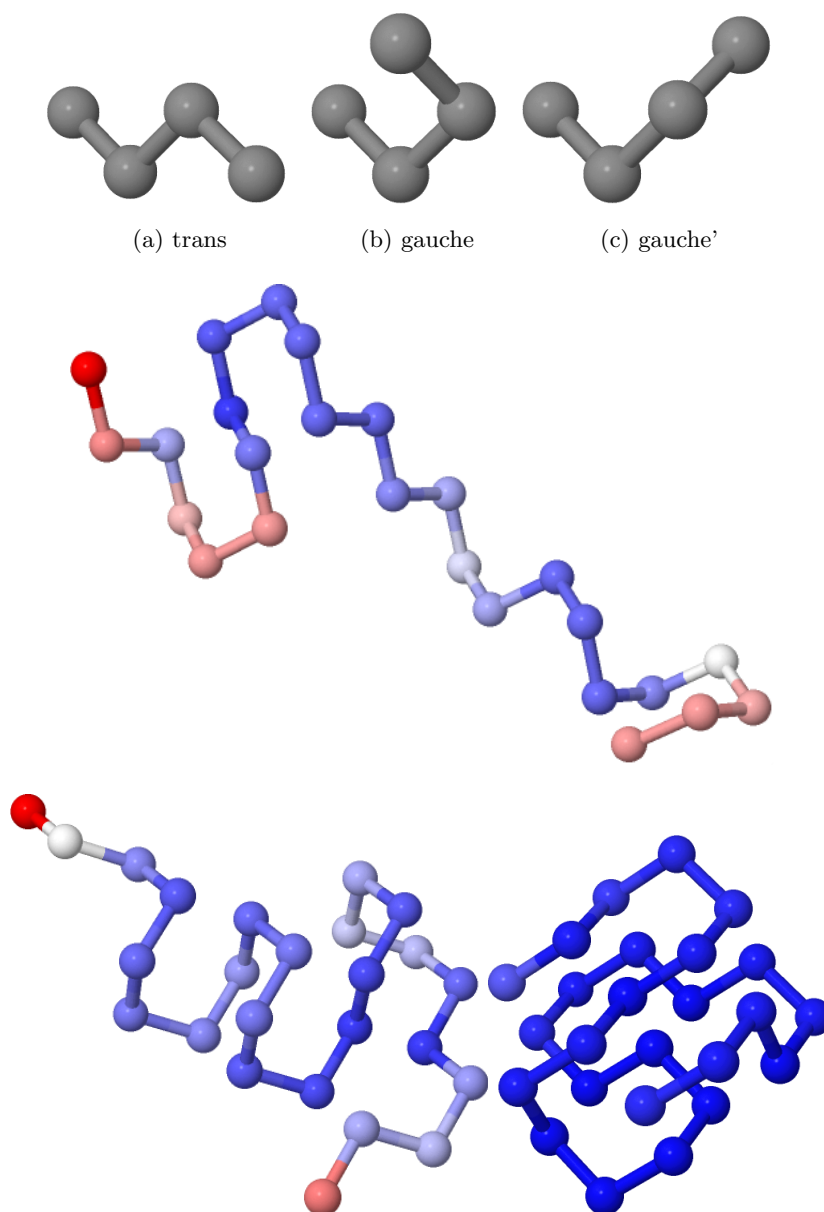


Figure 2.1: **Self Avoiding Walks on the tetrahedral lattice.** **Top:** four successive monomers in the *trans*, *gauche* or *gauche'* configuration. **Bottom:** three walks of 25 monomers and 3, 5, 11 contacts respectively. The colors represent the B-factors in an absolute blue-red scale.

We may apply the same reasoning to any polymer as well as to walks: irrespective to the nature of the local interactions, angle constraints and so on, the same description can be applied to a plethora of chains. The description changes only if a long range interaction, acting beyond the limits of the persistence length, is neglected. An important example of such interaction is the self avoidance: the self avoiding walk indeed belongs to a different universality class than random walk (without self avoidance). In general, the so called universal properties depend on the long range interactions, but they are independent of the microscopic properties, such as the nature of the lattice. This is a common aspect in critical phenomena, where a system manifests a long range correlation. In this case, again, the behavior of the system do not depend on the short range particular interactions: a typical approach consists in the coarse-graining of the short-distance degrees of freedom by introducing a description in terms of a new parameters which keep unchanged the physical aspect of interest. This general approach is at the basis of the lattice walk description of polymers.

In a somewhat opposite approach, the Elastic Network model (EN) considers the crystallography structure of a specific molecule. This consists essentially in the space coordinates of all the monomers detected when the molecule is in one specific compact conformation. The interactions between monomers are represented as springs connecting pairs of nodes located within a certain *cut-off* distance. In this way, the crystallography structure represents a configuration of minimum energy and the monomers oscillate around their equilibrium position. The network of the interactions is built on the basis of a *cut-off* distance,  $c^0$ : two nodes interact if their distance is shorter than  $c^0$ . In that case a spring of constant  $k$  connects the nodes, meaning that their distance  $d_{i,j}$  fluctuates around the equilibrium position  $d_{i,j}^0$  with energy<sup>1</sup>

$$V_{i,j} = \frac{1}{2}k(d_{i,j} - d_{i,j}^0)^2. \quad (2.1)$$

The total energy of the network is the sum of the elastic interactions (over all the interacting pairs). The harmonic approximation of the total energy is

$$V(\Delta\vec{r}) = \sum_{i,j}^{d_{i,j}^0 < c^0} V_{i,j} \sim \frac{1}{2}\Delta\vec{r}H\Delta\vec{r}, \quad (2.2)$$

where the vector

$$\Delta\vec{r} = (\Delta x_1, \Delta y_1, \Delta z_1, \Delta x_2, \Delta y_2, \dots)$$

represents the displacements of the monomers from the equilibrium position

---

<sup>1</sup>This is the *anisotropy* network model interaction [6].

Table 2.1: **Table of symbols**

monomer position	$\vec{r}_i$	$i = 1...N$
equilibrium monomer position	$\vec{r}_i^0$	$i = 1...N$
bond coordinates	$\vec{R}_i = \vec{r}_{i+1} - \vec{r}_i$	$i = 1...N - 1$
lattice spacing	$l$	
euclidean formula	$ \vec{r}_i  = \sqrt{(x_i)^2 + (y_i)^2 + (z_i)^2}$	

and  $H$  is the Hessian matrix

$$H_{ij} = \left. \frac{\partial^2 V}{\partial r_i \partial r_j} \right|_{\Delta \vec{r}=0} \quad (2.3)$$

calculated at zero displacement. The square root of each eigenvalue of the Hessian is the frequency of one normal mode:<sup>2</sup>

$$\omega_i = \sqrt{\lambda_i}. \quad (2.4)$$

Proceeding from the normal mode analysis, it is possible to calculate the B-factor of each monomer, which is a quantity proportional to the mean square fluctuation:

$$B_i = \frac{8\pi^2}{3} \langle (\Delta \vec{r}_i)^2 \rangle = \frac{8\pi^2}{3} \langle (\Delta x_i)^2 + (\Delta y_i)^2 + (\Delta z_i)^2 \rangle \quad (2.5)$$

and is at the same time a quantity measured by x-ray crystallography. Several applications have demonstrated that the fluctuations predicted by the EN model are in good agreement with experimental B-factors [16] and encourage us to continue on that way. In Figure 2.1 we illustrate three SAWs on the lattice and the B-factors calculated on the basis of the Elastic Model. In analogy with real polymers, regions with higher mobility (rose or red colors) are usually situated in the exterior of the molecule, while monomers in the bulk are constrained by many neighbors and are less mobile.

## 2.2 The partition function

The lattice model of the collapsing polymer is a SAW on the lattice with monomer-monomer attraction: for this reason it is called Interacting Self Avoiding Walk (ISAW). The energy of a walk with no contacts is maximum and equal to 0, otherwise it is negative and proportional to the number of contacts:

$$E = -\varepsilon n_{co}. \quad (2.6)$$

<sup>2</sup>The masses of the monomers are assumed to be identical and unitary. A more general treatment introduces mass-weighted coordinates [16].

The energy unit  $\varepsilon > 0$  is generally set to 1. Two monomers are in contact when they are not consecutive along the chain and are placed on nearest neighbor sites of the lattice. The canonical partition function (fixed length  $N$ ) of the ISAW is the sum over all the walks on the lattice:

$$Z_{ISAW} = \sum_{\gamma} e^{-\beta E(\gamma)}. \quad (2.7)$$

The number of contacts is the most important observable of the model because of some characteristics:

1. it is proportional to the energy;
2. it is a good descriptor of the geometry of the walk, a compact walk having many contacts (low energy), a coil having few (high energy);
3. it is an integer quantity which is easy to be treated.

An alternative way of writing the partition function is the sum over all the energies:

$$Z_{ISAW} = \sum_E e^{-\beta[E - T \ln \mathcal{N}(E)]} = \sum_E e^{-\beta[E - TS_{isaw}(E)]} \quad (2.8)$$

where the traditional entropy function  $S_{isaw}(E) = k_B \ln \mathcal{N}(E)$  is the natural logarithm of the number of configurations with energy  $E$ .

In our work, the lattice model is integrated with the dynamics of the small fluctuations, continuously perturbing the single monomer positions while leaving intact the global shape of the molecule. In order to do that, we describe a SAW as the minimum energy configuration of an elastic network and we add the contribution of small displacements from the equilibrium position,  $\Delta \vec{r}$ , to the contact energy of the walk:

$$V(\gamma, \Delta \vec{r}) = E(\gamma) + \frac{1}{2} \Delta \vec{r} H(\gamma) \Delta \vec{r}. \quad (2.9)$$

This relation is essentially equal to the elastic network potential (equation 2.2), except that the energy at zero displacement depends on the lattice configuration  $\gamma$  and is equal to its contacts energy  $E(\gamma)$ . In analogy with the EN model, the Hessian matrix is defined on the basis of the topology of the walk. The details of its definition involve some mathematical effort and are postponed to the section 2.5; hereafter we describe how the contribution of fluctuations is effectively included in the free energy of the single walk as an entropic term. The contact+elastic energy, equation (2.9), is defined on positions in the continuous space, so the partition function is to be integrated

over the lattice configuration and the small displacements from it:<sup>3</sup>

$$Z_{VIBR} = \sum_{\gamma} \int_{I(\gamma)} \left( \frac{d(\Delta\vec{r})d\vec{p}}{h^{3N}} \right) e^{-\beta \left[ \sum_{i=1}^N \frac{p_i^2}{2m} + V(\gamma, \Delta\vec{r}) \right]}. \quad (2.10)$$

The integration over the momentum coordinates results in a constant factor, independent on the configuration, hence irrelevant for the thermodynamics of the model. The integral over the space coordinates is a multi-dimensional Gaussian integral in the variables  $\Delta\vec{r} \in \mathbb{R}^{3N}$ . Since the Hessian matrix is semi-positive definite, a transformation from the Cartesian to the external+internal coordinates of the molecule has to be performed. The external coordinates describe the translation/rotation of the molecule as a rigid body; this is completely described by 3 coordinates in 2 dimensions and 6 coordinates in 3 dimensions, so there are  $2N - 3$  ( $2d$ ) or  $3N - 6$  ( $3d$ ) internal coordinates,  $\Delta\vec{q}$ . Expressing the Hessian matrix in terms of the new coordinates, the elastic part of the energy is  $\frac{1}{2}\Delta\vec{q}H^*\Delta\vec{q}$  where  $H^*$  is the positive definite sub-matrix corresponding to the internal coordinates only. Integration over external coordinates gives the constant contribution associated to the point motion in the space, while the integration over the internal coordinates gives the interesting part:

$$\begin{aligned} \int d(\Delta\vec{r}) e^{-\beta V(\gamma, \Delta\vec{r})} &= \int d(\Delta\vec{r}) e^{-\beta (E(\gamma) + \frac{1}{2} \Delta\vec{r} H \Delta\vec{r})} \\ &= C e^{-\beta E(\gamma)} \int d(\Delta\vec{q}) e^{-\beta \frac{1}{2} \Delta\vec{q} H^* \Delta\vec{q}} \\ &\propto e^{-\beta E(\gamma)} \prod_{\lambda_i > 0} \lambda_i^{-1/2}(\gamma). \end{aligned} \quad (2.11)$$

Note that the coordinate transformation is a formal passage: all what we need are the eigenvalues of  $H^*$  which are equal to the not null eigenvalues  $\{\lambda_i > 0\}$  of the Hessian  $H$ . It is important also to note that both the contact energy (2.6) and the fluctuation weight

$$W(\gamma) = \prod_i \lambda_i^{-1/2}(\gamma) \quad (2.12)$$

depend on the walk configuration. Within our approach, the fluctuations enrich the standard partition function with a weight of entropic nature:

$$Z_{VIBR} = \sum_{\gamma} W(\gamma) e^{-\beta E(\gamma)} = \sum_{\gamma} e^{-\beta [E(\gamma) - T \ln W(\gamma)]}. \quad (2.13)$$

This formulation, consisting in a partial integration (over the fluctuations) of the partition function, is interesting because the sum is performed over

---

<sup>3</sup>The limits of integration consist of a small neighborhood of the lattice configuration, but the calculation of the Gaussian integral is done on the whole space  $(-\infty, \infty)$  in order to have an exact result. A justification of this is presented in the following section 2.3.

the lattice walks and the difference between the two models is highlighted: the *vibrational entropy* of the single configuration ( $\ln W$ ) is included in its free energy, while in the standard model it is not considered at all, which corresponds to treating it as uniform over all the configurations. In terms of the contact energy, the partition function looks like (in analogy to equation (2.8)):

$$Z_{VIBR} = \sum_E e^{-\beta[E-T(S_{isaw}(E)+S_{vib}(E))]} \quad (2.14)$$

where the microcanonical vibrational entropy function is the logarithm of the average entropic weight at fixed contact energy:

$$S_{vib}(E) = \ln \frac{\sum_{\gamma|E} W(\gamma)}{\mathcal{N}(E)} = \ln \langle W \rangle_E. \quad (2.15)$$

### 2.2.1 Relation between B-factor and vibrational entropy

The B-factors of a SAW with vibrations can be calculated from equation (2.5) and the eigensystem of the Hessian matrix. From the properties of the multidimensional Gaussian distribution, we have:

$$\langle (\Delta x_i)^2 \rangle = k_B T \sum_{k=1}^{3N-6} \lambda_k^{-1} (u_{ix}^k)^2, \quad (2.16)$$

where  $\vec{u}^k$  is the eigenvector of  $H$  corresponding to the eigenvalue  $\lambda_k$ . The sum is performed on the  $3N - 6$  not null eigenvalues. Since the eigenvectors are normalized

$$\sum_{i=1}^N \sum_{x,y,z} (u_{ix}^k)^2 = 1, \quad (2.17)$$

the sum of the B-factors of the  $N$  monomers is proportional to:

$$\sum_{i=1}^N \langle (\Delta \vec{r}_i)^2 \rangle = k_B T \sum_{k=1}^{3N-6} \lambda_k^{-1} \quad (2.18)$$

and presents some resemblance with the vibrational entropy of the single configuration:

$$\ln W = \frac{1}{2} \sum_{k=7}^{3N} \ln \lambda_k^{-1}. \quad (2.19)$$

The similarity lies in the fact that both the quantities are principally affected by the slow motion modes (low  $\lambda$ s) which represent the broadest fluctuations. Both the quantities are expression of the total mobility of the molecule, but looking at the configurations of Figure 2.2 we can verify that a positive



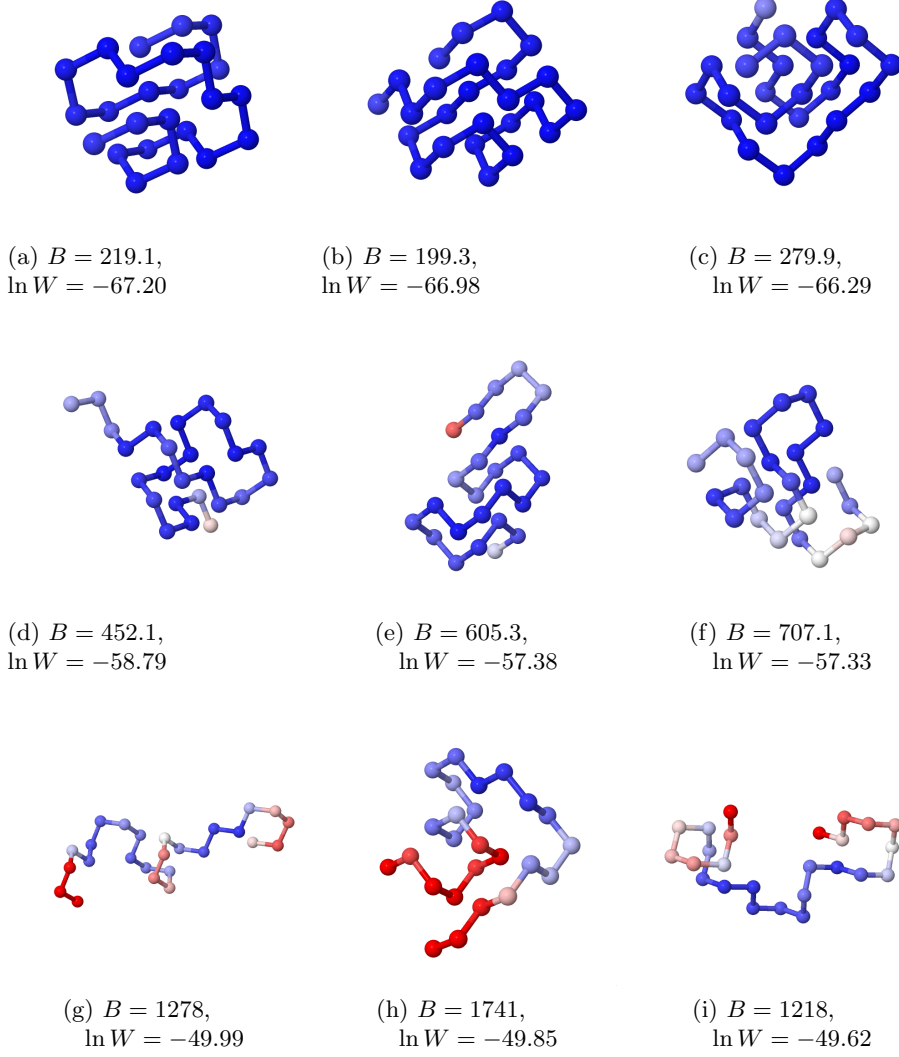


Figure 2.2: **Configurations, B-factor and Vibrational entropy.** Some ISAWs of length 25 monomers. **Top:** 11 contacts. **Medium:** 7 contacts. **Bottom:** 3 contacts. The B factor ( $B$ ) and the vibrational entropy ( $\ln W$ ) of each configuration is reported. The configurations are disposed with increasing entropy from left to right and from top to bottom. The configurations on the same line has the same number of contacts: in the traditional model (ISAW) they are thermodynamically equivalent because they have the same energy,  $E = -n_{co}$ ; in the model with vibration their free energies  $E - T \ln W$  differ in the vibrational entropy term.

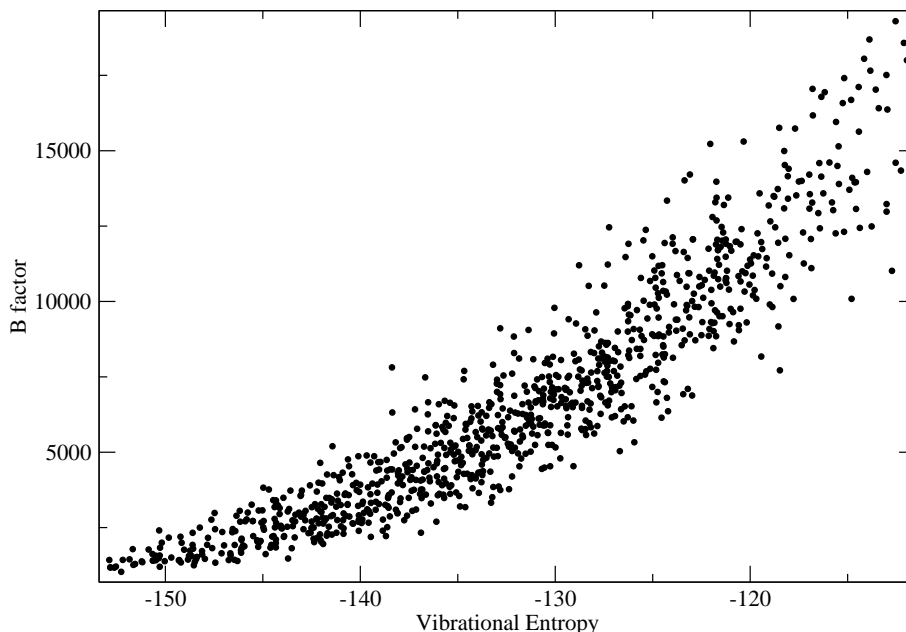


Figure 2.3: **Correlation between the B factor and the vibrational entropy.** Each point represents one configuration. The configurations have length 60 monomers and were sampled at inverse temperature  $\beta = 3.35$ . The graph is qualitatively representative of any set of random configurations. The correlation coefficient of the data in the figure is  $r = 0.92$ .

difference between the entropies of two walks does not always correspond to a positive difference in the total B-factor.

The correlation between the two quantities is well represented in Figure 2.3: a set of 10000 configurations of length 60 monomers sampled at inverse temperature  $\beta = 3.35$  is reported in the Vibrational Entropy - B-factor plane. The qualitative behavior would be identical for any other set of random configurations and the positive correlation is evident. The interest of the B factor is experimental, as it can be directly measured from x-ray crystallography. On the other hand, the important thermodynamic quantity is the vibrational entropy as it appears from the expression of the partition function, equation (2.13).

### 2.3 The observables and the assumption of *small fluctuations*

Any observable, such as the average radius of gyration, the number of contacts or the energy, should be calculated on the basis of the complete par-

Table 2.2: **Observables of a SAW of  $N$  monomers.**


---

Square radius of gyration	$R_{gyr}^2 = \frac{1}{N} \sum_{i=1}^N (\vec{r}_i - \langle \vec{r} \rangle)^2$
Center of mass	$\langle \vec{r} \rangle = \frac{1}{N} \sum_{i=1}^N \vec{r}_i$
Square end to end distance	$R_{end}^2 = (\vec{r}_N - \vec{r}_1)^2$
Energy	$E = -n_{co}$
Average contact length	$l_{co} = \frac{\sum_{i>j}^{\text{in contact}}  i - j }{n_{co}}$

---

tion function, equation (2.10), i.e. on the lattice+space coordinates. A fundamental assumption of our model is that *the fluctuations are very small* with respect to the lattice spacing. This assumption can be formulated also imposing that the spring constants participating in the definition of the elastic potential are strong enough that big fluctuations are associated to a prohibitive elastic energy and are negligible. In section 2.6.1 we verify that experimental parameters verify well this assumption. The assumption has few important consequences on the computation of the properties of the model.

First, the assumption justifies that the partial integration over the fluctuations within a small neighborhood of the lattice configuration is approximated by the integral over the whole space  $(-\infty, +\infty)$  in equation (2.11).

Second, any observable defined on the lattice, such as the radius of gyration or the number of contacts, can be calculated as an average over the lattice configurations:

$$\langle O \rangle_{VIBR} = \frac{1}{Z_{VIBR}} \sum_{\gamma} O(\gamma) W(\gamma) e^{-\beta E}. \quad (2.20)$$

It is interesting to write the same expression in terms of averages on the standard ensemble:

$$\langle O \rangle_{VIBR} = \frac{\langle OW \rangle_{ISAW}}{\langle W \rangle_{ISAW}}, \quad (2.21)$$

which looks as the correlation between observable and vibrational weight. In other words, if the calculated vibrational weight were a random number (uncorrelated with any observable:  $\langle OW \rangle_{ISAW} = \langle O \rangle_{ISAW} \langle W \rangle_{ISAW}$ ) or simply a constant (as implicitly assumed in the traditional model), the two models would bring exactly the same results. Of course, in the following we will show that this is not the case.

In the model with vibration the energy of a walk is composed by the contacts term and the elastic energy. The elastic energy is *not* constant on

the small fluctuations; on the opposite the small fluctuations assumption implies that it passes from zero to big values in the neighborhood of the lattice configuration. The total average energy of the walk should than be calculated on the original formulation (equation 2.10):

$$\begin{aligned}
 \langle V \rangle &= \frac{1}{Z_{VIBR}} \sum_{\gamma} \int d\vec{q} \left( E + \frac{1}{2} \vec{q} H^* \vec{q} \right) e^{-\beta [E + \frac{1}{2} \vec{q} H^* \vec{q}]} \\
 &= \frac{1}{Z_{VIBR}} \sum_{\gamma} E W(\gamma) e^{-\beta E} + \frac{1}{Z_{VIBR}} \sum_{\gamma} e^{-\beta E} \int d\vec{q} \left( \frac{1}{2} \vec{q} H^* \vec{q} \right) e^{-\beta \frac{1}{2} \vec{q} H^* \vec{q}} \\
 &= \langle E \rangle + (3N - 6) \frac{k_B T}{2},
 \end{aligned}$$

where the factor of the last term would be  $(2N-3)$  in 2 dimensions. The term linear in temperature  $T$  is typically associated with elastic vibrations. As we are concerned with the critical behavior, where the observables are highly non analytic, we won't consider the elastic part of the energy and we will *confuse* the *energy of the walk* with the contact energy alone  $\langle E \rangle$ .

Also, from now on we drop the subscripts *ISAW* or *VIBR*, implicitly meaning the second one if not otherwise specified.

## 2.4 An alternative approach: Energy Landscapes

To predict the behavior of a system, on the basis of laws of mechanics, one should know exactly the energy associated to every spatial arrangement of the system components

$$V(\vec{r}), \quad (2.22)$$

which is the Potential Energy Surface (PES) [75] of the system; formally it is a surface in a multidimensional space. A molecule, for example, is a system composed of many ( $N$ ) atoms, interacting through electrostatic interactions, as various as the covalent bonds, the van der Waals or the Hydrogen bond interactions. In general an exact and exhaustive description of the energy potential  $V$  as a function of the position of the  $N$  atoms,  $\vec{r} \in \mathbb{R}^{3N}$ , is unaffordable without make many assumptions and simplifications. One approach consists into characterize the *most important* configurations and built an approximation of  $V(\vec{r})$  from that. This *exploration* of the energy landscape may take place in two steps:

1. search of the minima of the energy, which are the most probable structures of the system, so the *most important*;
2. evaluation of the *basin of attraction* of each minimum, defined as the volume of the space of configuration which dynamically converges to that minimum.

In this section we show that the ISAW and VIBR models situate in the context of the *exploration* of the energy landscape and we will be able to illustrate which kind of approximation they are and how they differ, in particular how much better is VIBR with respect to ISAW.

The partition function of a polymer of  $N$  monomers

$$Z = \frac{1}{\Lambda^{3N}} \int d(\vec{r}) e^{-\beta V(\vec{r})}, \quad (2.23)$$

is the configurational integral of the Boltzmann factor ( $\beta$  is the inverse temperature) divided by the product of  $3N$  de Broglie wave lengths, which is the integration over the momenta. This equation is the ancestor of both equation (2.7), the ISAW partition function, and equation (2.10), the VIBR partition function. The first step towards them consists in splitting the total integral in a sum of integrals, each one limited to the *basin of attraction* of one minimum of the Potential Energy Surface, or inherent structure:

$$Z = \sum_{\gamma} Z_{\gamma}. \quad (2.24)$$

The next, and most important step, is the approximation of the integral over the neighborhood of the minimum  $\gamma$

$$Z_{\gamma} = \int_{I(\gamma)} d(\vec{r}) e^{-\beta V(\vec{r})}. \quad (2.25)$$

The Taylor expansion of  $V(\vec{r})$  around the minimum  $\gamma$  (equation 2.9) allows us to place ISAW and VIBR at two different truncations. Indeed, the ISAW approximation consists into truncate the expansion at the first term:

$$Z \sim \sum_{\gamma} e^{-\beta E(\gamma)} = Z_{ISAW}. \quad (2.26)$$

The bottom graph of Figure 2.4 represents the ISAW approximation of an hypothetical potential (top graph) with 4 minima ( $\gamma = x_1, x_2, x_3, x_4$ ); any other configuration is *thrown away* and the thermodynamics depends on the energy differences between the inherent structures.

If instead the expansion is truncated at the quadratic term (the linear term is zero), the VIBR approximation is obtained:

$$Z \sim \sum_{\gamma} \int_{I(\gamma)} d(\Delta\vec{r}) e^{-\beta(E(\gamma) + \frac{1}{2} \Delta\vec{r} H(\gamma) \Delta\vec{r})} \quad (2.27)$$

$$\sim \sum_{\gamma} \int_{[-\infty, +\infty]} d(\Delta\vec{r}) e^{-\beta(E(\gamma) + \frac{1}{2} \Delta\vec{r} H(\gamma) \Delta\vec{r})} = Z_{VIBR}, \quad (2.28)$$

where the ulterior approximation of the integration range allows to calculate the gaussian integral, resulting in the entropic weight  $W(\gamma)$  (equation 2.12).

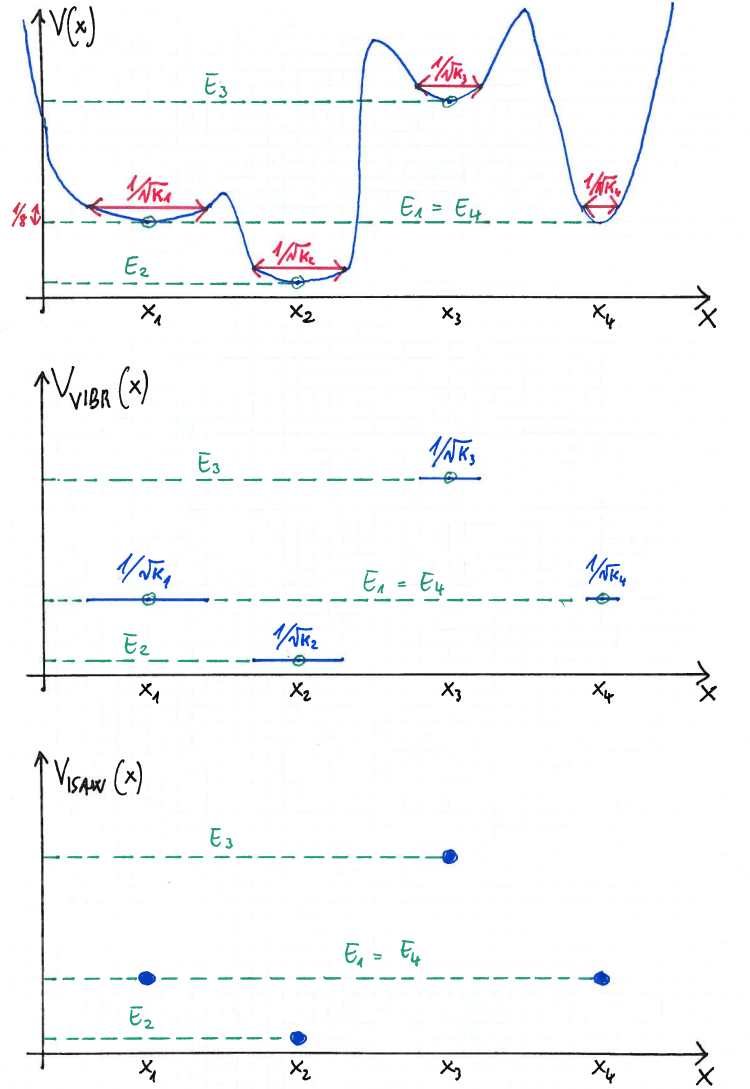


Figure 2.4: **Approximation of the energy landscape.** **Top:** an hypothetical Potential Energy Surface of one coordinate,  $V(x)$ . The PES has four minima with energy  $E_2 < E_1 = E_4 < E_3$ . In the neighborhood of each minimum the PES is well approximated by an harmonic function  $E_i + \frac{1}{2}k_i(x - x_i)^2 = E_i + \frac{1}{8}$  at the positions  $x_i \pm \frac{1}{2\sqrt{k_i}}$  (red arrows). The form of the potential between the harmonic approximation regions is arbitrary and does not enter in the calculations. **Middle:** the configurations in the interval  $x_i \pm \frac{1}{2\sqrt{k_i}}$  around each minimum are reassigned the energy  $E_i$ , other configurations are ignored. This is the VIBR approximation. **Bottom:** the configurations corresponding to the minima are only kept. This is the ISAW approximation.

This approximation is illustrated in the middle panel of Figure 2.4. Each minimum is *widened* over a region of configurations of size proportional to  $W_i = \frac{1}{\sqrt{k_i}}$ . This is the width of the parabola  $E_i + \frac{1}{2}k_i(x - x_i)^2$  at height  $E_i + \frac{1}{8}$ , and is illustrated by the red arrows in the original Potential (top panel). In other words, the configurations around  $\gamma$  which are most similar to it are added and they count as repetition of the same configurations, with the same Energy, the same Radius and so on. This looks crude, but it accounts of the entropic richness of each inherent configuration and it has significant consequences on the thermodynamics of the lattice model as we show in the rest of this Thesis.

## 2.5 The Hessian matrix

In this section we define the Hessian matrix of a Self Avoiding Walk on the lattice. The definition is based on the anisotropy elastic network model, but it differs from it in some important aspects.

1. The anisotropy elastic energy equation (2.1) is associated to each bond and each contact. As the bonds and contacts are defined by the lattice structure, the definition of a *cut-off* distance is not needed.
2. In addition, each angle between two consecutive bonds and each torsional angle defined by three consecutive bonds is associated with an angular elastic energy. A torsional network model (with rigid bond lengths and angles) has been recently proposed in [49].
3. Different elastic constants are associated to the bond, the bond angles, the torsional angles and the contacts elastic energies. On the opposite the traditional elastic network models are defined by a single parameter.

The expressions of each elastic interaction depend on the lattice: we treat the square lattice (2 dimensions) and the tetrahedral lattice (3 dimensions) separately.

### 2.5.1 2 dimensions

In 2 dimensions, a SAW of  $N$  monomers covers  $N - 1$  consecutive links ( $N$  sites) of a square lattice. Each bond has length  $l$  (the lattice spacing) and is associated with the interaction (equivalent to equation 2.1)

$$V_i^{bb} = \frac{1}{2}k_{BB}(|\vec{r}_{i+1} - \vec{r}_i| - l)^2 \quad i = 1 \dots N - 1. \quad (2.29)$$

The same energy, with a different spring constant, is attributed to the contacts:

$$V_{i,j}^{co} = \frac{1}{2}k_{co}(|\vec{r}_j - \vec{r}_i| - l)^2. \quad (2.30)$$

The angular elastic energy between two successive bonds is

$$V_i^{\text{angle}} = \frac{1}{2}k_{\text{angle}}(\theta_i - \theta_i^0)^2, \quad (2.31)$$

where the equilibrium angle between bonds  $i$  and  $i+1$  can be either  $\theta_i^0 = \pm\frac{\pi}{2}$  or  $\theta_i^0 = 0$ . On the 2 dimensional lattice there is no need to define a torsional potential.

In order to calculate the Hessian, we express the angular energy in terms of the coordinates of the monomers, through the cosine of the angles:

$$\cos \theta_i = \frac{\vec{R}_i \cdot \vec{R}_{i+1}}{|\vec{R}_i||\vec{R}_{i+1}|}. \quad (2.32)$$

If the equilibrium angle is  $\pm\frac{\pi}{2}$  ( $\cos \theta_i \sim \theta_i \mp \frac{\pi}{2}$ ), the potential is

$$V_i^{\text{right}} = \frac{1}{2}k_{\text{right}}\left(\theta_i \pm \frac{\pi}{2}\right)^2 \sim \frac{1}{2}k_{\text{right}}\cos^2 \theta_i; \quad (2.33)$$

if the equilibrium angle is 0 ( $\cos \theta_i \sim 1 - \frac{\theta_i^2}{2}$ ), the potential is

$$V_i^{\text{flat}} = \frac{1}{2}k_{\text{flat}}\theta_i^2 \sim k_{\text{flat}}(1 - \cos \theta_i). \quad (2.34)$$

The sum over the bonds, the angles and the contacts interactions gives the total elastic energy  $V$  of the SAW on the square lattice. The Hessian matrix is the second derivative with respect to the coordinates, calculated at zero displacement:

$$H_{ij} = \left. \frac{\partial^2 V}{\partial r_i \partial r_j} \right|_{\Delta \vec{r}=0}.$$

The derivation is reported in appendix A.

### 2.5.2 3 dimensions

In 3 dimensions a SAW of  $N$  monomers covers  $N - 1$  consecutive links of a tetrahedral lattice. The tetrahedral geometry is preferred to the simpler cubic geometry because a good definition of the elastic potential is possible only on the first one. Indeed, as shown in the 2 dimensional case, the elastic interactions are defined between two (bonds) or three (angles) consecutive monomers. The set of those interactions is enough to establish the lattice geometry. Instead, a SAW on the cubic lattice should be associated to many sites interactions, as we illustrate in the following example. Imagine



to grow a walk from the first step, along the x axes, followed by a step along the y axes: this turn defines the 3 directions of the lattice (the two directions of the steps and their common perpendicular). Imagine now that the walk grows straight for 2 other steps and turns at 5th step: in order to choose a direction, the 5th step should refer to the first 2 steps, which already defines the lattice. In general, in order to embed the walk on the lattice, any turn should refer to the previous turn which is possibly many steps before. The elastic interaction should then involve sites many steps away, depending on the turn, which does not have a clear physical meaning. The solution comes from adopting a lattice with no straight angles, such as the tetrahedral lattice, which is the geometry ideally formed by carbon atoms. This is also interesting because many polymer backbone bonds are between carbon atoms. In a diamond lattice each site (carbon atom) has four covalent bonds with as many other sites. The angle between any two bonds is  $\cos^{-1}(-\frac{1}{3}) \sim 109^\circ$ . Three consecutive bonds are disposed in such a way that the dihedral angle between the plane of the first two bonds and the plane of the second-third bonds is  $\pi$  (*trans*) or  $\pm\frac{\pi}{3}$  (*gauche*). A torsional potential associated to the dihedral angle fluctuations needs to be defined in addition to the bonds, contacts and bond angles interaction. All the interactions are defined on the basis of equations 2.29 (bonds and contacts) and 2.31 (bonds and dihedral angles). They are presented in Table 2.3, while the computation of the Hessian is demanded to appendix B.


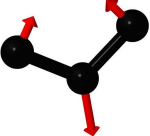
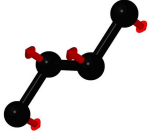

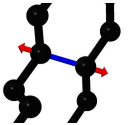
## 2.6 Comparison with off lattice force fields

Several force fields are used in molecular dynamics simulations, depending on the system to be studied and the desired degree of coarse graining. The aim of this section is to consider some of them and to compare the different definition of potentials. The coarse graining consists in general into represent a group of atoms (for example a heavy atom and a Hydrogen) with only one bead and is adopted to accelerate simulations. We take in consideration three of such force fields: the United Atom potential (UA, [62]), used in the study of polymethylene melts, a coarse grained model of proteins proposed by Honeycutt and Thirumalai for studying the nature of folded states (HT, [39]) and a more refined model of polyalanine-based peptides (HW, [36]).

The United Atom potential is a force field developed for investigating properties of alkane chains: the interactions defined in UA are very similar to the interactions defined in the Hessian of the model studied here. In UA, each carbon atom and the attached hydrogen atoms is represented as a single bead. The beads interact through:

1. a rigid bond between two successive beads;
2. an elastic potential for the angle between two successive bonds;

Table 2.3: **Interactions of VIBR.** The interactions defining the potential energy of small displacements of the tetrahedral walk with vibrations. A schematic representation, the functional form and the values of the fixed parameters are reported. The red arrows represents the normal mode of vibration associated to the interaction.

bonded interactions		functional form	parameter
bond		$V_i^{bb} = \frac{1}{2}k_{bb}( \vec{r}_{i+1} - \vec{r}_i  - l)^2$	$l = \text{lattice spacing}$
bend angle		$V_i^{be} = \frac{1}{2}k_{be}(\theta_i - \theta_i^0)^2$	$\theta_i^0 = \cos^{-1}\left(-\frac{1}{3}\right)$
dihedral angle		$V_i^{tr} = \frac{1}{2}k_{tr}(\alpha_i - \alpha_i^{tr})^2$	$\alpha_i^{tr} = -\pi$
		$V_i^{ga} = \frac{1}{2}k_{ga}(\alpha_i - \alpha_i^{ga})^2$	$\alpha_i^{ga} = \pm\frac{\pi}{3}$
non bonded interaction			
		$V_{i,j}^{co} = \frac{1}{2}k_{co}( \vec{r}_j - \vec{r}_i  - l)^2$	

3. a potential with three minima (the *trans*, *gauche*<sub>+</sub> and *gauche*<sub>-</sub>) for the dihedral angle between three successive bonds;
4. a Lennard Jones interaction between non bonded monomers.

The four interactions defined in the UA force field are reported in Table 2.4, where the equivalent spring constant is also calculated. The UA rigid bond

Table 2.4: **Interaction parameters of the United Atom Potential** [62] and equivalent parameter of VIBR.

interaction		parameters (kcal/mol)	equivalent spring constant (kcal/mol)
bond	constrained		$\infty$
bend angle	$\frac{1}{2}k_{\theta}(\theta - \theta_0)^2$	$k_{\theta} = 120$	$k_{be} = 120$
dihedral angle	$\frac{1}{2}[k_{\phi}^1(1 - \cos \Phi) + k_{\phi}^2(1 - \cos 2\Phi) + k_{\phi}^3(1 - \cos 3\Phi)]$	$k_{\phi}^1 = 1.6$ $k_{\phi}^2 = -0.867$ $k_{\phi}^3 = 3.24$	$k_{tr} = 13.6$ $k_{ga} = 14.8$
non bonded	$\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - 2 \left( \frac{\sigma}{r} \right)^6 \right]$	$\epsilon = 0.118$ $\sigma = 0.45 \text{ nm}$	$k_{co} = 41.9 \frac{\text{kcal}}{\text{mol nm}^2}$ $r_{min} = \sigma$

corresponds to an infinite spring constant. The UA bending elastic potential is identical to equation (2.31) and the equivalent spring constant can be immediately obtained. The equivalent elastic constant of the dihedral angle in the *trans* (*gauche*) configuration is obtained as the second derivative with respect to the angle of the UA definition calculated at the *trans* (*gauche*) angle. Similarly the equivalent spring constant of the Lennard Jones interaction is the second derivative of the potential calculated at the minimum energy distance ( $r_{min} = \sigma$ ).

HT is a coarse grained model used for studying proteins: the interactions are reported in Table 2.5. The bond is rigid, as in UA. Similarly to UA, the dihedral angle potential has three minima (despite the different functional form), the *trans* and the two *gauche* configurations. We remark that the equivalent spring constants of the bending and the dihedral angles have the same magnitude, while in UA the bending spring was one order of magnitude bigger than the dihedral spring. Substituting the UA values for  $\epsilon$  and  $\sigma$ , the contact spring constant gives  $k_{co} = 33.3 \frac{\text{kcal}}{\text{mol nm}^2}$ , of the same order of magnitude as UA, while the angles spring constant would result one order of magnitude smaller.

The HW potential [36] is more detailed. Each heavy atom with its attached hydrogen is represented as one bead (O, C, NH); the interactions keep in consideration the different nature of atoms and the parameter are

Table 2.5: **Interaction parameters of Coarse Grained protein potential** [39]. The energy  $\epsilon_h$  and the length  $\sigma$  are tunable parameters.

interaction		parameters (kcal/mol)	equivalent spring constant
bond	constrained		$\infty$
bend angle	$\frac{1}{2}k_\theta(\theta - \theta_0)^2$	$k_\theta = 20\epsilon_h$	$k_{be} = 20\epsilon_h$
dihedral angle	$A(1 + \cos \phi)$ $+ B(1 + \cos(3\phi))$	$A = B = 1.2\epsilon_h$	$k_{tr} = 12\epsilon_h$ $k_{ga} = 9.8\epsilon_h$
non bonded	$4\epsilon_h \left[ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right]$		$k_{co} = 57.14 \frac{\epsilon_h}{\sigma^2}$ $r_{min} = \sqrt[6]{2}\sigma$

obtained by optimization procedures. In Table 2.6 we show the parameters which are interpretable as one of the interactions defined in our model. The comparison is complicated by the atom dependent definitions of HW, while the model presented here is a homo-polymer. For example, three different expressions of the dihedral potential are proposed: they refer to the sequence of backbone atoms  $C_\alpha$ -C-N- $C_\alpha$  if one start from either  $C_\alpha$ , C or N. The non-bonded interactions depend also on the couple of atoms involved: in the case of van der Waals interaction we report the energy constant  $\epsilon$  which refers to atoms of the same type and a typical equilibrium distance, in the aim of evaluating order of magnitudes and not exact results. Finally the Hydrogen bond interaction between O and NH is enriched by repulsive terms between side atoms: they don't matter because the spring constant is calculated as the second derivative with respect to the distance between O and NH only. The angular interactions are comparable with the previous models. We remark in particular that the Hydrogen bond shows a stiffness tenth of times bigger than the van der Waals interaction.

### 2.6.1 Small fluctuations

With these data we are able to verify at which extent the small fluctuations assumption is valid (section 2.3). Let's consider a simple elastic potential between two monomers with equilibrium distance  $l$  and elastic constant  $k_{co}$ . The probability of the distance  $r$  is:

$$p(r) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{k_{co}}{2k_B T}(r-l)^2} \quad (2.35)$$

with average fluctuation  $\sigma = \sqrt{\frac{k_B T}{k_{co}}}$ . In Table 2.7 we show the average fluctuation calculated at room temperature  $T = 300\text{K}$  (corresponding to  $k_B T = 0.6$  kcal/mol). In the stiff hydrogen bond interaction, the fluctuation

Table 2.6: **Some terms of the interaction parameters of Coarse Grained protein potential** [36]. The dihedral angle potential depends on the atoms forming the quadruplet; in any of the three cases, the minima are equivalent and give the same  $k_{dih}$  spring constants. The non bonded interactions are Van der Waals (VdW) or Hydrogen bond (Hb) interactions. The parameter of VdW refers to atoms of the same kind (O-O, CH-CH, NH-NH); the equilibrium distance may vary slightly around the reported value, depending on the atoms. The Hb parameter refers to Lennard-Jones interaction between O and NH; the  $V_{rep}$  takes into account repulsion between side atoms: deriving twice with respect to the O-NH distance the repulsive term disappears.

interaction		parameters (kcal/mol)	equivalent spring constant (kcal/mol)
bond	constrained		$\infty$
bend	$\frac{1}{2}k_{\theta}(\theta - \theta_0)^2$	$k_{\theta} = 72$	$k_{be} = 72$
dihedral	$K_t(1 + \cos(2\phi - \pi))$	$K_t = 10$	$k_{dih} = 40$
	$K_t(1 + \cos(6\phi - \pi))$	$K_t = 0.2$	$k_{dih} = 7.2$
	$K_t(1 + \cos(6\phi))$	$K_t = 0.2$	$k_{dih} = 7.2$
VdW	$4\epsilon_{ij} \left[ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right]$	$\epsilon_{ij} = 0.25$	$k_{co} = 89.29 \frac{\text{kcal}}{\text{mol nm}^2}$
		$\sigma \sim 0.35\text{nm}$	$r_{min} = 0.39 \text{ nm}$
Hb	$4\epsilon_{attr} \left[ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right]$	$\epsilon_{attr} = 3.35$	$k_{co} = 3323 \frac{\text{kcal}}{\text{mol nm}^2}$
	$+V_{rep}$	$\sigma = 0.24\text{nm}$	$r_{min} = 0.27 \text{ nm}$

of the length is some percentile of the bond length; the van der Waals interaction are more mobile, but the average fluctuation is well below the half of the bond length and the lattice structure can be considered preserved.

The average fluctuation of the angle calculated from the gaussian probability

$$p(\alpha) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{k_{\text{angle}}}{2k_B T}(\alpha - \alpha_0)^2} \quad (2.36)$$

can be translated in a linear displacement if multiplied by the lattice spacing ( $l$ ), by definition of the radiant degree. So,  $\sigma$  calculated from  $p(\alpha)$  represents already an indication of the displacement of the monomer with respect to the lattice spacing.

Table 2.7: **Average fluctuations of contacts and angles** calculated with the parameters provided by the cited force fields.

contact	$\sigma$	$l$	$\sigma/l$	angle	$\sigma$
UA - VdW	$11.9 \cdot 10^{-2}$ nm	0.45 nm	26.5%	UA - bend	7.0%
HW - VdW	$8.1 \cdot 10^{-2}$ nm	0.39 nm	20.9%	UA - trans	20.9%
HW - Hb	$1.3 \cdot 10^{-2}$ nm	0.27 nm	4.9%		

## 2.7 Methods

### 2.7.1 Units of measurement

In this work we adopted adimensional units. In particular, the energy of non-bonded monomer contacts is  $\varepsilon = 1$  and the Boltzmann constant  $k_B = 1$ . This means that the temperature is expressed in units of  $\frac{\varepsilon}{k_B}$ , the energy constant depending on the specific experimental set. The lattice spacing is also set to  $l = 1$ , so the stiffness constants of the elastic potential are expressed in units of energy.

### 2.7.2 Simulation technique

An efficient dynamic Monte Carlo algorithm for generating SAWs is presented in [48]. It was used to study ISAWs on the cubic lattice in [67]; the Pivot algorithm was also defined for the ISAWs on the tetrahedral lattice [80]. The only modification needed to adapt the algorithm to the VIBRat-ing Interacting Self Avoiding Walk concerns the Metropolis acceptance ratio which has to take into account the total free energy of the single Walk,  $E(\gamma) - T \ln W(\gamma)$  (equation 2.13), sum of the contact energy and the vibra-

tional entropy:

$$p(\gamma_t \rightarrow \gamma_{t+1}) = \min \left( 1, e^{-\beta[\Delta E - T \Delta \ln W]} \right). \quad (2.37)$$

The algorithm is essentially the same in 2 and 3 dimensions. The starting configuration is a straight walk: in 2 dimensions it consists of a sequence of steps on the same axes, while on the tetrahedral lattice the *straight* walk is a sequence of steps in the *trans* configuration. As the algorithm is ergodic, the initial configuration does not matter and it is chosen for convenience because is easily defined. At each time step of the algorithm, one site of the walk is chosen at random and the *upper* part is modified by a symmetry group element. This is a transformation which leaves the walk on the lattice. There are eight such transformation on the square lattice (the identity included), while on the tetrahedral lattice they are limited to two rotations of  $\pm \frac{2\pi}{3}$ . If the new chain is self avoiding it is accepted with the probability (2.37), otherwise it is rejected.

In order to calculate the free energy of the new chain the algorithm calculates at each step:

1. the number of contacts between not neighbor sites ( $E = -\varepsilon n_{co}$ );
2. the sum of the logarithm of the not null eigenvalues of the Hessian matrix ( $\ln W = -\frac{1}{2} \sum_i \ln \lambda_i$ ).

The number of contacts is calculated by means of an hash table method: this requires a lot of memory (for the hash table, which is the list of the whole lattice sites) but is fast, order  $O(N)$ . The Hessian matrix of the new conformation is built step by step (time  $O(N)$ ) and the eigenvalues are calculated by means of a Lapack routine which is of complexity  $O(N^3)$  [3]. The complexity of the diagonalization algorithm slows down dramatically the simulations, so that it is difficult to obtain results at high  $N$ . The parallel version of the algorithm, ScaLAPACK [12], gains in efficiency only for matrix size bigger than  $1000 \times 1000$  which corresponds already to a too much long time in order to get many data.

### 2.7.3 Data analysis

The dynamic Monte Carlo algorithm produces a sequence of data which is analyzed following these steps:

1. the autocorrelation function is evaluated: from a series of  $S$  data  $\{D_i\}$ :

$$\rho(t) = \frac{\langle D_i D_{i+t} \rangle - \langle D_i \rangle^2}{\langle D_i D_i \rangle - \langle D_i \rangle^2}; \quad (2.38)$$

the average is over the series of data;

2. the autocorrelation time  $\tau$  is estimated as the time at which  $\rho(t) < e^{-1}$ ;
3.  $10\tau$  initial steps are considered to pertain the thermalization of the process and are discarded;
4. one datum every  $\tau$  is kept and used for the analysis (standard averages);
5. the data are elaborated with multiple histogram method technique and the error bars are evaluated with bootstrap method.

#### 2.7.4 Multiple Histogram Method

The MHM is *an optimized method for combining the data from an arbitrary number of simulations to obtain information over a wide range of parameter values in the form of continuous functions*[28]. In our work, a single simulation is a canonical sampling of the system of fixed length  $N$  and fixed inverse temperature  $\beta$ . We want to calculate some observable over a range of temperature, in the aim of finding for example the maximum of the specific heat. So, at a fixed length, we run several simulations at different temperatures and we combined the results through the MHM method. This approach extracts the maximum amount of information from the simulations and fits well with the modern trend of gain power from parallelization of codes. In practice, the configurations of a canonical simulation are sampled following the probability distribution

$$p(\gamma) = \frac{w(\gamma)e^{-\beta E(\gamma)}}{Z}. \quad (2.39)$$

In order to calculate the average energy at any temperature, one is interested in the distribution

$$p(E) = \frac{\mathcal{V}(E)e^{-\beta E}}{Z} \quad (2.40)$$

and in particular on the volume of the phase space at fixed energy,  $\mathcal{V}(E)$ , which is the only unknown of the system, but is shared by all the simulations at any temperature. The number of observation of a certain energy in a simulation is proportional to the probability (2.40). The algorithm elaborates the results of several distinct runs and finds the distribution  $\mathcal{V}(E)$  which fits at best all the simulation results. The method is based on a self consistency equation between the partition functions at different temperature and is generalizable to any other observable. We recommend the book [52] for the technical details.

We used the MHM to calculate all the observables, energy, Radii, heat capacity and so on, as well as the density of states  $\mathcal{V}(E)$  which allows to calculate the microcanonical entropy  $S(E) = \ln \mathcal{V}(E)$  (see section 4.6).



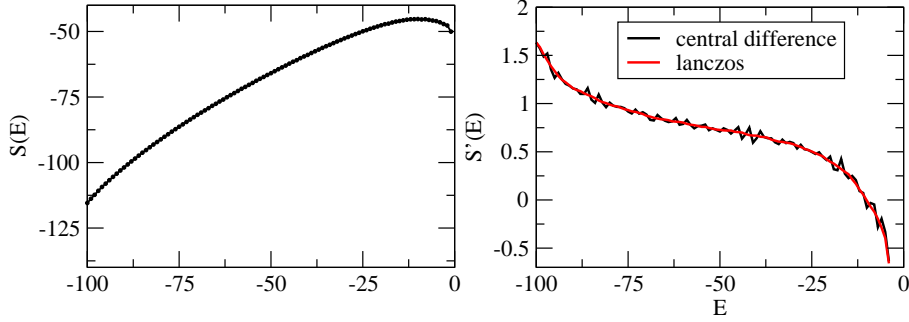


Figure 2.5: **The Lanczos low noise differentiator.** The function  $S(E)$  on the left is derived with respect to  $E$  with a standard central difference method and a Lanczos low noise differentiator. In both cases the derivative at each point is calculated from the value of the function at 7 points around.

The errors were calculated by means of the bootstrap method: the above procedure is repeated several times on group of data randomly sampled from the available; the errors is then traditionally calculated on the obtained results. The error converge after a certain time to a stable results: we found that 50 bootstrap extractions is already enough to get a stable error estimation.

### 2.7.5 Low noise differentiator

In the study of microcanonical entropy (section 4.6) we are concerned with its derivative. Nevertheless the microcanonical entropy is obtained from simulations and is affected by noise, like experimental results usually do. So calculating the derivative of a noisy function can not be done with a simple finite difference method. Low-noise differentiator can be used and the results are pretty nice. We used the so called Lanczos differentiator [59] and we found very useful the web page [38] where a complete explication of the method and reference formula can be found.



## Chapter 3

# Vibrational spectrum of linear chains

This chapter presents a quantitative analysis of the vibrational spectrum of polymers modeled as self avoiding walk with vibrations. Polymers have vibration profiles depending on the conformation they assume: the open and the compact configuration differ basically in the low frequency region of the spectrum. This behavior can be traced back to the behavior of lattice solids which depends on the dimensionality of the system (section 3.1). The analysis confirms that the coil (section 3.2) and the globule (section 3.3) populations have different spectral properties. The vibrational entropy is a function of the spectrum (equation 2.19); the observations of this chapter reveal that the intrachain contact elastic constant  $k_{co}$  plays a key role in triggering a new phenomenology in the model with vibrations.

### 3.1 Linear chains and solids

The vibrational spectrum of a system is the generalization of the frequency of one single spring. It is calculated from the eigenvalues of the Hessian matrix and depends on the geometry of the interactions. More precisely a system made by  $N$  particles with identical masses  $m$  in  $d$  dimensions has  $dN$  frequencies of vibration proportional to the square roots of the Hessian eigenvalues  $\lambda$ :

$$\omega_n = \sqrt{\frac{\lambda_n}{m}} \quad n = 1, \dots, dN. \quad (3.1)$$

In some special cases a simple analytical formula for the list of frequencies can be calculated. The case of interest here is the linear chain of identical particles [47]: a set of beads in a row linked by springs. The spectrum varies depending on the boundary condition, that is how the ends are treated (free ends, periodic condition, connected to a wall...). The spectrum of a linear

chain with free ends is:

$$\omega_n = 2 \sin\left(\frac{n\pi}{N}\right) \quad n = 1, \dots, N-1, \quad (3.2)$$

where the chain has  $N$  beads ( $N-1$  identical springs) and the masses and elastic constant are unitary.

The spectrum can be also described by the density of frequencies  $\rho(\omega)$ : this function gives the number of frequencies in the interval  $(\omega, \omega + d\omega)$  of width  $d\omega$ , normalized by the total number of frequencies  $dN$ . In the case of the linear chain with free ends the density of frequencies is ( $d=1$ )

$$\rho(\omega) = \frac{1}{N} \frac{dn}{d\omega} = \frac{1}{N} \frac{1}{\frac{d\omega_n}{dn}(n(\omega))} = \frac{1}{\pi \sqrt{1 - (\omega/2)^2}}. \quad (3.3)$$

The function  $n(\omega)$  is the inverse of the function  $\omega_n$  in equation (3.2). The spectrum and the density of frequencies of a linear chain of length 10000 is reported in Figure 3.1.

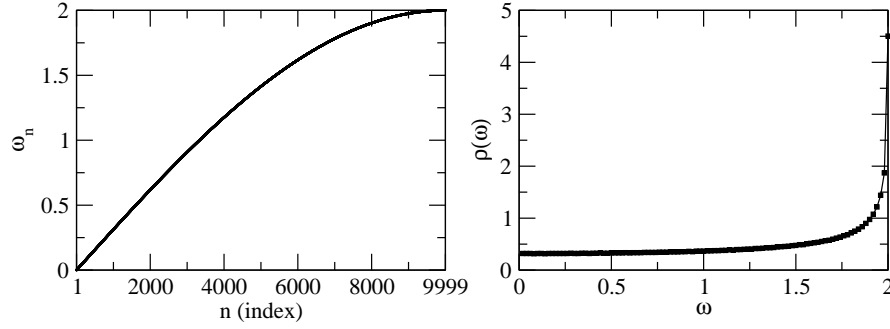


Figure 3.1: **Frequency spectrum of a linear chain.** The chain has 10000 beads and free ends. **Left:** list of the 9999 not null frequencies, equation (3.2). **Right:** density of frequency  $\rho(\omega)$  calculated from the frequency list over a bin  $d\omega = 0.02$ : the agreement with equation (3.3) is very good.

The linear chain is a 1 dimensional system and the main feature of its spectrum is that the density of modes is constant over a large portion of the spectrum, in the low frequency region:

$$\rho_1(\omega) \propto \frac{1}{\frac{d\omega_n}{dn}} \sim \text{const} \quad \omega \rightarrow 0 \quad (3.4)$$

In general, the low frequency region can be interpolated by a power of the index  $n$  (the frequencies  $\omega_n$  are listed in increasing order):

$$\omega_n \propto n^a. \quad (3.5)$$

Inverting it,  $n \propto \omega^{1/a}$ , the frequency density is:

$$\rho(\omega) \propto \frac{dn}{d\omega} \propto \omega^{1/a-1}. \quad (3.6)$$

The classical theory of the harmonic crystal [5] describes the distribution of the vibrational modes of a 3 dimensional Bravais lattice: a set of beads on a Bravais lattice interacting through harmonic interaction (springs). The wave vectors can take on its direction in a 3 dimensional space ( $i = x, y, z$ , but each component  $k_i$  is limited to a discrete series of values, because of the lattice nature of the system:  $k_i \propto n$ ,  $n = 1, 2, \dots$ , similarly to the linear chain, equation 3.2 Moreover, in the low region of the spectrum, the frequency is proportional to the wave vector in each direction: in  $d$  dimensions the density of frequencies in the interval  $(\omega, \omega + d\omega)$  is proportional to the number of wave vectors in the  $d$ -dimensional surface of radius  $\omega$ :

$$\rho_d(\omega) \propto \omega^{d-1}. \quad (3.7)$$

Comparing with the relation (3.6):

$$d = \frac{1}{a}. \quad (3.8)$$

In particular the spectrum is poor in low frequency modes when the spatial dimension increases. For example, relations (3.5) and (3.7) in  $d = 3$  are reported in Figure 3.2: the density  $\rho(\omega)$  vanishes at small  $\omega$ .

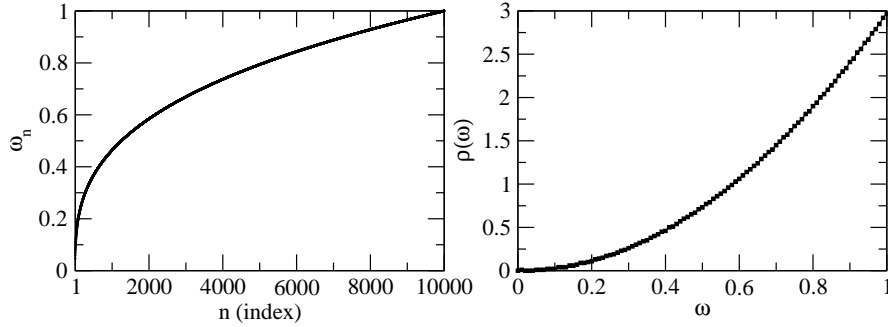


Figure 3.2: **Low frequency spectrum of a 3-dimensional solid. Right:** the frequency density  $\rho(\omega) \propto \omega^2$ , equation (3.7). **Left:** the corresponding frequencies are  $\omega_n \propto n^{1/3}$ .

Generalizing equation (3.8), we define the *vibrational dimension* of a system as

$$\bar{d} = \frac{1}{a}. \quad (3.9)$$

From this basic example we retain that in simple systems the *low frequency normal modes are more abundant in 1 dimensional objects than in 3 dimensions*. In the next section we will show that

- the spectrum of a coil polymer is richer in low frequencies than the standard linear chain;

- the spectrum of a compact polymer is
  - roughly the same as the spectrum of the coil polymer if the intrachain interaction are not taken into account;
  - poorer in lower frequencies when the whole set of interaction is taken into account.

## 3.2 Coils

Coils are conformations characterized by few contacts and are typical of the high temperature phase. A coil conformation is characterized by a vibrational spectrum rich in low frequencies. The vibrational spectrum is determined by the elastic fluctuations of the bonds, the bending and dihedral angles and the contacts (Table 2.3). In this section we show that if the non bonded interactions is not taken into account ( $k_{co} = 0$ ):

1. these spectra are rich in low frequencies: their vibrational dimension is considerably smaller than 1;
2. the spectrum is almost independent from the conformation of the walk: the elongated chain, a generic coil or a very compact conformation have very similar vibrational profiles.

The second observation means that from the point of view of the vibrational analysis, if  $k_{co} = 0$  all the configurations are *coil*, independently of the number of contacts. Indeed the information of a contact is contained in the elastic intrachain contactspring: this is candidate to be the key parameter of the model with vibrations.

In the two following sections of this chapter we will show how the spectra differentiate between the two phases (few and many contacts) when  $k_{co} > 0$ . In the next chapter we will see that  $k_{co}$  is the discriminating parameter between two universality classes.

### 3.2.1 2 dimensions

The straight polymer in 2 dimensions consists in a sequence of steps in the same direction. In this configuration the bond interactions give exactly the spectrum of the linear chain, while the angular vibrations (only flat angles) add modes of lower frequency. Indeed, while the bonds vibrations are almost proportional to the index,  $\omega_n \sim n$ , the flat angle vibrations fit very well with a quadratic function:

$$\omega_n^{flat} \propto n^2 \quad (3.10)$$

which implies  $\rho_{flat}(\omega) \propto \omega^{-1}$ . Altogether the fluctuations of the straight chain show an intermediate regime between the bonds and the flat angles:

$$\omega_n \propto n^a \quad 1 < a < 2. \quad (3.11)$$

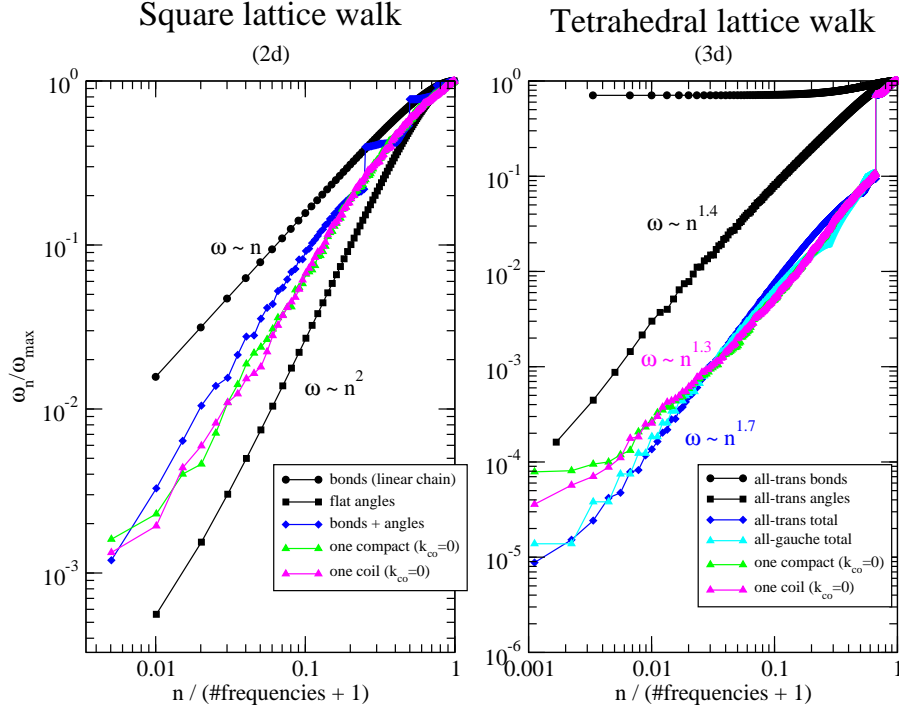


Figure 3.3: **Frequency spectrum of coil polymers.** The frequencies are relative to the highest one ( $\omega_n/\omega_{\max}$ ) and the index  $n$  is relative to the number of not null frequencies plus 1. Black curves are spectra of the bond (upper) and the angle (bottom) vibrations in the straight chain (a linear chain in 2 dimensions, an all-*trans* configuration in 3 dimensions). The other colors indicate the total spectra of the straight, a coil and a compact conformation, with no intrachain interaction ( $k_{co} = 0$ ). **Left:** walks on the square lattice: the bonds spectrum is identical to the linear chain. **Right:** walks on the tetrahedral lattice: the total spectrum of an all-*gauche* conformation is also reported. The coil and compact 3 dimensional walks have 300 monomers and 68 and 196 contacts respectively. The coil and compact 2 dimensional walk have 100 monomers and 30 and 80 contacts respectively.

Concerning the right angles, it is possible to show that the angular spectrum of a stairs-shaped walk on the square lattice (a configuration consisting of only right angles) is identical to the bond linear chain spectrum:  $\omega_n^{right} = 2 \sin\left(\frac{n}{N-1} \frac{\pi}{2}\right)$   $n = 1, \dots, N-2$ . As there are no other contributions to the vibrational spectrum of the square lattice walks, we guess that the spectrum of any coil behaves as in equation (3.11). In Figure 3.3 the spectrum of the straight configuration, an open coil configuration and a very compact configurations are reported: they are indeed similar and positioned between the flat angles and the bonds (equal to right angles) spectra.

### 3.2.2 3 dimensions

The model of the polymer in 3 dimensions is a chain on the tetrahedral chain, which is obviously not the simple extension of the square lattice to 3 dimensions. A fundamental difference is that there are no flat angles, so a perfectly linear chain does not exist. The most elongated walk is a sequence of *trans* dihedral angles, a zig-zag shaped chain. For this conformation the bond vibrations are bounded from below,  $\omega_n \sim n^0$ , so they do not contribute to the formation of global modes. Moreover the bond interactions are modeled to be very stiff with respect to other interactions and their presence can be recognized in a very high frequency region of the spectrum. Hence *the low frequency motion of the tetrahedral polymer is mainly associated to angular vibrations.*

The angular motions come from the variation of the bending angle between consecutive bonds and the dihedral angle between three successive bonds around their equilibrium value ( $\sim 109^\circ$  and the *trans/gauche* angle respectively). A separate analysis of their contribution is presented in appendix C. In total the backbone spectrum of the all-*trans* configuration decays to 0 faster than linearly. Similarly to the situation in two dimensions, the spectra of the coil and compact conformations are almost the same, all sharing the low frequency richness feature:

$$\omega_n^{3d \text{ coil}} \propto n^a \quad a > 1. \quad (3.12)$$

In conclusion we state that the walks on the lattice share a common vibrational footprint with the characteristic to be very rich in low frequencies. The vibrational dimension is

$$\bar{d}(\text{coil}) < 1 \quad (3.13)$$

## 3.3 Globules

The walks with many contacts are representative of the polymer globule state. The conformation influences the vibration spectrum as long as the



intrachain contacts springs are active ( $k_{co} > 0$ ); otherwise the results of the previous section show that the two cases are practically the same. We consider the same compact chain whose spectrum was shown in Figure 3.3 and we calculate its spectrum at increasing values of the contact spring constant  $k_{co}$ , Figure 3.4. The chosen compact configuration is a maximally compact (Hamiltonian walk) on the square lattice and the most compact configuration from simulations on the diamond lattice.

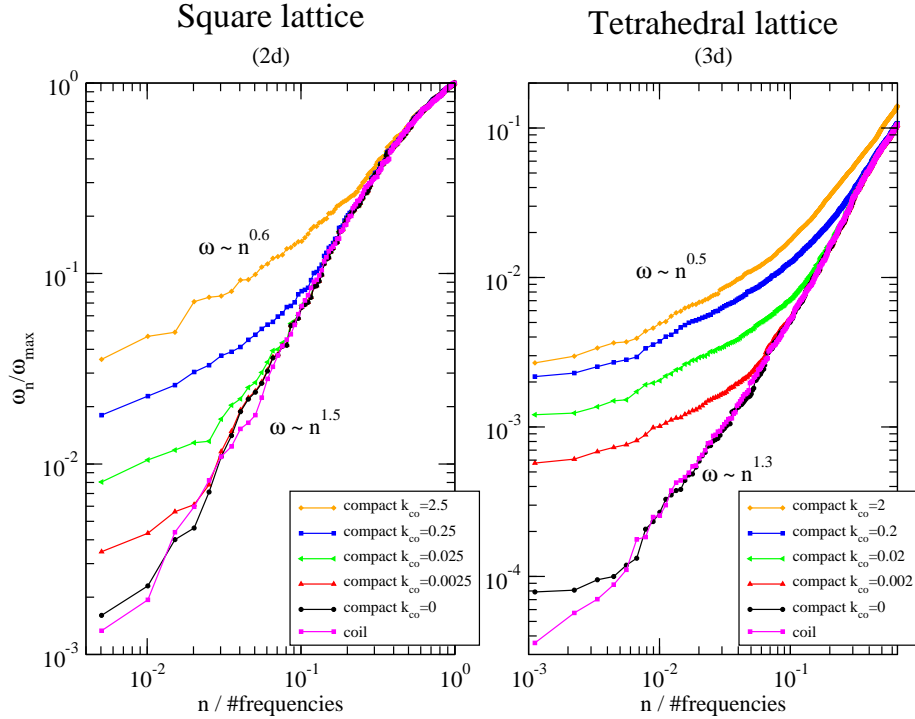


Figure 3.4: **Frequency spectrum of compact polymers.** The spectrum of a maximally compact configuration on the square lattice (**left**, 100 monomers, 80 contacts) and the tetrahedral lattice (**right**, 300 monomers, 196 contacts) at increasing values of the intrachain elastic constant  $k_{co}$ . The other constant are set to 25 : 4 : 6 (square) and 150 : 1 : 1 : 0.6 (tetrahedral). In the tetrahedral case the high frequency region (last 299 modes) is not shown.

We interpret this fact as a manifestation that if only the backbone interactions (bonds and angles) are considered then a polymer *does not feel* its compactness or coil-ness. Instead, the contacts introduce an interaction between monomers far away along the chain which holds the information of compactness. The main difference is observed in the low frequency region: while the coil spectrum decays faster than linearly to 0, the compact

spectrum decays with a smaller exponent:

$$\omega_n \propto n^a \quad a < 1. \quad (3.14)$$

We may expect  $d = 1$  for coils, but a smaller value is obtained in reason of the angular contribution:  $d(\text{coil}) \sim \frac{1}{1.5} \sim 0.67 < 1$ . Similarly the compact chain spectral dimension is lower than the spatial dimension:

$$\bar{d}_{2d}(\text{compact}) \sim 1.25 < 2 \quad (3.15)$$

$$\bar{d}_{3d}(\text{compact}) \sim 2.00 < 3 \quad (3.16)$$

### 3.4 Spectrum of coil and compact ensemble

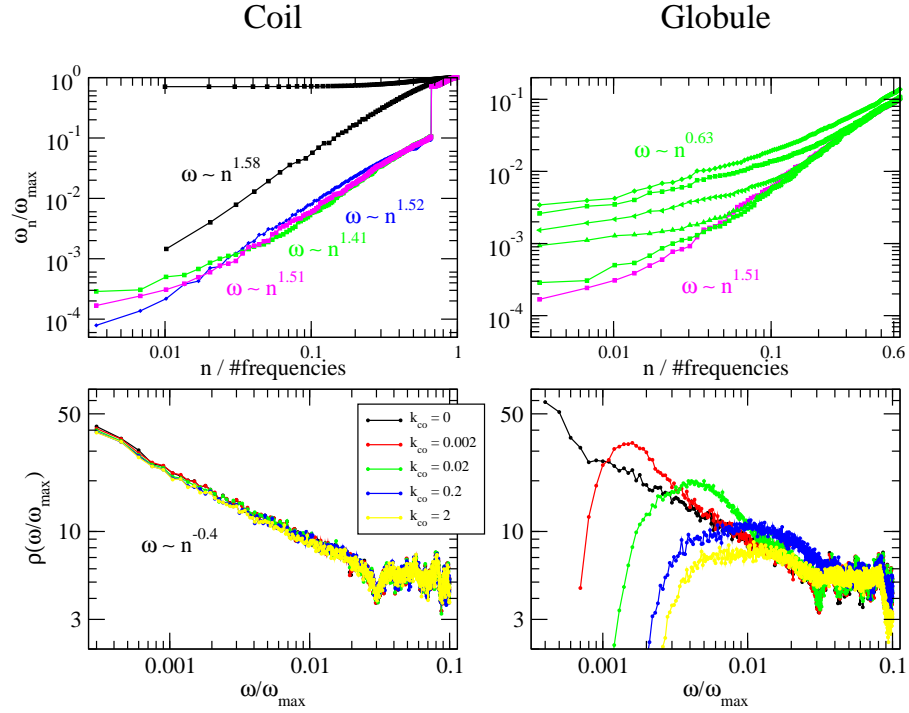


Figure 3.5: **Spectrum and density of frequencies of polymers in 3 dimensions.** **Upper row:** the same spectra as in Figure 3.3 and 3.4 (3 dimensions) for a polymer of length  $N = 100$  monomers. The coil configuration (magenta) has 5 contacts and the compact configuration (green) has 59 contacts. **Bottom row:** the frequency density is calculated on 1000 configurations of length  $N = 100$  monomers. **Left:** configurations with 5 contacts (coils). **Right:** configurations with more than 52 contacts (compact). Notice that the frequency density are normalized.

The previous observations are generalized with the computation of the vibrational spectrum of an ensemble of walks chosen either in the compact or

the coil state. Fixing the length to 100 monomers, we compute the spectra of 1000 walks with 5 contacts (coil sample) and with 52 to 59 contacts (compact sample) and we calculate the frequency density  $\rho(\omega)$  of the ensemble. The results reported in Figure 3.5 confirm the key role of the contact spring constant: while in the coil ensemble the exiguous number of contacts makes the spectrum almost independent to the value of  $k_{co}$ , the compact ensemble loses low frequencies with increasing  $k_{co}$ .

### 3.5 Fractons

The fact that the presence of intrachain contacts has an influence on the vibrational properties of the proteins is not new.

In the early eighties, experiments showed that the phonon dimension  $\bar{d}$  equation (3.9), agrees with the fractal dimension of the object, which is an index of the degree to which a structure fills the space in which it resides; in random/self-avoiding walks the fractal dimension is  $d_f = \nu^{-1}$ . In particular, the phonon dimension calculated from the temperature dependence of the Raman electron spin relaxation rate is in very good agreement with the fractal dimension calculated from x-ray data [63, 2].

The phonon density of states can be calculated also on fractals (fracton dimension  $d_{fr}$ ) [1]; if one calculates the fracton dimension of a random / self-avoiding walk, the value  $d_{fr} = 1$  is obtained irrespective to  $\nu$ : vibrations are sensitive to the connectivity and the walks are connected as a linear chain no matter their filling the space. It emerges [37] that the contribution of intrachain contacts is essential to obtain the experimental phonon dimension, which is bigger than 1.

In other words, if the self avoiding walk is considered without contacts (transposing to our language,  $k_{co} = 0$ ) the fracton (= phonon) dimension is NOT affected by the fractal nature of the object (in our language, the compactness) and the vibrational spectrum is the same as a linear chain, ( $\bar{d} = 1$ ), in contraddiction with experiments. The phonon dimension is raised if intrachain contacts through which vibrations can propagate are added. It is argued [37] that for a high enough density of contacts (bridges) the agreement with experiments can be restored  $\bar{d} = d_f$ . In our model the addition of contacts is equivalent to the case where the contact spring constant is  $k_{co} > 0$ . Concerning the actual values, the cited experiments report phonon dimensions ranging between 1.34 and 1.76. The vibrational spectrum of the globule ensemble presented in Figure 3.5 gives

$$\bar{d} \sim \frac{1}{0.63} = 1.58,$$

right in the middle of the experimental interval.

Because of the promising agreement between our result and the previous experimental and theoretical observations, and in light of the results shown

in the following chapters, the study of fractons on self-avoiding walks would deserve further analysis. One open question concerns the coil ensemble phonon dimension,  $\bar{d} \sim \frac{1}{1.51} = 0.6$  (Figure 3.5), which is lower than the expected  $\bar{d} = 1$ .

## Chapter 4

# The coil to globule transition

The local vibrations of a molecule around its equilibrium position is source of conformational entropy. This ingredient of the physics of macromolecules is not supposed to influence the thermodynamics as long as it does not display some specific traits. Indeed, if the vibrational weight would be uniformly distributed between all the conformations, like a constant or a random number, the probability of different conformations would be all affected by the same factor and this would not produce any consequence on the observables. This behavior of the vibrational weight corresponds to the observations made in the previous chapter concerning the spectra with  $k_{co} = 0$  (section 3.2). We will show that the thermodynamics of the corresponding model is very similar to the standard ISAW.

On the opposite, when  $k_{co} > 0$ , the spectrum of vibration changes appreciably between the compact or the coil configuration, specially in the low frequency region which is the most important source of vibrational entropy because it is associated with global motion. We will show that the coil to globule transition in both 2 and 3 dimensions is drastically affected. In 3 dimensions the transition becomes a two state process, akin to the all-or-none transition manifested by proteins (section 4.3); in 2 dimensions the transition is continuous but sharper than in the model without vibrations (section 4.5). These modification can only partially be understood on the basis of the redistribution of vibrational entropy privileging the open conformations with respect to the compact ones: in section 4.6 we individuate that the vibrational entropy has to satisfy the ulterior necessary condition of convexity with respect to the energy in order to explain the observed phenomenology. Interestingly the condition is the same in both 2 and 3 dimensions and the difference (continuous versus two state) may be attributed to the underlying ISAW.

## 4.1 Choice of the spring constants in 3 dimensions

From the analysis of MD force fields (section 2.6) we retain some general aspects:

- the bonds are very stiff ( $k_{bb} = \infty$ );
- there is some freedom in the value of the angular constants;
- the non bonded interactions show a remarkable different stiffness if refer to the van der Waals or the Hydrogen bond.

In the definition of the parameters for the VIBR model we have to take into account two additional mathematical aspects.

First, it is not possible to have rigid bonds in this model. The vibrational entropy of the single configuration is the sum of the logarithm of  $3N - 6$  not null eigenvalues, as many as the internal degrees of freedom:

$$\ln W = -\frac{1}{2} \sum_{i=1}^{3N-6} \ln \lambda_i.$$

Imposing the rigidity of one bond is equivalent to an infinite spring constant or null eigenvalue, which makes  $\ln W$  to diverge. Apparently, this problem could be solved by calculating the sum over the  $3N - 6 - N$  not null eigenvalues: this is equivalent to assign a constant amount of entropy to the vibrations of the bonds. This solution would work only if we don't consider non bonded interaction, that is in a model with only angle vibrations. But, when a non bonded interaction is added, also an additional degree of freedom could be added (up to a maximum of  $3N - 6$ ) and we would compare the entropy of different conformations with different number of degrees of freedom, which brings inconsistency. On the other hand, in MD the rigid approximation simplifies the simulations because it reduces the time step; in a gaussian network model, where all eigenvalues are calculated, there is no reason for introducing such approximation: we chose instead a very high value for  $k_{BB}$  (Table 4.1).

The second mathematical aspect concerns the dihedral angle spring constant and their associated eigenvalue. If we consider a triplet of monomers (in 2 dimensions) in the  $\alpha = 0$  or  $\alpha = \frac{\pi}{2}$  configuration interacting only through the dihedral potential, the relative Hessian have different eigenvalues. Since the entropy depends on eigenvalues (and not on spring constants), we chose to set the *trans* and the *gauche* configuration constants to a ratio such that the correspondent eigenvalues are equal. In such way we don't introduce any possible stiffness, corresponding to the preference for one direction.

The set of spring constants chosen for the 3 dimensional model are reported in Table 4.1. Because of its key role, we will study several cases for the contact spring constants.

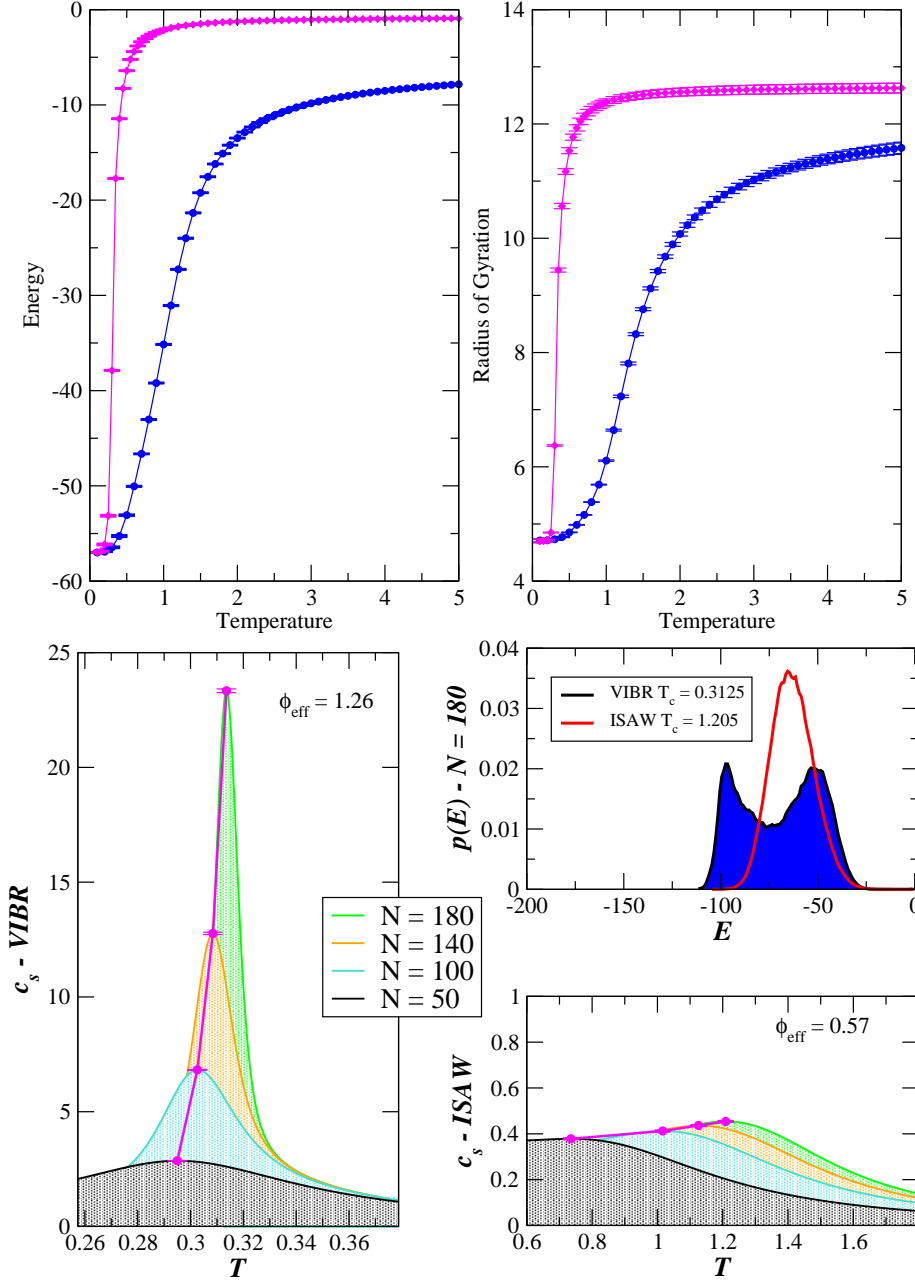


Figure 4.1: **Results in three dimensions. Top half:** The energy and the radius of gyration as a function of the temperature. The blue curve is ISAW, the magenta curve is  $k_{co} = 2$ . The length of the polymer is  $N = 100$  monomers and the data are reweighted with MHM method. **Bottom half:** The specific heat  $c$  of the model with vibrations (VIBR, left panel) and the classical model (ISAW, right bottom panel) as a function of the temperature; the distribution of the energy  $p(E)$  at the critical temperature (right top panel). The specific heat curves are plotted for several lengths of the polymer and are obtained from a multiple histogram method technique [52]. The maxima of the specific heat are marked with their error bar; the other points have similar errors. The result of the fit with a power law  $c = KN^{2\phi_{\text{eff}}-1}$  is reported. The population of the energy levels at the critical temperature are drawn from a simulation at the maximum heat capacity and length  $N = 180$ .

Table 4.1: **Typical spring constants in the VIBR partition function.**  
In adimensional units (see section 2.7.1).

spring constant	eigenvalue
$k_{bb} = 150$	$\lambda^{bb} = 300$
$k_{be} = 1$	$\lambda^{be} = 1.56$
$k_{tr} = 1$	$\lambda^{tr} = 1.5$
$k_{ga} = 0.6$	$\lambda^{ga} = 1.5$
$k_{co} \in [0, 2]$	

## 4.2 3 dimensions: high and low temperatures

In this section we analyze the very low and very high temperature situation.

The very low temperature corresponds to the compact phase; it is revealed by low energy and low Radius (see Figure 4.1). In the ISAW model, at temperature  $T = 0$  the average energy is equal to minus the maximum number of contacts, because the other configurations has a Boltzmann weight equal to 0. This is true also if entropic weights are taken into account, because they factorize:

$$\langle E \rangle = \frac{E_{min} \sum_{\gamma|E_{min}} w(\gamma)}{\sum_{\gamma|E_{min}} w(\gamma)} = E_{min} \quad (4.1)$$

The thermodynamics of the 0 temperature state is determined by all the maximally compact configurations which are the same in both models, as well as  $\langle E \rangle(T = 0)$ . Other observables are calculated on the same ensemble and weighted by the respective entropic factor. We may expect that, given the similar topology of the compact walks, the entropic weight may be quite uniform and the observables not much affected. The Radius of gyration for example is

$$\langle R_{gyr} \rangle = \frac{\sum_{\gamma|E_{min}} R_{gyr} w(\gamma)}{\sum_{\gamma|E_{min}} w(\gamma)} \neq \langle R_{gyr} \rangle_{ISAW}; \quad (4.2)$$

but Figure 4.1 shows that  $\langle R_{gyr} \rangle \sim \langle R_{gyr} \rangle_{ISAW}$ .

At high temperature, in particular in the limit  $\beta \rightarrow 0$ , the thermodynamic average of the model is realised over the whole ensemble of walks:

$$\langle E \rangle = \frac{\sum_{\gamma} w E}{\sum_{\gamma} w} \quad (4.3)$$

and similarly for the Radius of gyration. The result of Figure 4.1 reveals that the entropic vibrations tends to prefer the more open configuration, which



was already clear in the previous chapter. Nevertheless, the interest of this result is limited to that because the high temperature certainly corresponds to a state where the small fluctuations assumption is not valid (cfr criterion presented in section 2.6).

### 4.3 3 dimensions: first order phase transition

The contacts energy represents the *interesting* part of the total energy of the walk with vibrations, in the sense pointed out in section 2.3. Being proportional to the opposite of the number of contacts,  $E = -\varepsilon n_{co}$ , the contact energy is an increasing function of the temperature and changes from the lowest value equal to minus the number of contacts of the most compact walks at zero temperature, equation (4.1), to the highest value equal to the average number of contacts of all the walks weighted by the entropic weight, equation (4.3). The transition between the low and the high temperature is manifested by a pronounced variation of  $E$  at a temperature delimiting the two phases; in the thermodynamic limit (infinite number of monomers  $N \rightarrow \infty$ ) the variation approaches the singularity and is localized at *the* critical temperature, or  $\Theta$  point,  $T_\Theta$ . The singularity in the energy may be a discontinuity of the energy itself (discontinuous transition) or of one of its derivatives (continuous transition). In both cases, at finite  $N$ , the temperature at which the variation of the energy is more pronounced, is picked out by the maximum of its derivative, the specific heat

$$c = \frac{1}{N} \frac{dE}{dT} = \frac{\langle E^2 \rangle - \langle E \rangle^2}{Nk_B T^2}. \quad (4.4)$$

In the context of continuous transitions, scaling arguments relate the position and the height of the maximum with the length of the polymer. These arguments are based on the assumption of homogeneity of the free energy as a function of the temperature and the size. The free energy is obtained from the partition function

$$F(T, N) = -k_B T \ln Z(T, N), \quad (4.5)$$

and in the case of self avoiding walks it is a function of two parameters, the temperature and the number of monomers. Close to a critical point the free energy is assumed to be the sum of an analytical part, which is regular, and a singular part which is responsible of the divergence of some derivative. The homogeneity of the singular part of the free energy takes the form:

$$f_s(t, N) = N^{-1} h(tN^\phi), \quad (4.6)$$

where  $t = \frac{T-T_\Theta}{T_\Theta}$  is the reduced temperature and  $N$  the number of monomers. The specific heat is proportional to the second derivative of the free energy with respect to the temperature

$$c_s = \frac{\partial}{\partial t} \frac{\partial}{\partial \beta} \beta f_s(t=0) \sim N^{2\phi-1} \quad (4.7)$$

and it diverges in the thermodynamic limit if the cross-over exponent  $\phi > \frac{1}{2}$ . In figure 4.1 it is evident that the singularity of the specific heat grows much

faster in the model with vibrations ( $k_{co} = 2$ ) than in the standard ISAW. A linear fit of the data with the function  $\ln c_s = (2\phi_{eff} - 1) \ln N + A$  gives very diverse results in the two cases.

1. The effective exponent  $\phi_{eff}$  of the VIBR model is bigger than 1; this result is not physical because the specific heat is proportional to the variance of energy per monomers and should not be more than extensive ( $\phi \leq 1$ ). The found value indicates the coexistence between two populations, as we analyze later. If the transition is first order, the cross over exponent is  $\phi = 1$ .
2. The effective exponent of the ISAW model is higher than the theoretical value  $\phi = \frac{1}{2}$ ; this is manifestation of finite size and logarithmic corrections [26].

The top left graph of figure 4.2 reports the height of the critic specific heat as a function of  $N$  in a log-log scale. All these curves show a cross-over regime, driven by two parameters: the length and the contact spring constants.

At increasing length and fixed spring constant, the slope increases and in some cases ( $k_{co} = 2$ ) it corresponds to a cross over exponent bigger than 1, as already mentioned. This is the regime where the bimodal distribution of the energies is developing with increasing length: looking at the middle series of distributions (blue distributions in figure 4.2), two peaks are distinguishable only for lengths  $N \geq 140$ . In Figure 4.3 the bimodal distributions (lengths  $N \geq 140$ ) are represented as the sum of two gaussians functions

$$p(E) \sim \frac{1}{Z} \left[ x e^{-\frac{(E-E_-)^2}{2\sigma_-^2}} + (1-x) e^{-\frac{(E-E_+)^2}{2\sigma_+^2}} \right] \quad (4.8)$$

with averages and standard deviations given by the best least square fit of the data. To explain the super-extensive divergence of the specific heat, we represent the standard deviations growing with  $N$  as

$$\sigma_+ \propto N^{\alpha_+}, \quad \sigma_- \propto N^{\alpha_-},$$

and the energy difference between the two maxima also growing as

$$E_+ - E_- \propto N^x.$$

It can be shown that the energy variance (and the heat capacity) divergence depends on both the increasing distance between the two peaks,  $E_+ - E_-$ , and the standard deviation scaling:

$$\langle E^2 \rangle - \langle E \rangle^2 \propto \frac{(E_+ - E_-)^2}{N^{|\alpha_+ - \alpha_-|}}. \quad (4.9)$$

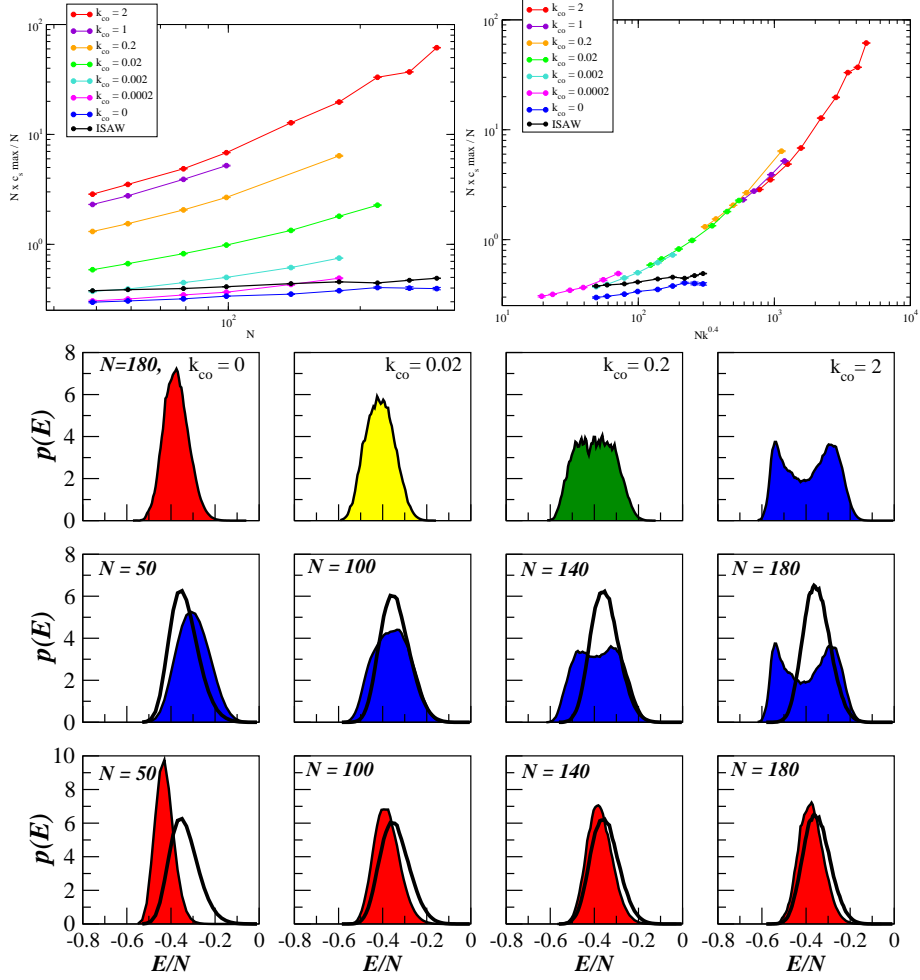


Figure 4.2: **Scaling properties of the transition in 3 dimensions.** **Top:** maximum of the specific heat with respect to the size at several contact spring constants. On the right the collapse of the curves on a function  $c(Nk^{0.4})$  is tried. **Series of distributions:** probability distribution of the contacts energy per monomer at the critical temperature (the maximum of specific heat). *Line filled with color:* VIBR, *Line not filled:* ISAW. **Top row:** fixed length (180 monomers) and varying contacts spring constant; **middle row:** fixed contacts spring constant ( $k_{co} = 2$ ) and varying length; **bottom row:** fixed contacts spring constant ( $k_{co} = 0$ ) and varying length. The distributions are calculated from canonical simulations and multiple histogram method.

The bimodal distribution is clearly manifested in the model with  $k_{co} = 2$  and length  $N \geq 140$ . The power law divergence calculated on those data (the  $N \geq 140$  part of the red curve of the top panel in figure 4.2) give an effective cross-over exponent  $2\phi_{eff} \sim 2.97$ , expressing the divergence of the left hand side of equation (4.9). The divergence of the average and standard deviation of the peaks was calculated on a fit of the distribution of the energy with a double gaussian. Inserting the values (see Figure 4.3) in the right hand side of equation (4.9) we obtain a good agreement with the observed cross-over exponent:

$$\frac{N^{2 \cdot 1.82}}{N^{1.39 - 0.69}} \sim N^{2.94}$$

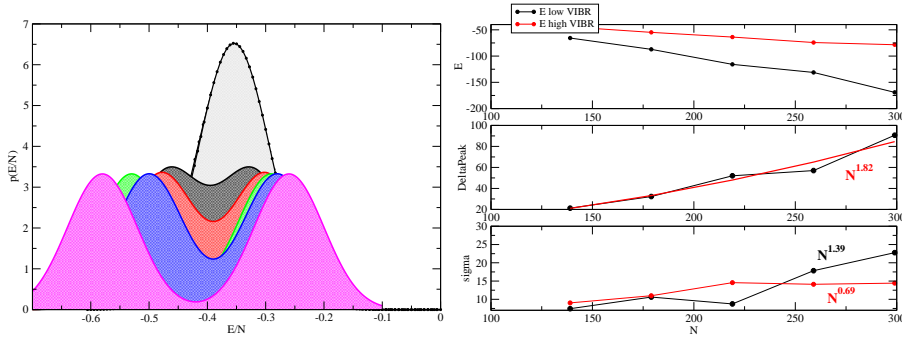


Figure 4.3: **Gaussian fit of double peaks.** **Left:** double gaussian fit of critical energy distributions for chain length  $N = 140, 180, 220, 260, 300$ . The estimated position of the two peaks ( $E_+$  and  $E_-$ ) are reported in the **top right graph**. The energy difference and the evaluated divergence power law is reported in the **middle right graph**, the standard deviations and their power law fit  $N^{\alpha_-}$  and  $N^{\alpha_+}$  are reported in the **bottom right graph**.

On the other hand the first order behavior is triggered also by the contacts spring constant: at fixed length ( $N = 180$ , for example, the top series of distributions of Figure 4.2), the double population is not distinguishable if  $k_{co} \leq 0.2$ . At fixed length (abscissa of the specific heat graph) and increasing contact spring constant  $k_{co}$ , the maximum of the specific heat grows. By generalizing the cross-over scaling relation to the new parameter  $k_{co} \equiv k$ , we write the specific heat maximum as a function of  $(N^{\phi_k} k)$ , which together with the classic scaling relation (eq. 4.7) gives:

$$c_s \sim (N^{\phi_k} k)^{2\phi-1} \sim k^{\frac{2\phi-1}{\phi_k}} \quad (4.10)$$

at fixed  $N$ . The maximum specific heat diverges as a power between  $k^{0.34}$  ( $N = 50$ ) to  $k^{0.42}$  ( $N = 100$ ). The variation is still imputable to the crossover situation. In particular, fixing  $\phi = 1$  (first order transition), the specific heat diverges as  $k^{1/\phi_k}$ . The top right graph of Figure figResults:gaussianFitOfPeaks

proposes a quite good collapse of the curves obtained with  $1/\phi_k = 0.4$ . The meaning of such a relation is that the stable regime where the two peaks are distinguishable can be reached either growing  $N$  and keeping fixed  $k$ , or viceversa. The exponent  $\phi_k > 1$  implies that the first method is more efficient.

At very small stiffness ( $k_{co} = 0.0002$ ) and length ( $N \leq 80$ ) the specific heat is very similar to the  $k_{co} = 0$  case, but it detaches from it as the length increases. This fact seems to exclude the possibility of a finite, though very small, critical  $k_{co}$  under which the system would behave as if  $k_{co} = 0$ . The available data show also that the  $k_{co} = 0$  specific heat is *parallel* to the ISAW model so it belongs to the standard universality class. The energy distribution at the critical temperature (bottom line of) also approaches well to the ISAW case. At the same time, the indication coming from the sensitivity of the vibrational spectra to the value of  $k_{co}$  (chapter 3) definitely corroborates the conclusion that two universality classes are individuated by  $k_{co} = 0$  and  $k_{co} \neq 0$ .

#### 4.3.1 Van't Hoff test

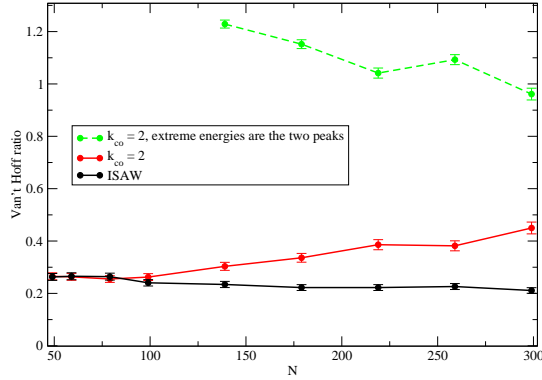


Figure 4.4: **The Van't Hoff ratio**, equation (4.14). In the ISAW (black curve) and  $k_{co} = 2$  (red curve) cases, the unfolded energy is the energy at infinite temperature  $H_U = \langle H \rangle (T = \infty)$  and the folded energy is the minimum energy obtained from any simulation at each length  $H_F = H_{min}$ , which is the best estimation of the energy at zero temperature. The green curve refers also to  $k_{co} = 2$ , but the folded and unfolded energies are the positions of the two coexisting peaks of the energy distribution at the critical temperature (if distinguishable:  $N \geq 140$ ), calculated from gaussian fits (see Figure 4.3).

The Van't Hoff test is an experimental criterion for determining how much the calorimetry of a process corresponds to a two-state transition [79, 45]. The criterion consists in the comparison between the measured

enthalpy difference between initial and final state and the enthalpy difference induced from the heat capacity curve in the hypothesis that the process is *purely* two state.

A pure two state transition from an initial state at energy  $H_U$  and a final state at energy  $H_F$  is described by the following partition function:

$$Z_2 = e^{-\beta(H_F - TS_F)} + e^{-\beta(H_U - TS_U)}. \quad (4.11)$$

The two states may also be endowed of different entropies,  $S_U$  and  $S_F$ , which don't take part in the following calculation. At the temperature of transition  $T_m$ , defined as the temperature at which the two populations coexist with identical probabilities, the heat capacity is

$$C(T_m) = \frac{\langle H^2 \rangle - \langle H \rangle^2}{k_B T_m^2} = \frac{\frac{1}{2}(H_F^2 + H_U^2) - (\frac{1}{2}H_F + \frac{1}{2}H_U)^2}{k_B T_m^2}. \quad (4.12)$$

Inverting this relation, the Van't Hoff enthalpy difference results

$$\Delta H_{VtH} = H_U - H_F = 2T_m \sqrt{k_B C(T_m)}. \quad (4.13)$$

If the enthalpy difference measured in a calorimetry experiment is equal to the van't Hoff enthalpy, than the process is most likely a two state process. Instead, if the transition passes through intermediate states, the experimental heat capacity broadens and the measured enthalpy difference is smaller than the van't Hoff enthalpy: in this case the ratio of the van't Hoff over the measured enthalpy

$$\frac{\Delta H_{VtH}}{H_U - H_F} \leq 1 \quad (4.14)$$

is sometimes interpreted as the number of cooperative units inside the system [68]. Exception to this empirical behavior (in particular ratio bigger than unity) are of course predicted and tested on simple models [45].

The models of homopolymers presented here are not the best candidate for a two state transition because they lack of a precise definition of the *unfolded* and the *folded* state. Nevertheless we can easily apply the van't Hoff criterion once we agree on the definition of them. An obvious choice is to consider the folded state as the state of zero temperature and the unfolded state as the state of infinite temperature. The energy difference is extensive

$$\Delta H = H(\infty) - H(0) \propto N; \quad (4.15)$$

on the other hand the heat capacity scales with the power law  $C \propto N^{2\phi}$  and

$$\frac{\Delta H_{VtH}}{H_U - H_F} \propto \frac{N^\phi}{N} \quad (4.16)$$

which vanishes if  $\phi < 1$ . This result is coherent with the interpretation of a two-state process as a first order like transition.

The Van't Hoff ratio of the VIBR model ( $k_{co} = 2$ ) increases from a value around 0.25 to a value around 0.45 (figure 4.4), mostly in the region where two peaks are manifested ( $N \geq 140$ ); the ISAW value instead is barely decreasing, but the theoretical results  $\phi = \frac{1}{2}$  affirms that it vanishes in the thermodynamic limit.

Another possible choice for the definition of unfolded and folded state is the average energy of the two peaks, when they are detected (gaussian fits of figure 4.3) [66]. With this choice the van't Hoff ratio is bigger and close to 1 (green curve).

#### 4.4 3 dimensions: metric properties

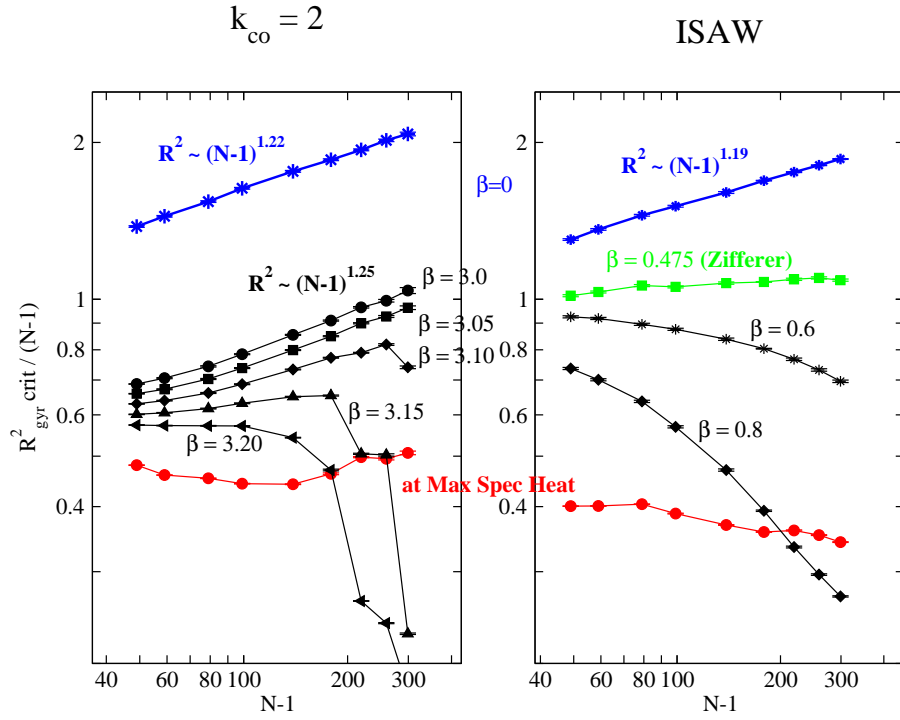


Figure 4.5: **Radius of gyration in 3 dimensions.** Average radius of gyration VS number of steps ( $N - 1$ ) at several temperatures (marked in the graph). The green series corresponds to the critical temperature calculated in [80]. The red series are calculated at the maximum of the specific heat and do not corresponds to a fixed temperature.

The Radius of gyration is the standard deviation of the monomer positions:

$$R_{gyr}^2 = \frac{1}{N} \sum_{i=1}^N (\vec{r}_i - \langle \vec{r} \rangle)^2 \quad (4.17)$$



for a chain of  $N$  monomers with position  $\vec{r}_i$  and average position (center of mass)

$$\langle \vec{r} \rangle = \frac{1}{N} \sum_{i=1}^N \vec{r}_i.$$

This is an important mathematical observable of a polymer because its scaling behavior (the function  $R_{gyr}(N)$ ) can be exactly predicted in several cases. In particular, for the ISAW, the three regimes of good solvent,  $\Theta$  point and compact phase correspond to the scaling relation

$$R_{gyr} \propto N^{\nu(T)}$$

with exponents of Table 1.1. The algorithm used in this work is adapted for the study of the high and critical temperature only, but the arguments presented in (section 4.2) indicates that the low temperature behavior of the two models should be the same.

#### 4.4.1 High temperature

The theoretical expected value for ISAW at high temperature is

$$\nu \sim 0.588;$$

this is a thermodynamic limit value, but it is already well reproduced by our simulations  $N \leq 300$ :

$$\nu(ISAW) = 0.595(5).$$

Notice that the given error estimation takes into account how much the data fits with the power law; finite size effects are not predictable from the quality of the fit and the exact value of the fit obtained at these length (to which the error estimation actually refers) may be still different than the thermodynamic limit value.

On the same lengths, the VIBR ( $k_{co} = 2$ ) high temperature scaling exponent is higher:

$$\nu(VIBR) = 0.610(5),$$

which may be traced back to the preference for more open configurations accorded by the configurational entropy. The behavior at the thermodynamic limit is still uncovered: a different metric exponent may stabilize or the ISAW regime may be approached. The second hypothesis is supported by the argument that in the thermodynamic limit the vibrational entropic cost of even a single contact could be prohibitive and the ensemble would reduce to the family of SAWs without contacts. As this ensemble maps on the standard ensemble of SAWs by splitting each step in two, it results that the metric exponent of no-contacts walk is the same as the SAW.

However, we remark again that the high temperature regime belongs to the region where the hypothesis of validity of the VIBR model (small fluctuations) is not valid, so the results are a good indication of the average influence of vibrational entropy (preference for open configurations) but they don't pretend to reproduce exactly the high temperature phase of polymers.

#### 4.4.2 $\Theta$ point

In the ISAW model, the polymer radius of gyration at the critical temperature scales as a random walk. Coherently with the hypothesis of homogeneity of the free energy at temperatures close to the critical point, the radius of gyration scales as

$$R_{gyr}(N, t) = bN^{1/2}g(Nt^{1/\phi}). \quad (4.18)$$

( $t$  is the reduced temperature). Following this relation, the  $\Theta$  point  $t = 0$  is the temperature at which the radius is proportional to the square root of the size; this is also the usual way for individuate the critical temperature of the  $N \rightarrow \infty$  polymer. The  $\Theta$  temperature of the ISAW on the diamond lattice was detected at  $\beta_{\Theta} \sim 0.475$  [80] and our simulations confirm the prediction (fig. 4.5, green curve).

In the VIBR model, we can not find a temperature with similar ideal behavior, if not limited to some finite size. For example, in Figure 4.5 (left panel) the radius of gyration at the temperature  $\beta = 3.20$  or  $\beta = 3.15$  is almost linear, similar to the ISAW  $\Theta$  point, up to the size at which the radius drops down crossing the red line of radius at the maximum of the specific heat. This is due to the fact that the finite size critical temperature (the maximum of the heat capacity) moves and is expected to approach the  $\Theta$  point only at infinite  $N$ , so a fixed temperature (like  $\beta = 3.20$  or  $\beta = 3.15$ ) is bigger than the maximum heat capacity temperature only up to a certain length of the polymer. Higher temperatures ( $\beta = 3.0$  or  $\beta = 3.05$ ), which don't drop down, are good candidates to represent the  $\Theta$  temperature: in those cases the square radius of gyration scales more than linearly and is compatible with a good solvent scaling.

It is important to observe that the scaling of the radius of gyration at the critical temperature corresponds to the scaling of the high energy peak. Indeed, assuming the two populations to scale with different laws,  $R_- \propto N^{\nu_-}$  and  $R_+ \propto N^{\nu_+}$  with  $\nu_- < \nu_+$ , at the critical temperature the average radius results:

$$\langle R_{gyr} \rangle(T_c) \propto \frac{1}{2}N^{\nu_+} + \frac{1}{2}N^{\nu_-} = \frac{1}{2}N^{\nu_+}(1 + N^{-(\nu_+ - \nu_-)}) \sim N^{\nu_+}. \quad (4.19)$$

Our results seem to indicate that the two populations correspond to a coil and a compact phase respectively.

## 4.5 2 dimensions: a more abrupt second order phase transition

The comparison between the coil to globule transitions manifested by the ISAW and the VIBR models highlights the not negligible contribution coming from the local fluctuations to the thermodynamic of the system. In 3 dimensions we have shown that the local fluctuations induces a first order phase transition where the standard ISAW transition is continuous with cross-over exponent  $\phi(3d - \text{ISAW}) = \frac{1}{2}$ . So, in terms of the universal quantities, we can resume the contribution of the entropic ingredient saying that the cross-over exponent is larger:  $\phi(3d - \text{VIBR}) = 1$ .

In this section we show that the same effect is induced in 2 dimensions where the cross over exponent

$$\phi(2d - \text{ISAW}) = \frac{3}{7}$$

is superseded by  $\phi(2d - \text{VIBR}) > 0.8$  which implies a new universality class. In terms of the standard classification of phase transitions by the degree of the discontinuous derivative, the 2d-ISAW model has singularity in the third derivative of the free energy, while the 2d-VIBR model is singular at the second derivative (the specific heat): similarly to the 3 dimensional case, the kind of transition gains one degree.

We study the case with contact spring constant  $k_{co} = 0.1k_{bb} > 0$  and the case with  $k_{co} = 0$  in order to show that this parameter has a key role in triggering the new universality class. The other spring constants are reported in Table 4.2 and are chosen following two criteria: the backbone constant is bigger than the others and the ratio  $k_{\text{flat}}/k_{\text{right}}$  provides no preference for one direction, similar to the equilibrium between *trans* and *gauche* defined in the 3 dimensional case.

Table 4.2: **Spring constants in the VIBR partition function in 2 dimensions.** In adimensional units (see section 2.7.1).

spring constant	eigenvalue
$k_{BB} = 25$	$\lambda^{bb} = 50$
$k_{\text{flat}} = 4$	$\lambda^{be} = 24$
$k_{\text{right}} = 6$	$\lambda^{tr} = 24$
$k_{co} = 2.5 \text{ or } 0$	$\lambda^{ga} = 5 \text{ or } 0$

The scaling hypothesis on the free energy, equation (4.6), is identical in 2 dimensions, with the proper critical exponent. The value  $\phi(2d - \text{ISAW}) < \frac{1}{2}$  implies that the singular part of the specific heat, equation (4.7), is vanishing: the singularity is in the third derivative of the free energy. As a

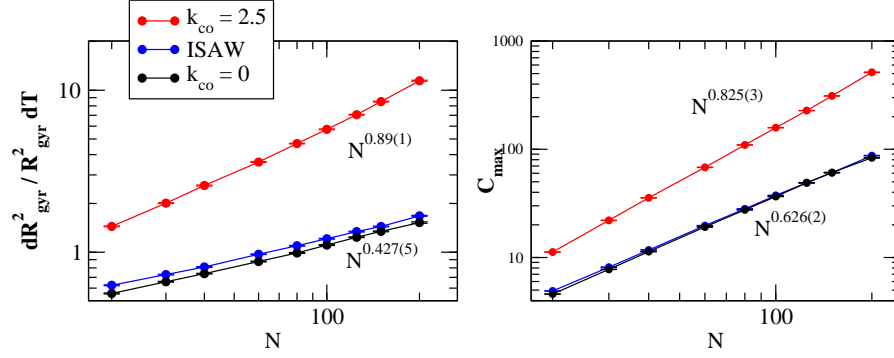


Figure 4.6: **Scaling of the model in 2 dimensions.** The maximum of the relative radius derivative (**left**) and the heat capacity (**right**) as a function of  $N$ . Least square fits with the power law  $N^\phi$  are reported in the graph.

consequence, the scaling of the maximum of the heat capacity is mixed with the analytical part and can not be interpreted as the power law  $N^{2\phi}$ : the observed exponent (Figure 4.6, left, blue line) describes the divergence of a function of which we don't know the behavior at these sizes, but it should be extensive ( $\phi = \frac{1}{2}$ ) in the thermodynamic limit: this datum is then not reliable for the evaluation of the cross-over exponent.

The scaling relation of the radius of gyration at small reduced temperature  $t$ :

$$R_{gyr}^2 \sim N^{2\nu} g(tN^\phi). \quad (4.20)$$

Deriving with respect to the temperature and taking the ratio with the radius itself we obtain the relative radius derivative

$$\frac{1}{R_{gyr}^2} \frac{\partial R_{gyr}^2}{\partial t} (T_{max}(N)) \sim N^\phi. \quad (4.21)$$

which diverges as  $N^\phi$ , whatever  $\phi$  is. The relative radius derivative is calculated from measures of the average energy and radius:

$$\frac{\partial R_{gyr}^2}{\partial t} \propto \frac{\partial R_{gyr}^2}{\partial T} = \frac{\langle R_{gyr}^2 E \rangle - \langle R_{gyr}^2 \rangle \langle E \rangle}{k_B T^2}. \quad (4.22)$$

Equation (4.21) is evaluated at the maximum  $T_{max}(N)$ , so that the value of the scaling function  $g(x_{max})$  is fixed and does not influence the critical divergence. This quantity permits a good evaluation of  $\phi$ : even at small sizes, we obtain value compatible with the theory (Figure 4.6).

With the same method the cross-over exponent of the VIBR model is

$$\phi(2d\text{- VIBR}) = 0.87 \pm 0.01.$$

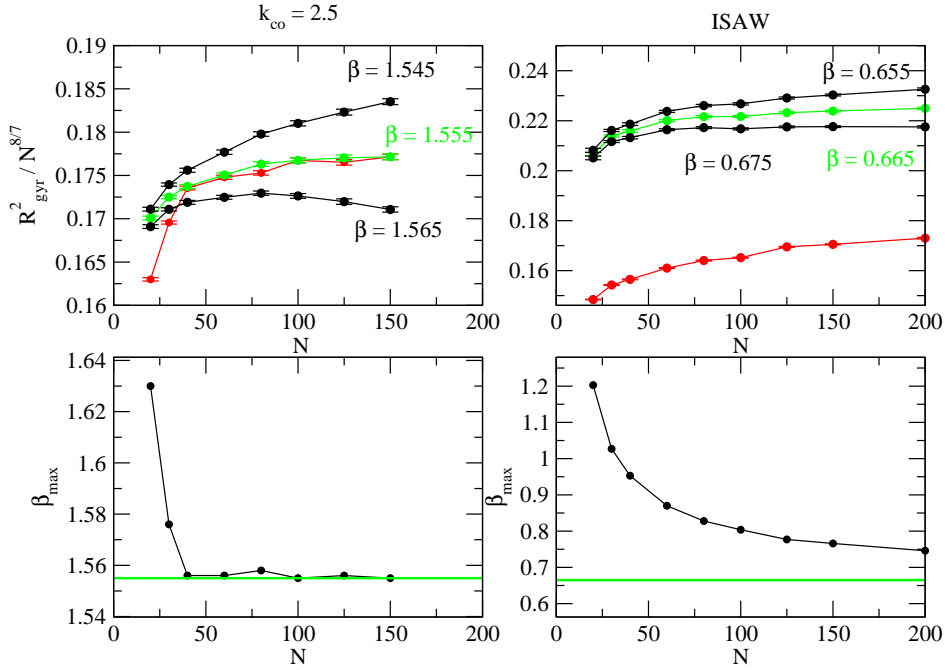


Figure 4.7: **Metric properties of the 2 dimensions model.** The average square radius of gyration divided by  $N^{8/7}$  (**top panels**) is calculated at fixed temperatures (black and green curves) or at the  $\beta_{\text{max}}(N)$  (red curves), which is reported in the **bottom panels**. The green continuous line in the bottom panels indicates the estimated critical temperature.

Interestingly, a very similar value results from the scaling of the heat capacity: despite the mixing with the regular extensive part the high cross-over exponent makes the singular part be predominant already at short sizes.

The radius of gyration at the critical temperature of the ISAW model scales as

$$\nu = \frac{4}{7}. \quad (4.23)$$

The critical temperature was estimated as  $\beta_\Theta \sim 0.665$  from previous works [35]; this can be detected by the temperature at which  $\frac{R_{gyr}}{N^\nu} \sim \text{const}$  and our data confirm the result (Figure 4.7, top left, green curve): note that the data picked at close temperatures (black curves) tend to leave the straight line.

The temperature  $T_{max}(N)$  is a finite size critical temperature: it is an empirical way for localize the transition point where the variation of the observables is maximum (the maximum of their derivative). The scaling relation (4.20) implies that the polymer behaves critically (i.e.  $R \propto N^\nu$ ) in a finite region of temperatures (and not just AT the critical temperature) such that  $tN^\phi$  is small. The finite size critical temperature is a priori different from the  $\Theta$  temperature, but it is contained in the critical region, so it approaches the  $T_\Theta$  as

$$\frac{T_{max} - T_\Theta}{T_\Theta} \propto N^{-\phi}. \quad (4.24)$$

The approach is faster, the bigger the value of  $\phi$ ; the diverse behavior depending on  $\phi$  is pointed out in Figure 4.7. In the VIBR model  $\beta_{max}$  reaches a stable value already at length around 40 monomers (bottom panel) at the inverse temperature:

$$\beta_\Theta = 1.555 \pm 0.003. \quad (4.25)$$

Surprisingly the scaling of the radius at that temperature (green curve in the top panel) is compatible with the scaling of ISAW, so we conclude that the metric exponent at the  $\Theta$  point is left unchanged by the vibrational contribution.

## 4.6 Convexity of the entropy

In the canonical ensemble the frequency of observation of a given energy, the probability  $p(E)$ , depends on the temperature  $T = \frac{\beta}{k_B}$  and the logarithm of the *volume* of the phase space at that energy,  $S = k_B \ln \mathcal{V}(E)$ :

$$p(E) = \frac{e^{-\beta(E-TS(E))}}{Z}. \quad (4.26)$$

In the ISAW model (partition function 2.8) the walks with the same energy have the same probability, so the volume  $\mathcal{V}(E)$  is the number of walks with  $(-E)$  contacts,  $\mathcal{V}(E) = \mathcal{N}(E)$ , and its logarithm is the microcanonical entropy  $S_{isaw}(E) = k_B \ln \mathcal{N}(E)$ . In the model with vibrations (partition function 2.14), each walk has statistical weight  $e^{-\beta E} W(\gamma)$  and the microcanonical entropy is

$$S(E) = k_B \ln \sum_{\gamma|E} W(\gamma) = k_B \ln \mathcal{N}(E) + k_B \ln \frac{\sum_{\gamma|E} W(\gamma)}{\mathcal{N}(E)} \quad (4.27)$$

$$= S_{isaw}(E) + S_{vib}(E). \quad (4.28)$$

The term  $S_{vib}(E)$  (equation 2.15) is the difference between the microcanonical entropy of VIBR and ISAW and it is the responsible of the different behavior manifested by the two models. In the first place, the critical heat capacities, studied in the previous sections for both 2 and 3 dimensions, reveals scaling properties indicating a sharper transition in the VIBR model than in ISAW. In this section we illustrate that this is due to a convexity of the vibrational microcanonical term  $S_{vib}(E)$ .

The relation between canonical (fixed temperature) and microcanonical (fixed energy) ensemble can be formulated through the saddle point approximation of the canonical probability of  $E$

$$p(E) \equiv e^{\ln p(E)}.$$

The saddle point approximation consists in a gaussian function centered at the maximum of the function:

$$e^{h(E)} \sim e^{h(E_{max}) + \frac{1}{2} \frac{d^2 h}{dE^2} (E-E_{max})^2} \quad (4.29)$$

$$= e^{h(E_{max}) - \frac{(E-E_{max})^2}{2\sigma^2}}. \quad (4.30)$$

The maximum is located at the point where  $\frac{dh}{dE} = 0$  and

$$\frac{dh}{dE} = 0 \quad \text{and} \quad \frac{d^2 h}{dE^2} \equiv -\frac{1}{\sigma^2} < 0. \quad (4.31)$$

In the case of the probability equation (4.26),

$$h(E) \equiv \ln p(E) = -\frac{1}{k_B T} (E - TS(E)) - \ln Z, \quad (4.32)$$

the energy of maximum probability is located at  $E_{max}$  such that

$$\frac{dh}{dE}(E_{max}) = \frac{S'(E_{max})}{k_B} - \frac{1}{k_B T} = 0 \iff S'(E_{max}) = \frac{1}{T} \quad (4.33)$$

and

$$\frac{d^2 h}{dE^2}(E_{max}) = \frac{S''(E_{max})}{k_B} < 0. \quad (4.34)$$

Through the standard relation between heat capacity and variance,

$$C(T) = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} = \frac{\sigma^2}{k_B T^2},$$

we obtain the relation between heat capacity and microcanonical entropy:

$$C(T) = -\frac{1}{T^2 S''(E)}, \quad (4.35)$$

where the second derivative of the microcanonical entropy,  $S''(E)$  is calculated at the average energy of temperature  $T$ :

$$\langle E \rangle(T) = \frac{\sum_E E e^{-\beta(E - TS(E))}}{Z}.$$

The one-to-one correspondence between canonical observables (average energy and heat capacity at *fixed temperature*) and microcanonical properties (first and second derivatives of  $S(E)$ ) is valid as long as the canonical distribution can be approximated with one Gaussian function. A necessary condition for this is the concavity of the microcanonical entropy,  $S''(E) < 0$ , which implies that  $S'(E) = \frac{1}{T}$  is a decreasing function of the energy: at each temperature, only one  $E_{max}$  exists and the saddle point approximation can be built around it. In the ISAW model the condition is verified: the derivative of the microcanonical entropy,  $\frac{dS}{dE}$  VS  $E$  (left top panel of Figure 4.8, **microcan.**), superposes perfectly on the canonical average energy, plotted as  $\beta$  VS  $\langle E \rangle$  (**canonical**). In the microcanonical ensemble one temperature corresponds to one energy (respectively the ordinate and the abscissa of the colored points reported on the microcanonical curve), while in the canonical ensemble one temperature corresponds to a distribution of energies (the distributions reported in the bottom panels).

The one-to-one correspondence breaks off if the microcanonical entropy is convex in some region ( $S''(E) > 0$ ): in this case a microcanonical temperature ( $\frac{dS}{dE}$ ) in the convex region corresponds to more than one energy (right



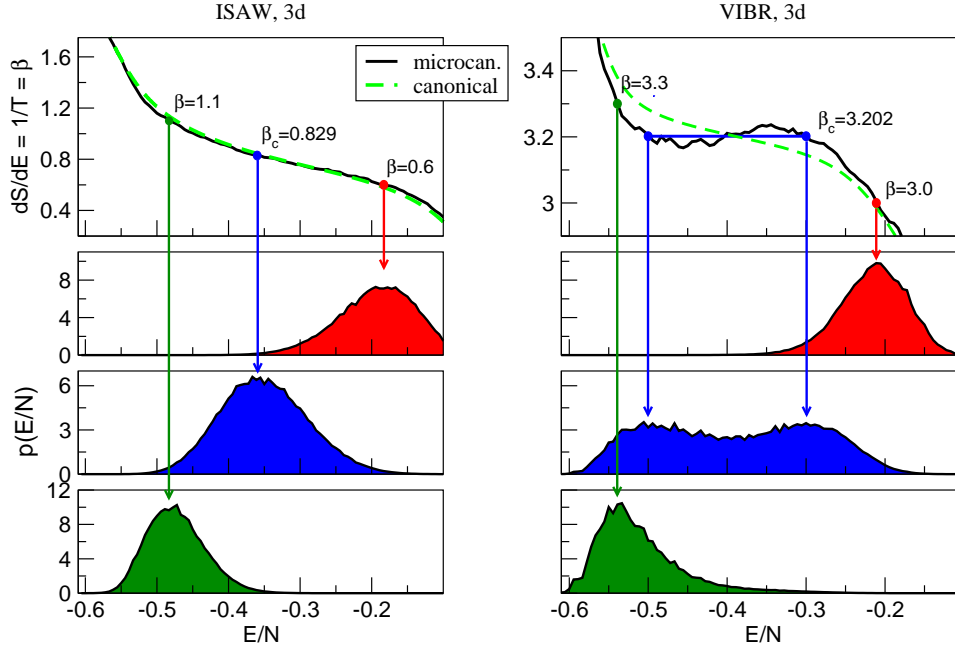


Figure 4.8: **Microcanonical entropy and canonical energy in 3 dimensions.** The top panels report the first derivative of the total microcanonical entropy  $S'(E)$  (black lines) calculated from canonical simulations through the Multiple Histogram Method and a low noise differentiator (Lanczos method, see section 2.7). The canonical average energy  $\beta(\langle E \rangle)$  is the green dashed line. The other panels report the distribution of energy at the critical temperatures and two others (reported in the graph). Arrows link the temperature and the energy.

top panel of Figure 4.8, *microcan.*). The canonical distribution of energies splits in two populations (right bottom panels): the position of the two maxima can be approximately found from a Maxwell construction on the microcanonical curve (the straight blue line in the top right panel). The average canonical energy (the green line) does not follow strictly the Maxwell distribution because of the finite size of the system: in the thermodynamic limit the coexistence would be observed only at the critical temperature and the average energy would manifest a finite jump.

Keeping this general analysis in mind, we proceed writing the relation between heat capacity and microcanonical entropy (equation 4.35) explicitly expressing the standard and the vibrational contributions, as in equation (4.30). At the critical temperature (of the VIBR model) the heat capacity is

$$C(T_c) = -\frac{1}{T_c^2 (S''_{isaw}(E_c) + S''_{vib}(E_c))} \quad (4.36)$$

where it should be clear that the heat capacity  $C$ , the critical temperature  $T_c$  and the average energy at the critical temperature  $E_c$  refer all to the model with vibrations. From the thermodynamic results of the previous sections, the 2 dimensional and 3 dimensional VIBR models show two behaviors:

- in **two dimensions**, the energy distribution is always single peaked, also at the transition point which is continuous; in this case equation (4.36) is valid and the microcanonical analysis corresponds exactly to the canonical results (Figure 4.9);
- in **three dimensions**, as shown before, the correspondence is broken close to the critical temperature; in this case the microcanonical analysis (right size of equation (4.36)) gives a negative result, while the canonical heat capacity is positive: the first is the derivative of the black (microcanonical) curve of the top right panel of Figure 4.8, the second is the derivative of the green (canonical) curve of the same figure.

We analyze separately the two cases.

#### 4.6.1 2 dimensions

In two dimensions the heat capacity of the model with vibrations diverges faster than in the classical model. From equation (4.36) we obtain the relation between the two cases:

$$C_{VIBR}(T_c) = \frac{C_{ISAW}(T_{isaw}(E_c))}{\left(1 + \frac{S''_{vib}}{S''_{isaw}}\right)} \left(\frac{T_{isaw}^2(E_c)}{T_c^2}\right). \quad (4.37)$$

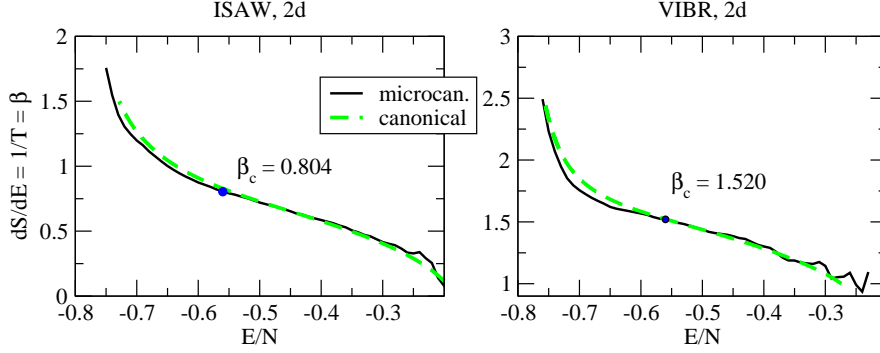


Figure 4.9: **Microcanonical entropy and canonical energy in 2 dimensions.** The first derivative of the total microcanonical entropy  $S'(E)$  calculated from canonical simulations through the Multiple Histogram Method and low noise differentiator (Lanczos method, see section 2.7).

Here the heat capacity of the VIBR model  $C_{VIBR}(T_c)$  is calculated at the critical temperature (maximum of the heat capacity), while the heat capacity of the ISAW model,

$$C_{ISAW}(T_{isaw}(E_c)) = -\frac{1}{T_{isaw}^2(E_c)S''_{isaw}(E_c)}$$

is calculated at the temperature  $T_{isaw}(E_c)$  obtained from the function  $S'_{isaw}$  at  $E_c$  (Figure 4.9, left). The energy  $E_c$  can be equal or not to the critical energy of the standard model,  $E_c^{ISAW}$ :

- if  $E_c = E_c^{ISAW}$ , the temperature  $T_{isaw}(E_c)$  is also the critical temperature of the standard model and the ISAW heat capacity diverges critically:

$$C_{ISAW}(T_{isaw}(E_c)) \propto N + N^{2\phi}; \quad (4.38)$$

- if  $E_c \neq E_c^{ISAW}$ , the temperature  $T_{isaw}(E_c)$  is *not* critical and the heat capacity is extensive:

$$C_{ISAW}(T_{isaw}(E_c)) \propto N \quad (4.39)$$

Nevertheless, in 2 dimensions  $\phi_{ISAW} < \frac{1}{2}$ , so the extensive part of the heat capacity is dominant in both cases. From equation (4.37) we obtain:

$$C_{VIBR}(T_c) \sim \frac{N}{1 + \frac{S''_{vib}}{S''_{isaw}}} \quad (4.40)$$

As the numerical results show that the heat capacity of the model with vibrations diverges faster than  $N$ ,

$$C_{VIBR}(T_c) \sim N^{2\phi(2d - \text{VIBR})} \quad \phi(2d - \text{VIBR}) = 0.87 > \frac{1}{2}$$

the denominator of equation (4.40) is expected to vanish:

$$S''_{vib} = -S''_{isaw} + o(S''_{isaw}) \quad (4.41)$$

with  $o(S''_{isaw})$  a function much smaller than  $S''_{isaw}$  in the limit  $N \rightarrow \infty$ ,  $\frac{o(S''_{isaw})}{S''_{isaw}} \rightarrow 0$ . It follows:

$$S''_{vib}(E_c) > 0. \quad (4.42)$$

In other words the vibrational part of the entropy is convex around the critical energy: this does not implies a convexity of the total entropy (which would be associated to a first order transition); we know instead that the transition is continuous. We conclude that in 2 dimensions

$$S''_{vib} + S''_{isaw} = o(S''_{isaw}) < 0.$$

Also, since the heat capacity is regular everywhere ( $C \propto N$ ), the vibrational entropy is expected to show a behavior like

$$S''_{vib}(E_c) = a \frac{1}{N} + o(1/N) \quad (4.43)$$

with  $a > 0$  and  $o(1/N) < 0$  a function vanishing faster than  $1/N$ . This result is valid *close to the critical energy*.

### 4.6.2 3 dimensions

In the 3 dimensional case, the relation between canonical heat capacity and microcanonical entropy is broken: the total microcanonical entropy is positive:

$$S''_{vib} + S''_{isaw} > 0,$$

which again implies the convexity of the additional term

$$S''_{vib} > 0. \quad (4.44)$$

It is interesting that assuming the same behavior as equation (4.43) can explain the change of phase in 3 dimensions, that is the convexity of the total microcanonical entropy. Indeed, in the critical region the ISAW heat capacity diverges as  $C \propto N \ln N$ , so the microcanonical entropy:

$$S''_{isaw} \propto -\frac{1}{N \ln N} \quad (4.45)$$

Assuming behavior equation (4.43) for the vibrational entropy at the energies close to the critical energies, the total entropy would follow

$$S'' = S''_{vib} + S''_{isaw} \propto \frac{1}{N} - \frac{1}{N \ln N} > 0 \quad (4.46)$$

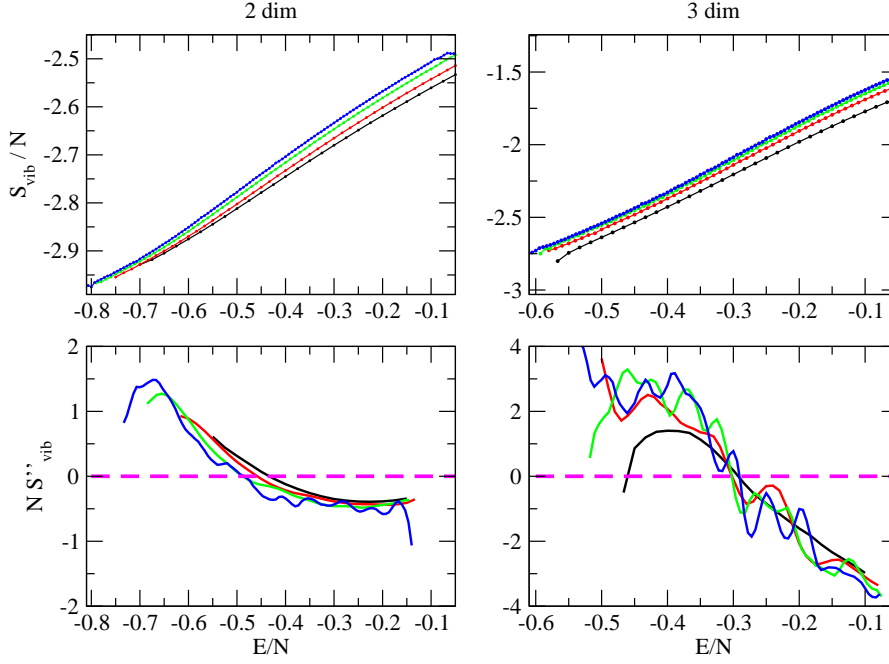


Figure 4.10: **Vibrational microcanonical entropy.** The vibrational entropy  $S_{vib}(E) = k_B \ln \langle W \rangle_E$  is calculated on configurations obtained from ISAW simulations (see the text for more details). The error bars are smaller than the size of the symbols (which are very small yet!). The second derivatives are calculated with low noise differentiator and some of the curves are smoothed by averaging over neighbor points. **Top:** vibrational microcanonical entropy per monomer with respect to the energy per monomer  $\frac{S_{vib}}{N}(E/N)$ . **Bottom:** the second derivative multiplied by  $N$  as a function of the energy per monomers  $NS''_{vib}(E/N)$ . The lengths of the chains are: **Left, 2 dimensions:**  $N = 40$  (black),  $N = 60$  (red),  $N = 100$  (green) and  $N = 150$  (blue); **Right, 3 dimensions:**  $N = 60$  (black),  $N = 100$  (red),  $N = 140$  (green) and  $N = 180$  (blue).

### 4.6.3 Vibrational microcanonical entropy

The observations of this section individuate that a common characteristic of convexity is shared by the vibrational microcanonical entropy term

$$S_{vib}(E) = k_B \ln \langle W \rangle_E. \quad (4.47)$$

In particular the convexity of  $S_{vib} = S_{VIBR} - S_{ISAW}$  around the critical energy is a necessary condition for the change of kind of transition in both 2 and 3 dimensions.

To verify that the condition is satisfied we run independent ISAW canonical simulations at several temperatures in order to collect configurations at all energies and calculate on those samples the average  $\langle W \rangle_E$  and the vibrational microcanonical entropy  $S_{vib}(E)$ . The result is very precise and illustrated in the top graphs of Figure 4.10: the error bars are smaller than the symbols. By careful inspection one can distinguish the convex region at low energies. Nevertheless, despite the high precision of this data, the second derivative (bottom graph) calculated by a low noise Lanczos differentiator were still too noisy for clearly distinguish the convex from the concave part ( $S'' > 0$  and  $S'' < 0$ ). The fluctuations increases with  $N$ . The reported curves are smoothed by averaging over neighbor points, in particular the average was performed over the 5 neighbors for the red curves ( $N = 60, 2d$  or  $N = 100, 3d$ ), 10 neighbors for the green curves ( $N = 100, 2d$  or  $N = 140, 3d$ ) and 15 neighbors for the blue curves ( $N = 150, 2d$  or  $N = 180, 3d$ ). The black curves were not smoothed. The numerical derivation is very sensitive to local fluctuations of the original data, but the message emerges: all the curves are convex in the low region of energies and concave in the high region. This explains the phenomenology of the coil to globule transition of the VIBR model.

## 4.7 Conclusions

The universality class of a model depends on long range interactions, while is not sensitive on the microscopic properties. As already mentioned in the first chapter, all ideal chains have the same metric exponent, no matter how the bonds are locally connected. But the excluded volume interaction introduces a new metric exponent; this interaction acts also between monomers far away along the chain, so it clearly belongs to the class of *long-range interactions*.

The elastic energy of small fluctuations adds a supplementary interaction to the self-avoiding walk, with two possible effects on the universality class:

1. left unchanged, if the elastic energy contains only backbone terms ( $k_{co} = 0$ );
2. changed, if the elastic energy contains also non bonded contact springs ( $k_{co} > 0$ ).

The elastic interaction is taken into account in terms of vibrational spectra, which are global properties, in the sense that they describe collective motions of the system; their long or short range nature is not so easy to be established.

The results presented in the present chapter prove that the universality class is actually affected by the elastic interactions; moreover the already mentioned works on fractons (section 3.5) have shown that the phonon dimension of self avoiding walks depends on the presence or not of the intra-chain bridges, in a similar way as the vibrational spectrum of compact walks depends on the value of  $k_{co}$ .

These observations bring to the conclusion that the elastic interaction acts effectively as a long range interactions, susceptible to modify the universality class of the model.





## Chapter 5

# Other results

In this chapter we present four results giving some more information about the contribution of the vibrational entropy in the physics of polymer, but which deserve further investigation for obtain a clearer answer.

First (section 5.1) we wonder about the influence of the backbone spring constants on the universal properties of the model and we find that most probably their variation does not induce any change in the universal behavior, and the key parameter is the contacts spring constant  $k_{co}$ . In section 5.2 we show that another critical exponent (the entropic exponent) is affected by the vibrations. In section 5.3 we present the influence of the vibrational entropy on the building of secondary structure. Finally in section 5.4 preliminary results indicate that an interesting phenomenology has to be expected when the inclusion of vibrations is extended to the problem of adsorption.

### 5.1 Different spring constants

The results of the previous chapter were calculated on the model in 2 and 3 dimensions with the specific choice for the backbone spring constants reported in Table 4.2 and Table 4.1. That choice is motivated by the similarity with off lattice potentials which are tuned for reproduce at best the experimental results.

We wonder if different choices may bring different results, in particular concerning the universality class to which the model belongs. This question has not been answered exactly, but we present some hints indicating that *only two universality classes exist: the ISAW, corresponding to  $k_{co} = 0$ , and the VIBR, corresponding to  $k_{co} > 0$ .*

#### 5.1.1 The case $k_{co} = 0$

The most interesting fact is that if the contact spring constant is 0 and the ratio  $k_{\text{flat}}/k_{\text{right}}$  (in 2 dims) or  $k_{\text{trans}}/k_{\text{gauche}}$  (in 3 dims) is fixed, the relative

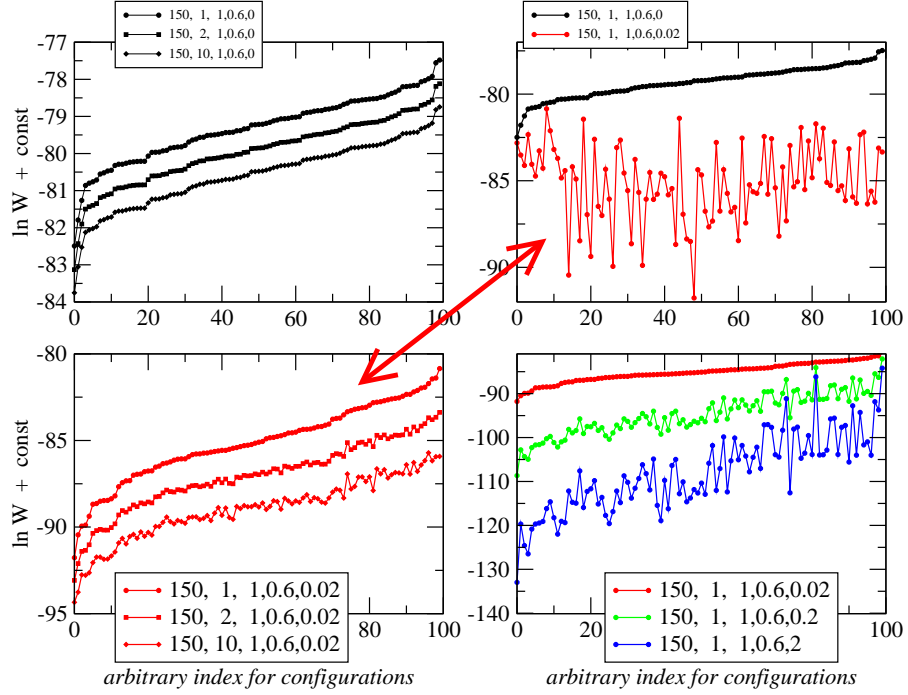


Figure 5.1: **Vibrational entropy and spring constants.** The vibrational entropy of hundred random configurations (3 dimensions, length 50 monomers) are calculated with several combination of the spring constants, reported in the legend. The curves are translated of arbitrary constants for facilitating the comparison. In the **top panels** the configurations are ordered for increasing entropy calculated with  $k_{BB} = 150$ ,  $k_{be} = 1$ ,  $k_{tr} = 1$ ,  $k_{ga} = 0.6$  and  $k_{co} = 0$ . In the **bottom panels** the configurations are ordered for increasing entropy calculated with  $k_{BB} = 150$ ,  $k_{be} = 1$ ,  $k_{tr} = 1$ ,  $k_{ga} = 0.6$  and  $k_{co} = 0.02$ . In other words the red series of circles linked by an arrow are the same values in the two different arrangements.

entropic weight  $W(\gamma)/W(\gamma')$  between any two configurations is constant. In other words, varying any of the  $k_{BB}$  or  $k_{\text{angle}}$ , the vibrational entropy  $\ln W$  of all the configurations varies of the same quantity. This is illustrated in the top left panel of Figure 5.1 where the vibrational entropy of hundred random configurations in 3 dimensions is calculated for different values of the constant  $k_{be}$ : each curve is the series of values obtained for one  $k_{be}$  and the different curves differ of a constant.

When the contacts spring constant is switched on, the relative entropic weights are completely shaken (same Figure, top right panel); moreover, even at fixed  $k_{co}$  if the angular constant changes, the relative weights are not conserved (same Figure, bottom left panel): in this case the different curves can not be superposed through a vertical translation. Finally, keeping fixed the backbone constants and increasing  $k_{co}$  ( $k_{co} = 0.02, 0.2, 2$  bottom right panel) the order of weights is not conserved.

In conclusion the model with  $k_{co} = 0$  presents a peculiar characteristics: the vibrational weight is strictly independent on the backbone spring constants, provided the ratio between the angular constants is kept fixed. More formally, the ratio between the weights of two distinct configurations is a function of the ratio  $k_{tr}/k_{ga}$  only:

$$\frac{W(\gamma, k_{BB}, k_{be}, k_{tr}, k_{ga})}{W(\gamma', k_{BB}, k_{be}, k_{tr}, k_{ga})} = r \left( \frac{k_{tr}}{k_{ga}} \right) \quad (5.1)$$

An identical phenomenology is observed in 2 dimensions as well, where the dependence is on the ratio  $k_{\text{flat}}/k_{\text{right}}$ .

In the reference set of spring constants the dihedral or angular spring constants  $k_{tr}/k_{ga}$  or  $k_{\text{flat}}/k_{\text{right}}$  are chosen in such a way that the Hessian matrix has the same eigenvalue for the *trans* (flat) and the *gauche* (right) angles. A variation of this ratio brings a variation of the vibrational entropy which is not a simple addition of a constant. Nevertheless, simulations of the 2 dimensional walk with  $k_{\text{flat}}/k_{\text{right}} = 1$  gives a cross-over exponent virtually identical to the reference case ( $k_{\text{flat}}/k_{\text{right}} = 3/2$ ), as illustrated in Figure 5.2: the two cases correspond to the yellow and black series of points respectively and they superpose almost exactly.

### 5.1.2 The case $k_{co} > 0$

When  $k_{co} > 0$  the vibrational entropy depends non trivially on the backbone spring constants: the bottom panels of Figure 5.1 illustrate how the weights change on variations of the angular or the contact spring constants. The non trivial behavior corresponds to a variation which is not the simple addition of a constants, which would turn out in identical relative weights and an identical model.

In 2 dimensions, Figure 5.2, the scaling of the relative radius derivative with  $N$  is almost similar in all the cases in consideration: the reference case

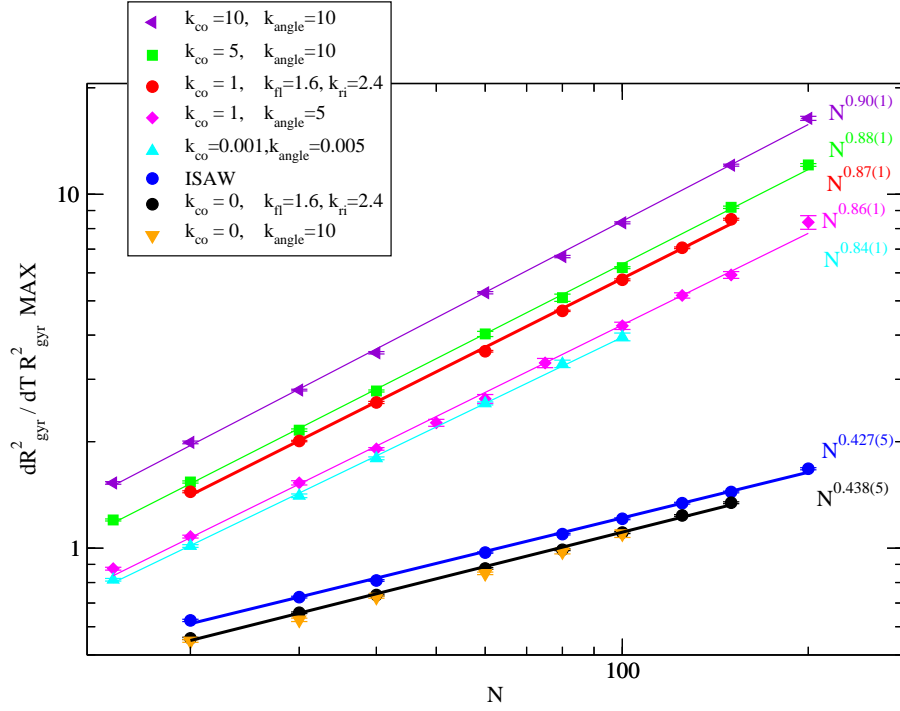


Figure 5.2: **Cross-over exponent at varying spring constants.** The maximum of the relative radius derivative VS  $N$ , two dimensions. The cross-over exponent of the different curves is reported in the graph and obtained from least square method: the straight lines between the point represent the fit. The backbone spring constant is  $k_{BB} = 10$  in all cases: the red points refer to the reference case Table 4.2.

(red points) and other cases with  $k_{\text{flat}}/k_{\text{right}} = 1$  scale in a very similar way, the cross-over exponents varying between 0.84 and 0.90; we are tempted to interpret this result as a hint of the existence of a unique universality class, in particular because of the net distinction from the ISAW universality class, even when the constant spring constant is very small ( $k_{co} = 0.001$ , cyan points). In this scenario, the precise value of the cross-over exponent should be found from simulations at very high  $N$  and would be independent on the choice of spring constants.

An alternative scenario, consisting of a continuous spectrum of cross-over exponents may also exist, but we are not able to discriminate now between the two.

In 3 dimensions, instead, the unique universality class of  $k_{co} > 0$  is the first order phase transition ( $\phi = 1$ ). In this case the only alternative scenario is that a critical very small constant spring constant exists, such that models with  $k_{co} < k_{co}^{\text{crit}}$  belong to the ISAW class, or even a third class. The simulations of very small  $k_{co}$  and the existence of a cross-over between the maxima of specific heats at different constants (Figure 4.2) seems to support the first case.

## 5.2 Entropic exponent in 2 dimensions

The entropic exponent is the power law divergence of the partition function:

$$Z(N) \simeq A\mu^N N^{\gamma-1}. \quad (5.2)$$

The canonical partition function depends on the temperature and increases with the number of monomers  $N$ . The parameter  $\mu$  depends on the lattice, the temperature and the nature of the walk (random, self avoiding,...). Equation (5.2) is the generalization of the simplest case, the number of random walk on a lattice with coordination number equal to  $\mu$ . On the square lattice each site has 4 neighbors, so the number of random walks with  $N$  steps is

$$Z_{RW}(N) = 4^N.$$

If the walk is self avoiding the possibilities are much less, because many directions are forbidden by the self avoidance constraint. Moreover the number of self avoiding walk does not follow a simple exponential law, but corrections are introduced: the first one is a power law characterized by the entropic exponent  $\gamma$ . Contrary to the connective constant, the exponent  $\gamma$  is a universal parameter which is expected to take on three values, one in the high temperature regime, one at the  $\Theta$  point and the third in the compact regime. Similarly to other critical exponents, it was evaluated in 2 dimensions through the mapping on magnetic systems [72], but the numerical confirmation, in particular at the  $\Theta$  point and the compact phase, is still challenging [57, 34, 9].

In this section we consider exact enumerations of ISAW and VIBR in 2 dimensions and we compare the entropic exponent at the infinite temperature and at the critical temperature with or without the contribution from the vibrational fluctuations. The analysis reveals that the meaningful quantity to compare is the difference between the two values  $\gamma(T = \infty) - \gamma(T_\Theta)$  and gives the occasion for better specify the definition of the entropic weight. We concentrate on a model with spring constants  $k_{BB} = 10$ ,  $k_{\text{right}} = k_{\text{flat}} = 5$  and  $k_{co} = 1$ .

### 5.2.1 The connective constant $\mu < 1$

At infinite temperature  $Z_\infty(N) = \sum 1$  for ISAW and  $Z_\infty(N) = \sum W(\gamma)$  for VIBR. At the critical temperature  $Z_\Theta(N) = \sum e^{\beta_c n_{co}} > Z_\infty(N)$ ; and similarly for VIBR. Looking at the exact values for the partition functions, and comparing with equation (5.2), in the logarithm form

$$\ln Z(N) \sim N \ln \mu + (\gamma - 1) \ln N + \text{const} + \dots \quad (5.3)$$

we find  $\mu < 1$  for the VIBR model (Figure 5.3).

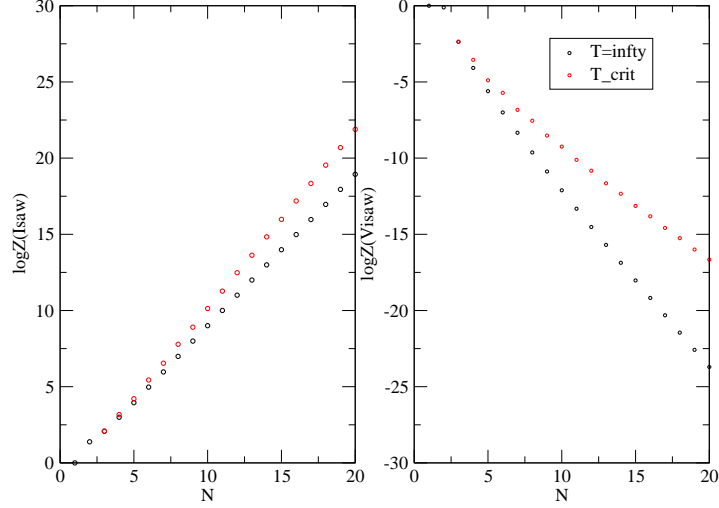


Figure 5.3: **Partition function with respect to the size in 2 dimensions.** The logarithm of the partition function  $\ln Z(N)$  is calculated at the temperatures  $\infty$  and critical (the maximum of the relative radius derivative) from exact enumerations of walks of length  $10 \leq N \leq 20$ , spring constants  $k_{BB} = 10$ ,  $k_{\text{flat}} = k_{\text{right}} = 5$ ,  $k_{co} = 1$ .

### 5.2.2 The exponent $\gamma < 0$

A basic, and rough, method for finding the parameters is a non-linear fitting with the formula 5.3 (without "..."); we also ignore data such that  $N < 10$  because first small  $N$  are particularly *out of fit*. It results:

	$\ln \mu$	$\gamma$	$const$	$\ln \mu(\text{"theor"})$	$\gamma(\text{theor})$
$T = \infty$	0.9717	1.312	-1.43	0.97008	1.34375
$T_{crit}$	1.1845	1.118	-1.45	1.17062	1.14286
VISAW					
$T = \infty$	-1.0216	-0.987	2.681		
$T_{crit}$	-0.5707	-1.452	2.087		

### 5.2.3 Normalization of data

The values of  $\mu$  and  $\gamma$  can be changed by *normalization* of the weights, that is multiplying the partition function by an  $N$ -dependent factor. This operation leaves invariant the thermodynamics (observables) and the critical exponents  $\nu$  and  $\phi$ , that is all the results shown in the previous chapter. The operation of *normalization* means to redefine the entropic weight equation (2.12) in relation with a reference weight. Two possibilities are for example:

1.  $Z' = \frac{\sum W(\gamma)e^{\beta n_{co}}}{W_N(\text{---})}$  where  $W_N(\text{---})$  is the weight of the straight configuration of  $N$  steps;
2.  $Z'' = \frac{\sum W(\gamma)e^{\beta n_{co}}}{\langle W \rangle}$  where  $\langle W \rangle = \frac{\sum W(\gamma)}{c_N}$  is the average weight over the  $c_N$  SAWs.

The physical meaning of such operation resides in the fact that only relative weights count for the thermodynamic of the system. So in the case (1.) the weight is redefined as  $W'(\gamma) = \frac{W(\gamma)}{W(\text{---})}$ , relative to the straight configuration; as the choice of the straight configuration as the reference weight is completely arbitrary, a better choice could be the average weight (case (2.)). In the case (2.) it is also interesting that the partition function at  $T = \infty$  reduces to the one of ISAW, by construction, so the entropic exponent at infinite temperature of the two models becomes identical. It results:

	$\gamma$	$\gamma_{(1.)}$	$\gamma_{(2.)}$	$\gamma_{ISAW}$	$\gamma_{theor}$
$T = \infty$	-0.987	1.528	1.312	1.312	1.34375
$T_c = 0.69$	-1.452	1.063	0.847	1.118 ( $T_c = 1.5$ )	1.14286

In general the partition function can be corrected with an  $N$ -dependent factor  $F_N$  whose behaviour can be described as:

$$F_N = B\mu^{*N}N^{\gamma^*-1} \quad (5.4)$$

analogous to equation (5.2). The corrected partition function results

$$Z' = \left( \frac{\mu}{\mu^*} \right)^N N^{\gamma - \gamma^*} \quad (5.5)$$

and the signs of  $\mu/\mu^*$  or  $\gamma - \gamma^*$  can be varied arbitrarily.

### 5.2.4 Conclusion

The arbitrariness in defining the entropic weight up to a constant depending on  $N$  (but of course not on the configurations of the same length) allows to adjust the value of  $\gamma$  at one temperature. Our choice is to set  $\gamma_{ISAW}(T = \infty) = \gamma_{VIBR}(T = \infty)$ . At this point it is meaningful to consider the difference  $\gamma(T = \infty) - \gamma(T = \Theta)$  and the previous tables show that this quantity calculated in the VIBR model is bigger than in ISAW.

Notice that for ISAW  $\gamma$  is expected to assume only 3 values: at  $T > T_c$ , at  $T = T_c$  and at  $T < T_c$ . If the same is valid for VISAW (which is an open question!) the equation (5.4) should not be  $T$ -dependent in the exponent  $\gamma^*$  (while in principle it could be in  $\mu$ ). In this way the difference  $\gamma(\infty) - \gamma(T_c)$  is conserved.

This result is in line with the conclusion that the entropy difference between the coil and the globule state is the responsible of the new universality class.

## 5.3 Topological properties: preliminary results

The secondary structure is a constituent part of a protein: it consists basically on special arrangements of monomers in the form of an  $\alpha$ -helix or a  $\beta$ -sheet [29]. These structures are characterized by a high packing of monomers and they are observed at high frequency in the compact state of proteins. Previous studies on lattice models aimed to answer a question which could be re-formulated as: how much obvious is the presence of the secondary structure?

Indeed, in the compact phase the high packing of monomers makes plausible the formation of *regular* structures. The result depends also on the definition of *regular structures*, but complete enumerations of compact polymers on the square lattice [13] and simulation of longer polymers [53] have confirmed that some motif, such as the  $\alpha$ -helix represented in Figure 5.4, are more likely to be found in the compact than in the coil state.

We have studied (Figure 5.4) the influence of the vibrational entropy on the formation of helix: on the 2 dimensional lattice the helix is a sequence of step in an angle configuration: the smaller helix contains 4 monomers (black portion), a longer helix is made by 6 monomers (black+red portion) and so on. All of these structures are enhanced by the vibrations: already



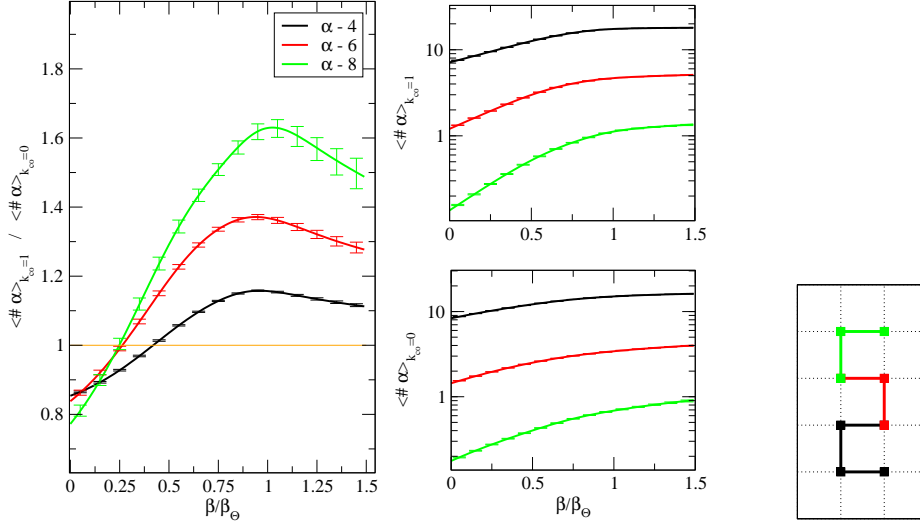


Figure 5.4: **Helices in 2 dimensions.** *Definition:* The helix of length 4 ( $\alpha - 4$ ) is the black portion of the sketch on the right. A helix of length 6 (8) is the black+red (black+red+green) portion. *Graphs:* In the **middle** we report the average number of helices in the case  $k_{co} = 1$  and the case  $k_{co} = 0$ ; on the **left** the ratio between the two. The other spring constants are  $k_{BB} = 10$  and  $k_{right} = k_{flat} = 5$ .

in the coil region (reduced beta  $\frac{\beta}{\beta_{\Theta}} < 1$ ) the ratio is bigger than 1. The comparison is made between the case  $k_{co} = 0$  and  $k_{co} = 1$  in order to show that also in this context the existence of the contact spring constant is crucial. This preliminary result will motivate further research in this direction, for example the extension to secondary structure in 3 dimensions or to different definitions of structure.

The definition of helix on the square lattice as given in Figure 5.4 or the definition of any other structure are intuitive but arbitrary. With the aim of finding an objective characterization of the topology we define the average contact length of a walk as

$$l_{co} = \frac{\sum_{contacts} |i - j|}{n_{co}}. \quad (5.6)$$

The  $l_{co}$  is calculated over the contacts between monomers  $i$  and  $j$ . The minimum contact length is given by the minimum number of steps necessary for the walk to make a ring: in the square lattice it is 3, while in the tetrahedral lattice is 5. The maximum contact length is  $N$ , corresponding to a ring chain (the first and the last monomer are in contact). The definition of this observable was inspired by the fact that helices in both 2 and 3 dimensions are characterized by short distances between contacts, but it does not require any *intuitive similarity* with any structure.

The investigation of the distribution of  $l_{co}$  in 3 dimensions furnishes the interesting result illustrated in Figure 5.5. The configurations are sampled

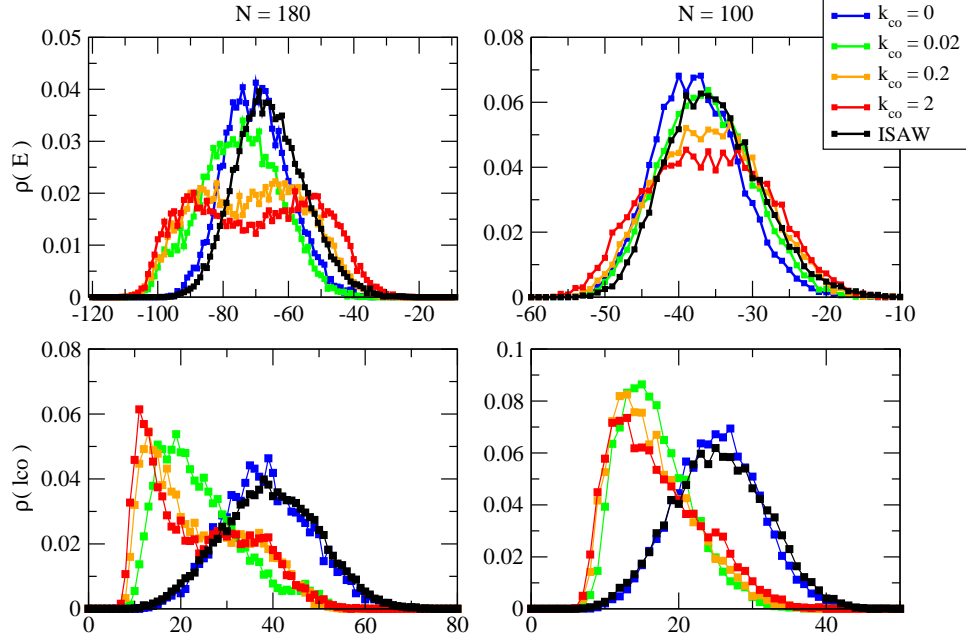


Figure 5.5: **Average contact length distribution** of 3 dimensional chains of length 100 and 180 monomers and at the critical temperature. **Top:** the distribution of the energy; **Bottom:** the distribution of the average contact length. The colors refer to different  $k_{co}$ .

at the critical temperature (the maximum of the heat capacity) at several values of the contact spring constants. The first evidence is that the entropic vibration enhances the localization of contacts (distributions shifted on smaller lengths). The result is coherent with the observed enhancement of helix in 2 dimensions. An ulterior interesting fact is the discernibility of the two universality classes: the average contact length of ISAW and VIBR- $k_{co} = 0$  are almost the same, while VIBR- $k_{co} > 0$  collapses rapidly on a state with prevalently short distance contacts. When the bi-modal distribution of the energy is still not evident, as  $N = 100$  for example, the average contact length behaves in a well distinct way between the two categories. This result may motivate further research for example on the definition of local structure or, given the sensitivity to the universality class, the use of it as an important order parameter.

## 5.4 Chain adsorbed on a wall: preliminary results in 2d

### 5.4.1 The model

A polymer in presence of a surface behaves in a different way than in the bulk of a solution. If monomers are attracted by the surface, the polymer can be adsorbed. In this case, as in the coil-globule transition, the energetically convenient adsorbed state is at the same time entropically unfavorable, because associated to a reduction of the spatial freedom.

What is the general scheme of the adsorption onto a surface? Is it a gradual process or an abrupt transition? The answer to these general questions is searched on lattice models retaining the essential ingredients of the phenomenon: one chain (a self avoiding walk), a surface and an attraction interaction between them. The energy of a walk is proportional to the number of surface-polymer contacts:

$$E_{Ads} = -\varepsilon n_{sc} \quad (5.7)$$

and in the simplest model the solvent is good, so the monomers do not interact between them.

The first discovery is that a transition indeed exists: in the thermodynamic limit, the SAW is expected to collapse from the bulk situation to the adsorbed configuration at a critical temperature. The two phases are distinguished by a different density of contacts:

$$\frac{n_{sc}}{N} \begin{cases} = 0 & \text{if } T > T_c \\ > 0 & \text{if } T < T_c \end{cases} \quad (5.8)$$

Theoretical and numerical studies [42] indicate that the adsorption lattice model undergoes a continuous transition at the critical temperature. As a consequence the same scaling arguments as for the coil-globule case can be developed (section 4.3). In particular the energy at the critical temperature is governed by a cross-over exponent:

$$E_{Ads}(T_c) \propto N^\phi \quad (5.9)$$

As long as  $\phi < 1$  the density at the critical temperature

$$\frac{n_{sc}}{N}(T_c) \propto \frac{E}{N}(T_c) \rightarrow 0, \quad (5.10)$$

and the two phases connect continuously; instead, the unitary cross-over exponent  $\phi = 1$  would denote a finite jump from 0 at  $T < T_c$  to  $\propto N$  at  $T = T_c$ , as it is customary for discontinuous transitions. The value indicated by simulations is

$$\phi_{Ads} = \frac{1}{2} \quad (5.11)$$

in both 2 and 3 dimensions [42].

The metric properties of the walk represent an other discriminant between the two phases. The high temperature phase is characterized by the good solvent metric exponent,  $\nu = \frac{3}{5}$  in 3 dimensions or  $\nu = \frac{3}{4}$  in 2 dimensions, while in the low temperature phase the parallel and perpendicular radius scale differently: in the direction perpendicular to the surface, the polymer does not propagate ( $\nu = 0$ ) and in the parallel direction the polymer is constrained in the  $d-1$  space and take the corresponding good solvent metric exponent; this is 1 (straight line) when the polymer lives in  $d = 2$  and is adsorbed onto a line.

### 5.4.2 Vibrations

The partition function of the adsorbing lattice polymer is

$$Z_{ADSO} = \sum_{\gamma} e^{\beta n_{sc}} \quad (5.12)$$

and corresponds to the 0 order approximation of a complete potential with inherent structures corresponding to the lattice self avoiding walks. A better approximation introduces the vibrations and the adsorbing polymer with vibrations, VIAD, is represented by the following partition function:

$$Z_{VIAD} = \sum_{\gamma} e^{\beta n_{sc}} W(\gamma). \quad (5.13)$$

The entropic weight depends on the conformation of the polymer and the elastic constants associated to the bonds, the angles and the contact-surface interactions. The backbone interactions are identical to the coil-globule model, while the contact-surface term depends on the displacement of the monomers in contacts with the surface. If the surface is fixed (the degrees of freedom of the surface molecules are neglected) the elastic potential of the monomer  $i$  in contact with the surface is

$$V_i^{sc} = \frac{1}{2} k_{sc} \Delta y_i^2. \quad (5.14)$$

depending only on the monomer displacements. The corresponding Hessian submatrix is

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (5.15)$$

### 5.4.3 Results

This model was studied in [73]. Simulations of walks of length up to 100 monomers, on the square lattice and adsorbed to a line were performed at several temperatures with (VIAD) and without (ADSO) vibrations. The

discontinuity specific heat at the critical temperature shows an abrupt discontinuity and the cross over exponent is much bigger than in the standard model

$$\phi_{VIAD} = 1.15 \pm 0.02 \quad (5.16)$$

$$\phi_{ADSO} = 0.46 \pm 0.04; \quad (5.17)$$

the second result is compatible with the predicted  $1/2$ , while the VIAD case is compatible with a first order transition scaling, as in the 3 dimensional coil to globule. Nevertheless the data did not allow to distinguish the coexistence typical of first order transitions: the length of the chains are probably too short.

The critical inverse temperature is significantly shifted up:

$$\beta_c^{VIAD} = 1.22 \pm 0.02 \quad (5.18)$$

$$\beta_c^{ADSO} = 0.477 \pm 0.022. \quad (5.19)$$

This shifting is index of a bigger unbalance between the entropy and enthalpy differences of the free and the adsorbed state; in addition to the evident similarity with the frequency analysis of chapter 3, this show that the modified behavior is attributable to the additional entropic gap added by taking into account the contribution of vibrations.

This preliminary work deserves to be developed both in 2 and 3 dimensions; in 2 dimensions the question of the coexistence between the populations is still open, even if strongly supported by the cross over exponent bigger than 1. In 3 dimensions the phenomenology could be similar, given the similarity of the cross-over exponent in the standard problem ( $1/2$  in both cases), but only a direct investigation can answer.



## Chapter 6

# Viscosity experiments

In this chapter we show that the cross-over exponent defines the formation of the thermal blobs (section 6.1) and that the size of thermal blobs is measurable from intrinsic viscosity experiments (section 6.2). From these ideas we examined data from the literature of viscosity measurements and estimate some experimental cross-over exponent. The results (section 6.3) confirm that an experimental cross-over exponent bigger than the standard ISAW (1/2) is a plausible situation.

### 6.1 Thermal Blobs

At the  $\Theta$  temperature in 3 dimensions, the size  $R$  (Hydrodynamic Radius, Radius of Gyration...) of a polymer behaves as an ideal chain:

$$R = bN^{\nu_{\Theta}} \quad (6.1)$$

At high temperature  $T \gg \Theta$  the polymer swells because of the self avoidance between the monomers:

$$R = bN^{\nu} \quad (6.2)$$

$b$  and  $N$  are the length and the number of Kuhn monomers which describe the size of the polymer as if it was a chain of  $N$  spheres, each of radius  $b$ . The metric scaling exponents are:

	$\nu_{\Theta}$	$\nu$
2d	4/7	0.75
3d	1/2	0.588

Power laws such as equation (6.1) and equation (6.2) show the property of self-similarity. This means that the same law is valid when the measures are taken with different units. For example, if we measure the radius with a unit  $\xi$  much bigger than the monomer size  $b$ , in  $\xi$  there is the number of

monomers  $g = \left(\frac{\xi}{b}\right)^{1/\nu}$  and in  $\xi$  units the polymer measures  $N' = \frac{N}{g}$ . In the new units, the law (6.2) is the same:

$$R = bN^\nu = b\frac{g^\nu}{g^\nu}N^\nu = \xi(N')^\nu. \quad (6.3)$$

in particular the universal exponent  $\nu$  is unchanged. If the unit is long enough, any local detail (stiffness, conformation...) is lost, but the measured scaling is still correct.

At intermediate temperatures,  $T \gtrsim \Theta$ , if the chain is not very long the self-avoidance can not manifest and the chain behaves as ideal, equation (6.1), while if the chain is long enough the SAW statistics is recovered. Changing the length of the chain determines a **crossover** from the ideal to the SAW statistic.

A simple mathematical description of the crossover proceeds as follows. The size  $R$  is a function of the number of monomers  $N$  and the reduced temperature  $t = \frac{T-T_\Theta}{T_\Theta}$  (a measure of the distance from the  $\Theta$  temperatures):

$$R(N, t) = bN^{\nu_\Theta} g(Nt^{1/\phi}) \quad (6.4)$$

At fixed  $t$ , if  $N$  is small the size behaves ideally, while if it is high the size grows as a SAW. This is expressed by the equation:

$$g(Nt^{1/\phi} \ll 1) \propto \text{const} \Rightarrow R \propto N^{\nu_\Theta} \quad (6.5)$$

$$g(Nt^{1/\phi} \gg 1) \propto (Nt^{1/\phi})^{\nu-\nu_\Theta} \Rightarrow R \propto N^\nu \quad (6.6)$$

From these expressions the relation between the critical region and the temperature can be read:

$$\text{ideal: } N \ll t^{-1/\phi} \quad \text{SAW: } N \gg t^{-1/\phi}$$

Even if the chain is enough long to behave as a self avoiding chain, a small section of the chain covering  $N_T \sim t^{-1/\phi}$  monomers behaves ideally:

$$R_T = bN_T^{\nu_\Theta} \quad (6.7)$$

The **thermal blob** is biggest subsection of the chain which behaves ideally. The existence of the thermal blob destroys the relation of self-similarity, equation (6.3). Nevertheless, choosing a measure unit bigger than the thermal blob the law is valid:  $R = R_T N'^\nu$  with  $N' = \frac{N}{N_T}$ . That is:

$$R = bN_T^{\nu_\Theta} \left( \frac{N}{N_T} \right)^\nu \quad (6.8)$$

Comparing with the cross-over relation, equation (6.4), in the SAW region (assumption equation 6.6) we can write:

$$R = b(t^{-1/\phi})^{\nu_\Theta} \left( \frac{N}{t^{-1/\phi}} \right)^\nu \quad (6.9)$$



and we conclude that the thermal blob size is a simple power law of the reduced temperature:

$$N_T \propto t^{-1/\phi}. \quad (6.10)$$

From Flory arguments (see section 6.1.1) the cross over exponent  $\phi$  is

$$\phi = \frac{1}{2} \quad (6.11)$$

The literature reports investigations of the thermal blob size as a function of the solvent quality, but the relation (6.11) has never been verified. If the thermal blob size is calculated independently from the assumption (6.11) than the effective value of  $\phi$  can be calculated.

### 6.1.1 Flory Argument for $\phi = \frac{1}{2}$

A thermal blob contains so many monomers as they don't overlap each other: the mean field probability of overlap is  $N_T v \cdot \frac{N_T}{R_T^3}$ . Allowing 1 overlap per thermal blob we obtain

$$R_T \sim \frac{b^4}{v} \quad (6.12)$$

The excluded volume can be approximated with  $v \sim t \cdot b^3$  from which, as  $R_T = bN_T^{\nu_\Theta}$  we conclude

$$\phi = \nu_\Theta = \frac{1}{2} \quad (6.13)$$

### 6.1.2 Our result

Our lattice model with local vibrations gives a clear result at the  $\Theta$  temperature: the crossover exponent  $\phi$  is higher than the exponent predicted by the standard theory:

	ISAW	VIBR
2d	3/7	$0.87 \pm 0.01$
3d	1/2	1

## 6.2 Intrinsic Viscosity

The intrinsic viscosity of a solution is:

$$[\eta] = \lim_{c \rightarrow 0} \frac{1}{c} \frac{\eta - \eta_s}{\eta_s} \quad (6.14)$$

where  $c = \frac{M(\text{solute})}{V(\text{solution})}$  is the mass concentration,  $\eta$  and  $\eta_s$  are the viscosity of the solution and the solvent (alone, which means at  $c = 0$ ) respectively.

The Einstein formula

$$\eta = \eta_s \left(1 + \frac{5}{2}\phi\right) \quad (6.15)$$

relates the viscosity of a suspension with the volume fraction  $\phi = \frac{v(\text{pervaded})}{V(\text{solution})}$  of suspended impermeable particles.

A polymer solution can be represented as a suspension of molecules whose volume is not simply equal to the volume occupied by its monomers,  $v(\text{occupied}) = Nb^3$ , but it is proportional to the size of the whole molecule  $R^3 \propto N^{3\nu}$ . In other words the molecule is assumed to drag the solvent inside its pervaded volume with it. From equations (6.15) and (6.14) the intrinsic viscosity can be expressed as a function of the linear dimension of the molecule. Introducing the solute density  $\rho_s = \frac{M(\text{solute})}{v(\text{occupied})}$ :

$$[\eta] = \frac{5}{2} \frac{\phi}{c} = \frac{5}{2c} \frac{c}{\rho_s} \frac{v(\text{pervaded})}{v(\text{occupied})} \quad (6.16)$$

$$= \frac{5}{2\rho_s} \frac{R^3}{Nb^3} \quad (6.17)$$

where  $R$  is the radius of the molecule and  $b$  the radius of one monomer. The obtained relation (6.17) makes possible to introduce cross over arguments in the viscosity. In particular the radius of the polymer scales with  $N^{1/2}$  as in equation (6.7) in ideal conditions (at  $\Theta$  temperature) or within a thermal blob, while it scales as in equation (6.8) above the critical temperature.

The intrinsic viscosity is also empirically described as a function of the molar mass by the Mark-Houwink-Sakurada (MHS) equation:

$$[\eta] = K_\eta M^a \quad (6.18)$$

The MHS exponent

$$a = 3\nu - 1 \quad (6.19)$$

depends on the quality of the solvent and should take values between 0.5 (in a  $\Theta$  solvent,  $\nu = \nu_\Theta$ ) and 0.76 (in a good solvent,  $\nu = 0.588$ ). The constant  $K_\eta$  depends both on the solvent quality and the characteristics of the monomers, but not on the length of the polymer. Comparing with equation (6.17) in ideal conditions ( $R^3 = b^3 N^{3/2}$ ) we obtain:

$$K_\Theta = \frac{5}{2\rho_s M_b^{1/2}} \quad (6.20)$$

where the mass of a single monomer  $M_b$  appears. In other conditions the same equation (6.17) with the polymer size equal to  $R^3 = (bN_T^{1/2-\nu})N^\nu$  gives:

$$K_\eta = \frac{5N_T^{3(1/2-\nu)}}{2\rho_s M_b^{3\nu-1}} \quad (6.21)$$

Putting together (6.20) and (6.21) one gets the number of monomer within

one thermal blob:

$$N_T = \left(\frac{2.5}{\rho_s}\right)^{-2} (K_\Theta^{-2a} K_\eta)^{1/(0.5-a)} \quad (6.22)$$

$$= \left(\frac{2.5}{\rho_s}\right)^{-2} ([\eta]_\Theta^{-2a} [\eta])^{1/(0.5-a)} \quad (6.23)$$

$$= \frac{1}{M_b} \left(\frac{K_\Theta}{K_\eta}\right)^{1/(a-0.5)} \quad (6.24)$$

$$= \frac{1}{M_b} \left(\frac{[\eta]_\Theta}{[\eta]}\right)^{1/(a-0.5)} M \quad (6.25)$$

The four equations (6.22-6.25) corresponds to equations (16), (17), (4) and (20) of reference [33]. They can be used indifferently depending on the available data.

The MHS exponent  $a = 3\nu - 1$  in equation (6.18) is of phenomenological nature: theoretically (for polymers of infinite size) one should measure either the ideal ( $a = \frac{1}{2}$ ) or the good solvent ( $a = 0.76$ ) value, but in real conditions any intermediate value is found. *In the argument presented by [33] even at intermediate states ( $1/2 < a < 0.76$ ) the chain size scaling exponent is considered to cross over from the thermal blob exponent ( $\nu = 1/2$ ) to the effective exponent  $\nu = (a + 1)/3$ . And not to cross over directly to the good solvent exponent  $\nu = 0.588$ .*

In [33] it is shown that the thermal blob size is a unique function of the MHS exponent: as a consequence *the MHS exponent can be determined from intrinsic viscosity measured **at single molecular mass***. By means of equation (6.22) the viscosity data of several polymer-solvent systems with known solvent quality (exponent  $a$ ) were treated and the unique relation between  $N_T$  and  $a$  is found as the function which fit at best the data of all the systems:

$$N_T = e^{1.13a(a-0.5)^{-2/3}} \quad (6.26)$$

This relation is illustrated in figure 6.1a.

## 6.3 Results

### 6.3.1 Poly(*p*-methylstyrene): pPMS

The MHS exponent  $a$  is an index of the quality of the solvent. At a fixed temperature different solvents correspond to different exponents (qualities), depending on the polymer. When the temperature increases the quality of the solvent changes. Usually the quality is good at high temperature and bad at low temperature (there are exception to that, such as cold denaturation). In order to compare the size of a thermal blob with the temperature we need to measure the viscosity at several temperatures. One of the system

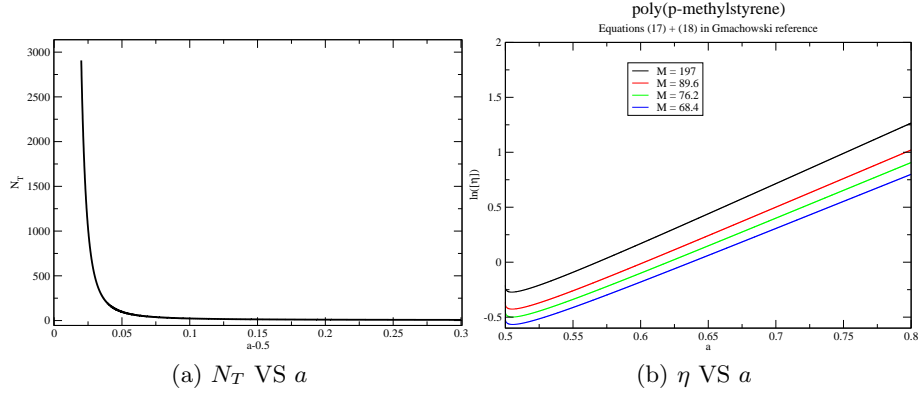


Figure 6.1: **Intrinsic Viscosity and thermal blob size.** **Left:** relation (6.26) between thermal blob size  $N_T$  and MHS exponent  $a$ . **Right:** relation (6.27) between intrinsic viscosity of poly(*p*-methylstyrene) of 4 masses (expressed in  $10^{-4} \frac{g}{mol}$ ) in diethyl succinate  $[\eta]$  and MHS exponent  $a$ .

treated in [33], poly(*p*-methylstyrene) in diethyl-succinate (data taken from [65]), is studied while varying the temperature. In particular by substituting equation 6.26 into equation 6.23 and taking the logarithm:

$$\ln[\eta] = (1 - 2a) \ln \frac{2.5}{\rho_s} + 2a \ln[\eta]_{\Theta} - 1.13a(a - 0.5)^{1/3}, \quad (6.27)$$

we find a direct relation between the intrinsic viscosity and the MHS exponent, with no need of taking polymers at different sizes and comparing with the scaling relation (6.18). The relation (6.27) is illustrated in figure 6.1b. The density of poly(*p*-methylstyrene) (pPMS) is  $\rho_s = 102 \frac{g}{dl}$  [18] and the curves refer to the four different lengths of pPMS reported in [65]. Note that  $\rho_s$  and  $[\eta]$  have to be expressed in the same unit (if  $\rho_s$  is expressed in  $\frac{g}{dl}$  than the viscosity has to be expressed in  $\frac{dl}{g}$ ). From the plot of equation 6.27 we numerically find the MHS exponent corresponding to each temperature. The results are in Table 6.1. The MHS exponent is almost the same in the four sets of data, validating the hypothesis of being a function of only the solvent quality. From equation (6.26) we calculate the size of the thermal blob (in number of monomers) using the average MHS exponent (last column of Table 6.1). The thermal blob size is compared with the relation  $N_T \sim t^{-1/\phi}$  and by least square method we obtain

$$\phi = 1.21 \pm 0.04 \quad (6.28)$$

(see Figure 6.2a). We can use also equation (6.25) and calculate the thermal blob mass separately as in [33, Fig. 7]. The result is reported in Figure 6.2b. The good agreement with a power law validates the cross-over relation equation 6.10, or:

$$t \sim M_T^{-\phi}.$$

t	$M \times 10^{-4}$				
	197	89.6	76.2	68.4	
16.4	0.500	0.500	0.500	0.500	0.500(0)
18	0.535	0.536	0.536	0.536	0.536(1)
20	0.545	0.544	0.545	0.544	0.544(1)
25	0.562	0.561	0.558	0.556	0.559(3)
30	0.575	0.573	0.572	0.570	0.572(2)
40	0.596	0.591	0.587	0.588	0.591(4)
50	0.610	0.603	0.601	0.601	0.604(4)
60	0.622	0.612	0.610	0.614	0.615(5)

Table 6.1: **MHS exponent of pPMS in diethyl succinate** at several temperatures, calculated from equation 6.27. The right column gives the arithmetic average of the 4 results.

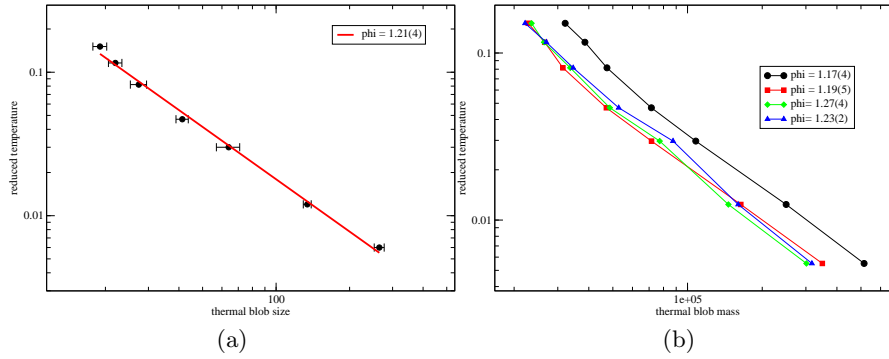


Figure 6.2: **Thermal blob size of Poly(*p*-methylstyrene) in diethyl succinate** as a function of the reduced temperature. **Left:** from equation (6.26) and the exponent in Table 6.1. Error bars are from the errors on  $a$ , last column of Table 6.1. **Right:** from equation (6.25). The four masses are the same as in Figure 6.1b:  $M = 197, 89.6, 76.2, 68.4 \cdot 10^4$ .

From these data the cross over exponent results:

$$\phi \in (1.17, 1.27). \quad (6.29)$$

### 6.3.2 Poly(methyl methacrylate) in acetonitrile

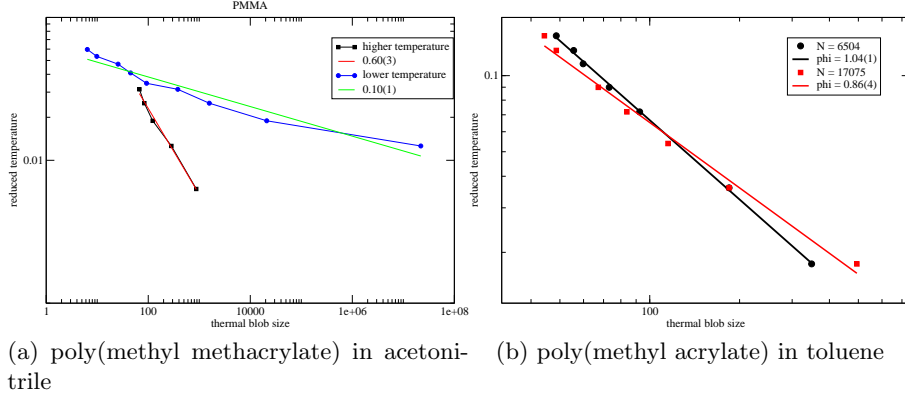


Figure 6.3: **pMMA and pMA: reduced temperature and thermal blob mass.** **Left:** mass  $M = 3.15 \cdot 10^6$ . **Right:** mass  $M = 56 \cdot 10^4$  or  $M = 147 \cdot 10^4$ .

In [22, Table 1] data of viscosity of **poly(methyl methacrylate) in acetonitrile** above and under  $\Theta$  condition are reported. When more than one data is reported at a same temperature we calculate the arithmetic average of them. The result is illustrated in Figure 6.3a.

### 6.3.3 Poly(methyl acrylate) in toluene: pMA

The data are from [41, Fig. 2]. The resulting reduced temperature and thermal blob mass are illustrated in Figure 6.3b.

### 6.3.4 Poly(2-cinnamoyloxyethyl methacrylate): pMMA

This polymer is studied in [40]. In particular its intrinsic viscosity is measured at several lengths (160, 190, 320, 330, 650, 1180 monomers) and temperature (at and above the  $\Theta$  temperature estimated to be at  $60.0^\circ\text{C}$ ). The measured MHS exponent  $a$  and constant  $K_\eta$  are reported ([40, Table 3]). We calculate the size of the thermal blob from equation (6.26) and we represent it as a function of the temperature in Figure 6.4. Assuming for the dependence between reduced temperature and thermal blob size the power law  $t \sim N_T^{-\phi}$  we estimate the cross-over exponent (see the legend). We are in doubt about considering the lowest  $t$  point which looks very far from the

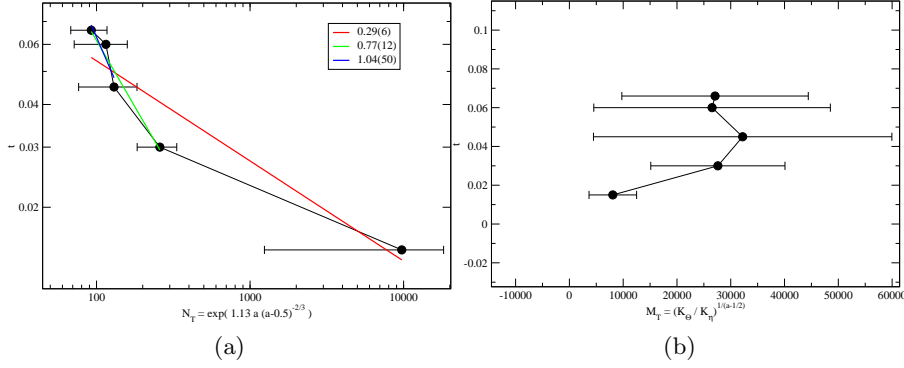


Figure 6.4: **Thermal blob size of Poly(2-cinnamoyloxyethyl methacrylate) as a function of the reduced temperature.** The empiric relation (6.26) is used in 6.4a and the cross-over relation (6.24) is used in 6.4b.

other. One good reason to exclude it is that the estimated thermal blob size is much bigger than the size of the polymers used for the estimation. The results obtained keeping into account only the first 3 or 4 data are both compatible with a model in which the cross-over exponent is bigger than 0.5.

On the right figure 6.4b the thermal blob size is estimated relying directly on the scaling arguments, equation (6.24). The strange result is that the two lowest temperatures seem to give a thermal blob size growing with the temperature, while we should expect the opposite.

The error bars are based on the error bars over  $K$  (Figure 6.4b) or  $a$  (Figure 6.4a).

## 6.4 Conclusions

We report in table 6.2 the cross-over exponents obtained following the procedure described in the previous sections. We exclude from this summary the results of section 6.3.4 because of their clear imprecision.

The obtained values, ranging between 0.60 and 1.20, indicate that a transition more cooperative than predicted by the standard model is plausible. A definitive proof should be searched in calorimetry experiments of the transition, but we are not aware of their existence. The cross-over exponent is 1 in a first order transition and is smaller in a second order transition. The exact value refers to the thermodynamic limit, but finite size corrections are unpredictable. The analysis of the VIBR model presents values ranging between 1/2 (ISAW) and values bigger than 1, depending on both the finite

Table 6.2: **Results of viscosity experiments.**

$\phi$	polymer length (monomers)	polymer name
1.23(2)	$5.8 \times 10^3$	pPMS
1.27(4)	$6.5 \times 10^3$	
1.19(5)	$7.6 \times 10^3$	
1.17(4)	$16.7 \times 10^3$	
1.04(1)	$6.5 \times 10^3$	pMA
0.86(4)	$17.1 \times 10^3$	
0.60(3)	$31.5 \times 10^3$	pMMA

size and the value of the contact spring constant (section 4.3). Similarly, the variability of the experimental results can be attributed to the variability of the nature of polymers, which may realize in different contact spring constants, and the variability of their lengths.

On the other hand, the formation of the thermal blob concerns the standard model and its formation should be justified in the context of the model with vibrations. In the standard model the size of the thermal blob increases approaching the  $\Theta$  temperature where it diverges and polymers are expected to behave ideally in the thermodynamic limit. In the model with vibrations we state that the ideal behavior is a finite size effect, while in the thermodynamic limit one should expect a discontinuous jump from the coil to globule state. Referring to figure 4.5, we observe that a region of ideal behavior is indeed present at finite size: at inverse temperature  $\beta = 3.20$  it is particularly pronounced (from length 50 to 100). In models with smaller contact spring constant, where the bimodal distribution is observed at much higher lengths, the ideal behavior extends longer. In these cases, the ideal behavior presented by the metric exponent comes with a non standard value of the the cross-over exponent (figure 4.2).

The finite size limit of the VIBR model suits to the interpretation of the viscosity experiments, presenting at the same time the ideal behavior at the  $\Theta$  temperature and the cross-over exponent bigger than  $1/2$ . In other words, the ideal behavior (and the formation of thermal blobs) is preliminary to the discontinuous jump at the thermodynamic limit.



## Chapter 7

# Conclusions

In this Thesis we have improved a well known model consisting of the Self Avoiding Walk with attractive Interaction between segments on a square lattice in 2 dimensions and a tetrahedral lattice in 3 dimensions. The model is sensed to reproduce the phenomenology of the coil to globule transition of long molecules. This is an important phenomenon occurring for example in solutions of homo-polymers and in proteins; in these one, in particular, the phenomenon looks like a binary response to the quality of a solvent: either a protein is in its unique native state, or it is denatured. The process is often reversible and the accuracy with which proteins fold can have important effects on the health of an organism.

Such characteristics motivate the study of the phenomenon from many points of view. The approach of the statistical physics consists into find the simplest model on which the transition of a long chain from the coil state (denatured, in the language of protein physics) to the globule state (native protein) can be investigated. The ISAW on the lattice retains the basic ingredients of the phenomenon: the lattice approximation reproduces the disposition of the components of the molecule, as allowed by the strong covalent bonds of the backbone or the long range interactions between monomers which are distant along the chain, but close in space; the attraction between nearest neighbor monomers establishes a phase transition belonging to the universality class of the  $\Theta$  point.

Reconsidering the long jump from the physics of real molecules to the definition of the model, we may interpret the standard ISAW as the result of an approximation of order zero of the global molecular potential around its inherent structures, the walks on the lattice. A natural improvement of the model considers the next term in the approximation, which is represented by the Elastic Network built around the Self Avoiding Walk. In other words, the walk on the lattice is the configuration at the minimum of the potential, while the elastic network is the approximation of the potential with a quadratic form.

In more general terms, the improvement of the approximation consists in a more precise description of the configurational entropy of the molecule: as the refinement comes from harmonic vibrations around the inherent configurations, we described the characteristics of the spectrum of vibration of the walks, with particular attention to the coil and the compact structures, as they play the role of the two opposing phases in the transition.

Afterwards the thermodynamics of the transition was investigated by means of numerical simulations and scaling arguments. The global scenario states that the vibrational entropy induces a behavior belonging to a universality class different from the standard  $\Theta$  point. The result can be explained in a similar way in 2 and 3 dimensions, sign of the existence of a general characteristics of the vibrational entropy. More in details, the elastic constant associated to the intrachain contacts is individuated as the parameter responsible of switching on the new universality class when  $k_{co} > 0$ . The result in 3 dimensions in particular shows that an all or none transition akin to the protein folding can be represented by a general model like the self avoiding walk without recurring to *ad-hoc* modifications of the interaction energies.

In the final parts of the work we have presented some preliminary results which could inspire further development in the study of the secondary structure formation, an other important aspect of the protein physics, or in the different phase transition presented by the polymer adsorbed onto a surface. We also illustrated the re-elaboration of experimental measures of intrinsic viscosity and the results show that a model with a degree of sharpness of the coil to globule transition higher than the standard model is plausible.

# Appendix A

## Hessian matrix of 2d walk

We describe here the step by step construction of the Hessian matrix of a Self Avoiding Walk on the square lattice. The SAW has  $N$  monomers in 2 dimensions, so the Hessian is a  $2N \times 2N$  matrix. The matrix is built by adding a sub-matrix corresponding to each of the interactions:

1. bond between sites  $i$  and  $i + 1$ : one of the  $4 \times 4$  matrices of table A.1 is added at element  $(2i, 2i)$ ;
2. angle between sites  $i, i + 1, i + 2$ : if the angle is right, one of the  $6 \times 6$  matrices of table A.2 is added at element  $(2i, 2i)$ ; if the angle is straight, one of the matrices in table A.3 has to be chosen;
3. contact between sites  $i$  and  $j$ : one of the matrices of table A.1 is *spanned* on the elements  $(2i, 2i + 1, 2j, 2j + 1) \times (2i, 2i + 1, 2j, 2j + 1)$ .

### A.1 bond

Consider the bond between sites 1 and 2, along the x axes:

$$\vec{R}_1 = \vec{r}_2 - \vec{r}_1 = \begin{pmatrix} l + \Delta x_2 - \Delta x_1 \\ \Delta y_2 - \Delta y_1 \end{pmatrix}$$

The bond potential (equation 2.29) can be approximated for small displacements ( $\frac{\Delta x_i}{l} \ll 1$ ) with:

$$\begin{aligned} \frac{1}{2}k_{BB} \left( |\vec{R}_1| - l \right)^2 &= \frac{1}{2}k_{BB} \left[ l \left( \sqrt{1 + 2\frac{\Delta x_2 - \Delta x_1}{l}} + o\left(\frac{1}{l}\right) - 1 \right) \right]^2 \\ &\sim \frac{1}{2}k_{BB} (\Delta x_2 - \Delta x_1)^2 \end{aligned} \quad (\text{A.1})$$

The expression (A.1) is a quadratic form in the variables  $\Delta \vec{r} = (\Delta x_1, \Delta y_1, \Delta x_2, \Delta y_2)$  and can be written in the matrix form

$$\frac{1}{2}k_{BB} \Delta \vec{r}^t H \Delta \vec{r}$$

with  $H$  defined as in table A.1.

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}$$

Table A.1: Hessian submatrix corresponding to a bond along the x (left) and y (right) axes.

## A.2 angle

Consider the bending energy of a right angle, eq. (2.33), formed by the first step along the positive x axes and the second step along the positive y axes:

$$\vec{R}_1 = \begin{pmatrix} l + \Delta x_2 - \Delta x_1 \\ \Delta y_2 - \Delta y_1 \end{pmatrix} \quad \vec{R}_2 = \begin{pmatrix} \Delta x_3 - \Delta x_2 \\ l + \Delta y_3 - \Delta y_2 \end{pmatrix}$$

Again assuming small displacements, the potential can be approximated by the quadratic form:

$$\frac{1}{2}k_{right} \left( \frac{\vec{R}_1 \cdot \vec{R}_2}{|\vec{R}_1||\vec{R}_2|} \right)^2 \sim \frac{1}{2}k_{right} ((\Delta x_3 - \Delta x_2) + (\Delta y_2 - \Delta y_1))^2$$

The corresponding matrix is the top left of table A.2.

Instead when two steps are in the straight conformation, for example along the x axes:

$$\vec{R}_1 = \begin{pmatrix} l + \Delta x_2 - \Delta x_1 \\ \Delta y_2 - \Delta y_1 \end{pmatrix} \quad \vec{R}_2 = \begin{pmatrix} l + \Delta x_3 - \Delta x_2 \\ \Delta y_3 - \Delta y_2 \end{pmatrix}$$

the energy is:

$$\frac{1}{2}k_{str} \left( 1 - \frac{\vec{R}_1 \cdot \vec{R}_2}{|\vec{R}_1||\vec{R}_2|} \right) \sim \frac{1}{2}k_{str} ((\Delta y_3 - \Delta y_2) - (\Delta y_2 - \Delta y_1))^2$$

This and the similar case sub-matrices are in table A.3

## A.3 contact

The energy of a contact is similar to the bond energy, eq. (A.1), but with a different elastic constant and involving two sites which are *far away* along the chain. If the contact is un the x direction, the interaction is:

$$\frac{1}{2}k_{co} (|\vec{r}_j - \vec{r}_i| - l)^2 \sim \frac{1}{2}k_{co} (\Delta x_j - \Delta x_i)^2$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 & -1 & 0 \\ 0 & -1 & -1 & 1 & 1 & 0 \\ 0 & -1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & -1 & 1 & 1 & -1 & 0 \\ 0 & -1 & 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 & -1 \\ -1 & 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & -1 & 0 & 1 \\ -1 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & -1 & 0 & 1 \end{pmatrix}$$

Table A.2: Hessian sub-matrices corresponding to a right angle: (top left) first step along positive (negative) x, second positive (negative) y; (top right) positive (negative) x, negative (positive) y; (bottom left) positive (negative) y, negative (positive) x; (bottom right) positive (negative) y, positive (negative) x.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 4 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 4 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Table A.3: Hessian sub-matrices of straight angles between steps aligned with x (left) or y (right) axes

The corresponding submatrix is one of table A.1, but it will have to be added at elements  $(2i, 2i + 1, 2j, 2j + 1) \times (2i, 2i + 1, 2j, 2j + 1)$  of the global Hessian.



## Appendix B

# Hessian matrix of 3d walk

A Self Avoiding Walk covering  $N$  sites of a 3 dimensional lattice is associated with a  $3N \times 3N$  Hessian matrix. The Hessian is built step by step by adding the sub-matrices corresponding to each interaction. The sub-matrix contains the derivative of the energy with respect to the coordinates of the involved monomers and depends on the absolute direction of the steps between the monomers.

1. The bond between  $i$  and  $i + 1$  sites contributes with a  $6 \times 6$  matrix to be added at the element  $(3i, 3i)$  of the total matrix. As there are 8 absolute directions in the tetrahedral lattice, there are 8 possible matrices.
2. The fluctuation of the angle between two successive bonds at sites  $(i, i + 1)$  and  $(i + 1, i + 2)$  contributes with a  $9 \times 9$  matrix to be added at the element  $(3i, 3i)$  of the total matrix. The possibilities are 24, they can be calculated considering the possible absolute directions of two bonds.
3. The fluctuation of the dihedral angle formed between three consecutive bonds, from site  $i$  to site  $i + 3$ , contributes with a  $12 \times 12$  matrix. Depending on the absolute direction of the bonds, there are 24 possibilities for the *trans* triplets and 48 possibilities for the *gauche* triplets.
4. The contact between site  $i$  and  $j$  is treated similarly to the bond.

The total of the 104 possible matrices were calculated with MATHEMATICA<sup>TM</sup> following the procedure described in the following.

## B.1 bond

Let's take the bond between sites 1 and 2:

$$\vec{R}_1 = \vec{r}_2 - \vec{r}_1 = \vec{R}_1^0 + \begin{pmatrix} \Delta x_2 - \Delta x_1 \\ \Delta y_2 - \Delta y_1 \\ \Delta z_2 - \Delta z_1 \end{pmatrix}.$$

The elastic energy is

$$V_1^{bb} = \frac{1}{2} k_{bb} (|\vec{R}_1| - l)^2.$$

The Hessian matrix is

$$H_{i,j}^{bb} = \frac{d^2 V_1^{bb}}{da_i da_j}$$

where  $i, j = 1, \dots, 6$  and  $a_1 = \Delta x_1$ ,  $a_2 = \Delta y_1$ , ...  $a_6 = \Delta z_2$ . The eigenvalue of  $H^{bb}$  is

$$\lambda^{bb} = 2. \quad (\text{B.1})$$

## B.2 bend

Let's take the first and second bonds:

$$\vec{R}_1 = \vec{R}_1^0 + \begin{pmatrix} \Delta x_2 - \Delta x_1 \\ \Delta y_2 - \Delta y_1 \\ \Delta z_2 - \Delta z_1 \end{pmatrix}, \quad \vec{R}_2 = \vec{R}_2^0 + \begin{pmatrix} \Delta x_3 - \Delta x_2 \\ \Delta y_3 - \Delta y_2 \\ \Delta z_3 - \Delta z_2 \end{pmatrix}.$$

The bending energy is

$$V_1^{be} = \frac{1}{2} k_{be} (\theta_1 - \theta_1^0)^2 \sim \frac{1}{2} k_{be} \frac{9}{8} (\cos \theta_1 - \cos \theta_1^0)^2 \quad (\text{B.2})$$

where

$$\cos \theta_1 = \frac{\vec{R}_1 \cdot \vec{R}_2}{|\vec{R}_1| |\vec{R}_2|}.$$

The factor  $\frac{9}{8}$  appears because

$$\left. \frac{d^2}{d\theta^2} \frac{1}{2} (\theta_1 - \theta_1^0)^2 \right|_{\theta^0} = 1$$

while

$$\left. \frac{d^2}{d\theta^2} \frac{1}{2} (\cos \theta_1 - \cos \theta_1^0)^2 \right|_{\theta^0} = \frac{8}{9}.$$

The Hessian matrix of (B.2) has eigenvalue

$$\lambda^{be} = \frac{14}{9}. \quad (\text{B.3})$$



### B.3 trans

The cosine of the dihedral angle between the plan of  $\vec{R}_1, \vec{R}_2$  and the plan of  $\vec{R}_2, \vec{R}_3$  is

$$\cos \alpha = \frac{(\vec{R}_1 \times \vec{R}_2) \cdot (\vec{R}_2 \times \vec{R}_3)}{|\vec{R}_1 \times \vec{R}_2| |\vec{R}_2 \times \vec{R}_3|}.$$

The potential energy is

$$V^{tr} = \frac{1}{2} k_{tr} (\alpha - \alpha^0)^2 \sim k_{tr} (\cos \alpha - \cos \alpha^0)$$

and the Hessian matrix has eigenvalue

$$\lambda^{tr} = \frac{3}{2}. \quad (\text{B.4})$$

### B.4 gauche

The cosine of the dihedral angle is the same as before, the potential energy is

$$V^{ga} = \frac{1}{2} k_{tr} (\alpha - \alpha^0)^2 \sim \frac{1}{2} k_{tr} \frac{4}{3} (\cos \alpha - \cos \alpha^0)^2$$

the eigenvalue of the Hessian is:

$$\lambda^{ga} = \frac{5}{2} \quad (\text{B.5})$$



## Appendix C

# Angular contribution to the vibrational spectrum

We consider two peculiar instances of tetrahedral walk with  $N = 300$  monomers (299 steps):

- a walk in the *trans* configuration (all dihedral angles equal to  $\pi$ ), which also is the maximally elongated tetrahedral walk;
- a walk in the *gauche* configurations (all dihedral angles equal to  $\pi/3$ ).

In a log-log plot each spectrum is represented by a curve, which is roughly a line for small frequencies, that is

$$\omega_n \sim n^a \quad n/N \rightarrow 0. \quad (\text{C.1})$$

The graph reports some estimates of the exponent  $a$  which were obtained by application of the least square method on a portion of the curve which is *approximately* linear. The progressive index  $n$  is normalized by the number of not null frequencies plus 1, which is the natural normalization parameter in the spectrum of the linear chain (equation 3.2). The frequencies are relative to the highest frequency.

A fundamental difference exists between the bending and dihedral modes of the *trans* and the *gauche* configuration. The spectra of the *gauche* walk is bounded from below, while the spectra of the *trans* configuration approaches 0 through a power law of the index  $n$ . Nevertheless, the backbone interactions all together (angles and bonds, red curves) give very similar results. The spring constants are in the ratio  $k_{bb} : k_{be} : k_{tr} : k_{ga} = 1 : 1 : 1 : 0.6$ . This differs from the traditional case (reference ...) where the bond interaction  $k_{bb}$  is much stiffer. Imposing  $k_{bb} = 150$  gives a spectrum where the very high frequency region is well distinct, but the slope of the low frequency region is almost the same (see figure 3.3).

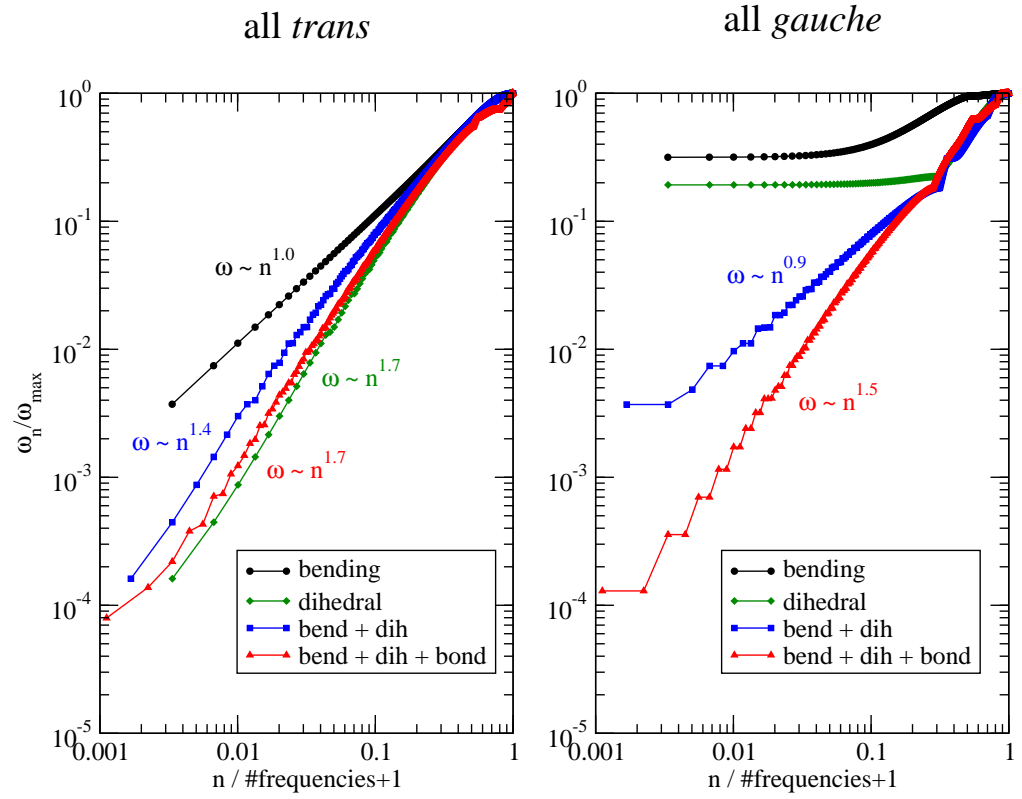


Figure C.1: **Spectrum of angular vibrations of polymers in 3 dimensions.** The chain has 300 monomers or 299 steps. **Left:** all steps in the *trans* configuration, **right:** all steps in the *gauche* configuration. The contribution of bending modes and dihedral modes are shown separately and globally.

# Bibliography

- [1] S. Alexander and R. Orbach. density of states on fractals: fractons. *Journal de Physique Lettres*, 43:625, 1982.
- [2] J. P. Allen, J. T. Colvin, D. G. Stinson, C. P. Flynn, and H. J. Stapleton. Protein conformation from electron spin relaxation data. *Biophysical journal*, 38(3):299–310, June 1982.
- [3] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. D. Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. LAPACK Users’ guide, third edition. *Numerical Algorithms*, 1999.
- [4] C. Anfinsen and E. Haber. Studies on the reduction and re-formation of protein disulfide bonds. *The Journal of biological chemistry*, 236(5):1361–3, May 1961.
- [5] N. Ashcroft and N. Mermin. *Solid state physics*. Harcourt College Publishers, 1976.
- [6] A. R. Atilgan. Anisotropy of Fluctuation Dynamics of Proteins with an Elastic Network Model. *Biophysical Journal*, 80(1):505–515, 2001.
- [7] R. H. Austin, J. P. Brody, E. C. Cox, T. Duke, and W. Volkmuth. Stretch Genes. *Physics Today*, 50(2):32, 1997.
- [8] M. Bachmann and W. Janke. Multicanonical Chain-Growth Algorithm. *Physical Review Letters*, 91(20):1–4, Nov. 2003.
- [9] M. Baiesi, E. Orlandini, and A. Stella. Scaling of a Collapsed Polymer Globule in Two Dimensions. *Physical review letters*, 96(4):40602, 2006.
- [10] I. N. Berezovsky, W. W. Chen, P. J. Choi, and E. I. Shakhnovich. Entropic stabilization of proteins and its proteomic consequences. *PLoS computational biology*, 1(4):e47, Sept. 2005.
- [11] D. J. Bergman, Y. Imry, and L. Gunther. Exactly soluble magnetoelastic lattice with a magnetic phase transition. *Journal of Statistical Physics*, 7(4):337–360, Apr. 1973.

- [12] L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley. *ScaLAPACK Users' Guide*. Society for Industrial and Applied Mathematics, 1997.
- [13] H. S. Chan and K. Dill. Intrachain loops in polymers: Effects of excluded volume. *J. Chem. Phys*, 90(1):492, 1989.
- [14] G. Chikenji, M. Kikuchi, and Y. Iba. Multi-Self-Overlap Ensemble for Protein Folding: Ground State Search and Thermodynamics. *Physical Review Letters*, 83(9):1886–1889, Aug. 1999.
- [15] B. Chu, Q. Ying, and A. Y. Grosberg. Two-Stage Kinetics of Single-Chain Collapse. Polystyrene in Cyclohexane. *Macromolecules*, 28(1):180–189, Jan. 1995.
- [16] Q. Cui and I. Bahar, editors. *Normal mode analysis*. Chapman & Hall/CRC, 2006.
- [17] P. De Gennes. *Scaling concepts in polymer physics*. Cornell University Press, 1979.
- [18] C. De Rosa, V. Petraccone, F. Dal Poggetto, G. Guerra, B. Pirozzi, M. L. Di Lorenzo, and P. Corradini. Crystal Structure of Form III of Syndiotactic Poly(p-methylstyrene). *Macromolecules*, 28(16):5507–5511, July 1995.
- [19] J. Des Cloizeaux and G. Jannink. *Polymers in solution: their modelling and structure*. Oxford University Press, USA, 1990.
- [20] K. Dill. Dominant forces in protein folding. *Biochemistry*, 29(31):7133–7155, 1990.
- [21] K. Dill, S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yee, P. D. Thomas, and H. S. Chan. Principles of protein folding -A perspective from simple exact models. *Protein Science*, 4:561–602, 1995.
- [22] M. Dogan and A. Kuntman. Study on conformational transition phenomena of poly(methyl methacrylate) in acetonitrile near theta conditions. *Polymer International*, 49(12):1648–1652, Dec. 2000.
- [23] S. Doniach, T. Garel, and H. Orland. Phase diagram of a semiflexible polymer chain in a  $\theta$  solvent: Application to protein folding. *The Journal of Chemical Physics*, 105(4):1601, 1996.
- [24] J. Doye, R. Sear, and D. Frenkel. The effect of chain stiffness on the phase behaviour of isolated homopolymers. *J. Chem. Phys*, 108:2134–2142, 1998.

- [25] J. S. Dugdale. *Entropy and its physical meaning*. Taylor and Francis, 1996.
- [26] B. Duplantier. Tricritical polymer chains in or below three dimensions. *EPL (Europhysics Letters)*, 1986.
- [27] B. Duplantier and H. Saleur. Exact tricritical exponents for polymers at the Theta point in two dimensions. *Physical Review Letters*, 59(5):539, 1987.
- [28] A. Ferrenberg and R. Swendsen. Optimized monte carlo data analysis. *Physical Review Letters*, 63(12):1195–1198, 1989.
- [29] A. Fersht. *Structure and mechanism in protein science*. Freeman, 1999.
- [30] P. J. Flory. *Principles of polymer chemistry*. Cornell University Press, 1953.
- [31] K. K. Frederick, M. S. Marlow, K. G. Valentine, and A. J. Wand. Conformational entropy in molecular recognition by proteins. *Nature*, 448(July):325–330, 2007.
- [32] J. Gibbs. *Elementary principles in statistical mechanics*. Yale U. Press, 1902.
- [33] L. Gmachowski. Thermal blob size as determined by the intrinsic viscosity. *Polymer*, 50(7):1621–1625, Mar. 2009.
- [34] P. Grassberger. Pruned-enriched Rosenbluth method: Simulations of  $\theta$  polymers of chain length up to 1 000 000. *Physical Review E*, 56(3):3682–3693, Sept. 1997.
- [35] P. Grassberger and R. Hegger. Simulations of  $\theta$ -Polymers in 2 Dimensions. *J. Phys. I France*, 5:597–606, 1995.
- [36] W. Han and Y.-D. Wu. Coarse-Grained Protein Model Coupled with a Coarse-Grained Water Model: Molecular Dynamics Study of Polyalanine-Based Peptides. *Journal of Chemical Theory and Computation*, 3(6):2146–2161, Nov. 2007.
- [37] J. Helman, A. Coniglio, and C. Tsallis. Fractons and the fractal structure of proteins. *Physical review letters*, 53(12):1195–1197, 1984.
- [38] P. Holoborodko. Smooth Noise Robust Differentiators. <http://www.holoborodko.com/pavel/numerical-methods/numerical-derivative/smooth-low-noise-differentiators/>.
- [39] J. D. Honeycutt and D. Thirumalai. The nature of folded states of globular proteins. *Biopolymers*, 32(6):695–709, June 1992.

- [40] L. Hong and G. Liu. Viscometric Study of Poly(2-cinnamoyloxyethyl methacrylate). *Macromolecules*, 43(8):3941–3946, Apr. 2010.
- [41] H. Ikeda. Solution properties of polymethyl acrylate I. Properties of PMA chain by viscometric method. *European Polymer Journal*, 40(7):1565–1574, July 2004.
- [42] E. J. Janse Van Rensburg and A. R. Rechnitzer. Multiple Markov chain Monte Carlo study of adsorbing self-avoiding walks in two and in three dimensions. *Journal of Physics A: Mathematical and General*, 37:6875–6898, 2004.
- [43] M. Karplus, T. Ichiye, and B. M. Pertirr. Configurational entropy of native proteins. *Biophysical Journal*, 52(December):1083–1085, 1987.
- [44] M. Karplus and J. N. Kushick. Method for Estimating the Configurational Entropy. *Macromolecules*, 14:325–332, 1981.
- [45] H. Kaya and H. S. Chan. Polymer principles of protein calorimetric two-state cooperativity. *Proteins*, 40(4):637–61, Sept. 2000.
- [46] K. Lau and K. Dill. A lattice statistical mechanics model of the conformational and sequence spaces of proteins. *Macromolecules*, 22(10):3986–3997, 1989.
- [47] J. Louck. Exact Normal Modes of Oscillation of a Linear Chain of Identical Particles. *American Journal of Physics*, 30:585, 1962.
- [48] N. Madras and A. D. Sokal. The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk. *Journal of Statistical Physics*, 50(1-2):109–186, Jan. 1988.
- [49] R. Mendez and U. Bastolla. Torsional Network Model: Normal Modes in Torsion Angle Space Better Correlate with Conformation Changes in Proteins. *Physical Review Letters*, 104(22):1–4, June 2010.
- [50] J. Morris and R. Gooding. Exactly solvable heterophase fluctuations at a vibrational-entropy-driven first-order phase transition. *Physical review letters*, 65(14):1769–1772, 1990.
- [51] P. Munk and T. Aminabhavi. *Introduction to macromolecular science*. Wiley-Interscience, 2002.
- [52] M. Newman and G. Barkema. *Monte Carlo methods in statistical physics*. Oxford University Press, USA, 1999.
- [53] R. Oberdorf, A. Ferguson, J. Jacobsen, and J. Kondev. Secondary structures in long compact polymers. *Physical Review E*, 74(5):1–9, Nov. 2006.



- [54] I. H. Park, Q. W. Wang, and B. Chu. Transition of linear polymer dimensions from .THETA. to collapsed regime. 1. Polystyrene/cyclohexane system. *Macromolecules*, 20(8):1965–1975, Aug. 1987.
- [55] P. Poole, A. Coniglio, N. Jan, and H. Stanley. Universality classes for the theta and theta' points. *Physical Review Letters*, 60(12):1203, 1988.
- [56] F. Rampf, K. Binder, and W. Paul. The phase diagram of a single polymer chain: New insights from a new simulation method. *Journal of Polymer Science Part B: Polymer Physics*, 44(18):2542–2555, 2006.
- [57] A. Rechnitzer and E. Rensburg. Canonical Monte Carlo determination of the connective constant of self-avoiding walks. *Journal of Physics A: Mathematical and General*, 35(42):L605, Oct. 2002.
- [58] M. Rubinstein and R. Colby. *Polymer Physics*. Oxford University Press, USA, 2003.
- [59] A. Savitzky and M. J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639, 1964.
- [60] F. Seno and A. Stella. theta point of a linear polymer in 2 dimensions: a renormalization group analysis of monte carlo enumerations. *J. Phys. I France*, 49:739–748, 1988.
- [61] F. Seno, A. Stella, and C. Vanderzande. universality class of the d=2 Theta point of linear polymers. *Physical Review Letters*, 61(13):1520, 1988.
- [62] G. D. Smith and D. Y. Yoon. Equilibrium and dynamic properties of polymethylene melts from molecular dynamics simulations. I. n-Tridecane. *The Journal of Chemical Physics*, 100(1):649, 1994.
- [63] H. Stapleton, J. Allen, C. Flynn, D. Stinson, and S. Kurtz. Fractal form of proteins. *Physical Review Letters*, 45(17):1456–1459, 1980.
- [64] G. Swislow, S. Sun, I. Nishio, and T. Tanaka. Coil-globule phase transition in a single polystyrene chain in cyclohexane. *Physical Review Letters*, 44(12):796–798, 1980.
- [65] G. Tanaka. Experimental Test of the Two-Parameter Theory of Dilute Polymer Solutions: Poly-p-methylstyrene. *The Journal of Chemical Physics*, 52(5):2639, 1970.
- [66] M. P. Taylor, W. Paul, and K. Binder. Two-state protein-like folding of a homopolymer chain. *Physics Procedia*, 4(2009):151–160, 2010.

- [67] M. Tesi, E. Rensburg, E. Orlandini, and S. Whittington. Monte Carlo study of the interacting self-avoiding walk model in three dimensions. *Journal of statistical physics*, 82(1):155–181, 1996.
- [68] E. I. Tiktopulo, V. E. Bychkova, J. Ricka, and O. B. Ptitsyn. Cooperativity of the Coil-Globule Transition in a Homopolymer: Microcalorimetric Study of Poly(N-isopropylacrylamide). *Macromolecules*, 27(10):2879–2882, May 1994.
- [69] E. I. Tiktopulo, V. N. Uversky, V. B. Lushchik, S. I. Klenin, V. E. Bychkova, and O. B. Ptitsyn. "Domain" Coil-Globule Transition in Homopolymers. *Macromolecules*, 28(22):7519–7524, Oct. 1995.
- [70] F. Valle, M. Favre, P. De Los Rios, A. Rosa, and G. Dietler. Scaling Exponents and Probability Distributions of DNA End-to-End Distance. *Physical Review Letters*, 95(15):1–4, 2005.
- [71] C. Vanderzande. *Lattice models of polymers*. Cambridge University Press, 1998.
- [72] C. Vanderzande, A. Stella, and F. Seno. Percolation, the special Theta' point and the Theta-Theta' universality puzzle. *Physical Review Letters*, 67(20):2757, 1991.
- [73] J. R. P. Vieira. Markov chain monte carlo simulation for polymer adsorption on a flat surface, 2010.
- [74] S. A. Vshivkov and A. P. Safronov. The conformational coil-globule transition of polystyrene in cyclohexane solution. *Macromolecular Chemistry and Physics*, 198(10):3015–3023, Oct. 1997.
- [75] D. J. Wales. *Energy landscapes*. Cambridge University Press, 2003.
- [76] X. Ye, Y. Ding, and J. Li. Scaling of the molecular weight-dependent thermal volume transition of poly(N-isopropylacrylamide). *Journal of Polymer Science Part B: Polymer Physics*, 48(12):1388–1393, June 2010.
- [77] J. Zhang and J. S. Liu. On side-chain conformational entropy of proteins. *PLoS computational biology*, 2(12):e168, Dec. 2006.
- [78] Y. Zhou, C. Hall, and M. Karplus. First-Order Disorder-to-Order Transition in an Isolated Homopolymer Model. *Physical review letters*, 77(13):2822–2825, Sept. 1996.
- [79] Y. Zhou, C. K. Hall, and M. Karplus. The calorimetric criterion for a two-state process revisited. *Protein science : a publication of the Protein Society*, 8(5):1064–74, May 1999.

- [80] G. Zifferer. Monte carlo simulation of tetrahedral chains, 6 Linear and star-branched polymers near to theta conditions. *Macromol. Chem., Theory Simul.*, 2:653–671, 1993.



# Curriculum Vitae

**Carlo Maffi**

Born on **22 February 1980**

**Italian** citizenship

Mobile: +41 76 713 39 06

carloarlen@gmail.com

## Education

---

**2007-2011**      PhD in physics, Ecole Polytechnique Fédérale de Lausanne, Laboratoire de biophysique statistique  
**1999-2006**      Bachelor and Master in Physics, Università di Milano, Italy

## Academic Projects

---

**PhD thesis**, focused on the entropic motion in biomolecules.

This topic is relevant for understanding the dynamics of complex biological systems, such as proteins or DNA. The study was based on Monte Carlo simulations of models of molecules. I implemented the algorithm in a serial and a parallel version (Parallel Tempering). The simulations run on dedicated clusters.

**Master thesis**, study of a model of gene networks. Grade 110/110 cum laude.

Boolean networks represent a system of variables subject to internal constraints which models the network of genes and transcription factors. I described the behavior of a particular kind of network by means of numeric and analytic analysis.

## Technical Skills

---

Programming languages: C++, C, Fortran, Message Passing Interface (MPI), OpenMP, Bash, Eclipse  
Software: Mathematica, Matlab, Latex, Office  
Operative Systems: Linux, Windows

## Languages

---

Italian	French	English
Native language	Currently used language (C1)	Good knowledge (B2)

## Publications

---

C. Maffi, M. Baiesi, L. Casetti, F. Piazza, P. De Los Rios, *First order coil-globule transition driven by vibrational entropy*, Submitted (2011).

L. Ciandrini, C. Maffi., A. Motta, B. Basseti and M. Cosentino Lagomarsino, *Feedback topology and XOR-dynamics in Boolean networks with varying input structure*, Phys. Rev. E 80, 026122 (2009).

C. Maffi and V. Rezzonico, *Simulation de chaines de Markov pour la biophysique*, Flash Informatique F18/2009 (EPFL).

## Assistant activity

---

Biophysique II, Prof. G. Dietler (spring 2008)

Physique Statistique I, Prof. D. Ivanov (fall 2008, fall 2009)

Statistical physics of biomacromolecules, Prof. P. De Los Rios (spring 2009, spring 2010)

Statistical physics III, Prof. P. De Los Rios (fall 2010)

**Tutor** of the master student Joana Vieira during her thesis titled “Markov chain Monte Carlo simulation for polymer adsorption on a flat surface” (spring 2010)

## Poster

---

*An interacting self-avoiding walk with vibrations at*

International summer school “Fundamental problems in statistical physics XII”, Leuven, Belgium, August 31<sup>st</sup> – September 11<sup>th</sup> 2009

Winter doctoral school “Studying biomolecules by experiment and theory: a biophysicist’s view”, Crans-Montana, Switzerland, 14<sup>th</sup>-20<sup>th</sup> February 2010

## Oral presentations

---

*Conformational entropy in macromolecules*, Swiss soft day, Fribourg, Switzerland, October 20<sup>th</sup> 2010

*Entropy driven first order coil to globule transition*, Dynamics of complex systems workshop, Cergy-Pontoise, France, September 8<sup>th</sup> 2011

## Work Experience

---

- For several years I taught mathematics and physics in individual courses and classes for high school and university level students. In 2005-2006 I prepared a class of working people for the attainment of the high school degree.
- 2006: Teaching Italian for foreign people in a voluntary organization

## Other activities

---

I have a strong passion for classical music, to which I dedicated 10 years of studies for achieving my **Bachelor of Piano (2002)**. Now I play regularly in amateur chamber music formations.

I love to dance Tango Argentino and I am member of the organizing committee of the EPFL student association Architango.