

---

# Dynamics of discrete variables in random graphs

---



## DOCTORAL THESIS

Author: David Machado Pérez  
Tutor: Dr. Roberto Mulet Genicio

Center for Complex Systems  
Physics Faculty  
University of Havana

May 2024



# Acknowledgements

I thank Mulet for everything I learned in science and beyond. To Eduardo for his original ideas and his help from the beginning. To Erik and Federico for welcoming me there, for their always useful comments, and for the work experience. To Lage and Ernesto, colleagues in the PN223LH010-013), for their interest and contribution to the thesis. To Martín for the shared work. To Luisa for dedicating time to me amidst Roman bureaucracy. To the friends who tried and gradually understood the titles of the presentations. To the friends who don't care about the titles. To my family for supporting me without question. To my brother who is starting now. To my parents, responsible for all of this, for their constant effort. To Cristina for giving me the motivation to continue.



# Abstract

The dynamics of sets of discrete variables defined on random graphs is a topic with relatively few results in the scientific literature. The development of tools to study these systems, however, has direct implications for problems in combinatorial optimization, opinion dynamics, epidemic propagation, condensed matter physics, among other fields. This doctoral thesis in physical sciences aims to consolidate the recently introduced Cavity Master Equation (CME) as a relevant technique in this area. It also proposes designing new methods to overcome its limitations. Regarding the CME, this work recovers exact solutions known in the literature both in equilibrium and in the dynamics of some models. Its equations are simplified to obtain an average description of the temporal evolution of systems, which saves computational cost and is easier to interpret. Additionally, a hierarchical system of increasingly precise equations is introduced to address some difficulties arising from approximations made in deriving the CME. After this first part dedicated to consolidating the CME in more familiar models, the thesis returns to its main motivations: the dynamics of combinatorial optimization algorithms and spin glasses. The CME is applied to two algorithms for the famous K-Satisfiability problem. These are the Focused Metropolis Search and WalkSAT. In both cases, predictions about the algorithmic transition fail in the region where a spin glass phase is expected to appear. This leads to the development of a new technique: Approximate Conditional Dynamics (CDA). With this technique, the spin glass dynamics of the p-spin ferromagnetic model is described more accurately than with the CME. Finally, the thesis returns to the K-Satisfiability problem to compare average versions of CME and CDA. The results indicate that CDA is more advisable for continuing studies on the behavior of algorithms. Thus, the thesis contributes to refining and establishing techniques for studying disordered systems out of equilibrium, a field that until now lacks a general methodology and has scarce exact solutions.



# Author's articles

## Articles directly related to the thesis

- 2019 E. Aurell, E. Domínguez, D. Machado and R. Mulet. A theory of non-equilibrium local search on random satisfaction problems. *Physical Review Letters*, 123:230602.
- 2020 E. Domínguez, D. Machado and R. Mulet. The cavity master equation: average and fixed point of the ferromagnetic model in random graphs. *Journal of Statistical Mechanics: Theory and Experiment*, 2020:073304.
- 2021 D. Machado and R. Mulet. From random point processes to hierarchical cavity master equations for stochastic dynamics of disordered systems in random graphs: Ising models and epidemics. *Physical Review E*, 104:054303.
- 2022 E. Ortega, D. Machado and A. Lage-Castellanos. Dynamics of epidemics from cavity master equations: susceptible-infectious-susceptible models. *Physical Review E*, 105:024308.
- 2023 E. Aurell, D. Machado and R. Mulet. A closure for the master equation starting from the dynamic cavity method. *Journal of Physics A: Mathematical and Theoretical*, 56, 17LT02.
- 2023 D. Machado, R. Mulet and F. Ricci-Tersenghi. Improved mean-field dynamical equations are able to detect the two-step relaxation in glassy dynamics at low temperatures. *Journal of Statistical Mechanics: Theory and Experiment*, 2023:123301

## Other articles by the author

- 2018 E. Aurell, E. Domínguez, D. Machado and R. Mulet. Exploring the diluted ferromagnetic p-spin model with a cavity master equation. *Physical Review E*, 97:050103(R)
- 2023 Y. Hernández-Hernández, D. Machado and R. Mulet. Modeling the neural activity of *Caenorhabditis Elegans* through neural message passing. *Revista Cubana de Física*, 40(1):39



# Contents

<b>Acknowledgements</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Author's articles</b>	<b>vii</b>
<b>I Introduction</b>	<b>1</b>
<b>1 Discrete Variable Systems in Random Graphs</b>	<b>2</b>
1.1 A Useful Model: The p-Spin Ferromagnetic Model . . . . .	5
1.2 The random K-Satisfiability . . . . .	8
1.3 Dynamics . . . . .	10
<b>2 Random Point Processes and Master Equations</b>	<b>15</b>
2.1 Random point processes and master equations . . . . .	16
2.1.1 One single variable . . . . .	17
2.2 Dynamic cavity method . . . . .	19
2.2.1 Trees and dynamic cavity messages . . . . .	19
2.2.2 Differentiating the update rule . . . . .	21
2.3 Cavity Master Equation . . . . .	23
<b>II Cavity Master Equation</b>	<b>26</b>
<b>3 Equilibrium and average case</b>	<b>27</b>
3.1 Equilibrium cavity method . . . . .	27
3.2 Equivalence with message passing . . . . .	29
3.2.1 CME's stationary state and equilibrium cavity method	30
3.3 Average case . . . . .	31
3.3.1 Fully asymmetric ferromagnet . . . . .	31
3.3.2 Symmetric ferromagnet . . . . .	34
3.3.3 SIS model for epidemics . . . . .	38

<b>4 Hierarchical system of cavity master equations</b>	<b>43</b>
4.1 Derivation . . . . .	43
4.1.1 Equations for the cavity probabilities . . . . .	46
4.1.2 Closure and hierarchical approximations . . . . .	47
4.2 Numerical results . . . . .	49
4.2.1 First and second levels of approximation . . . . .	50
4.2.2 Comparison with the Dynamical Independent Neighbors Approximation . . . . .	53
4.2.3 SIS model for epidemics . . . . .	56
<b>5 Local search algorithms in the K-SAT problem</b>	<b>59</b>
5.1 Difficult instances and focused algorithms . . . . .	59
5.2 Adaptation of the Cavity Master Equation . . . . .	61
5.3 Focused Metropolis Search (FMS) . . . . .	63
5.4 WalkSAT algorithm with greedy steps . . . . .	68
5.4.1 Dynamic Independent Neighbor Approximation (DINA)	69
5.5 Cavity Master Equation and numerical results . . . . .	71
<b>III Conditional Dynamic Approximation</b>	<b>76</b>
<b>6 Spin glass dynamics</b>	<b>77</b>
6.1 Conditional Dynamic Approximation . . . . .	78
6.2 New time scale . . . . .	82
<b>7 Average case equations in K-SAT</b>	<b>89</b>
7.1 Derivation of the equations . . . . .	89
7.2 Average phase diagrams . . . . .	92
<b>Conclusions</b>	<b>97</b>
<b>Recommendations</b>	<b>99</b>
<b>IV Appendixes</b>	<b>101</b>
<b>A Point-to-set correlation in the <math>p</math>-spin model</b>	<b>103</b>
<b>B Derivation of the master equation for <math>N</math> spins</b>	<b>107</b>
<b>C The stationary CME</b>	<b>111</b>
<b>D Computation of the sums for the average case</b>	<b>115</b>

<b>E</b> Populations dynamics for the asymmetric ferromagnet	<b>117</b>
<b>F</b> Other average equations	<b>121</b>
<b>G</b> General closure for the hierarchy of master equations	<b>127</b>
<b>H</b> Average equations for random regular graphs	<b>131</b>
<b>I</b> Simplification of the CME in the K-SAT problem	<b>137</b>
<b>J</b> Determination of the algorithmic transition	<b>139</b>
<b>K</b> The Instantaneous Master Equation	<b>143</b>
K.1 Closure of the Master Equation . . . . .	143
K.2 Numerical results . . . . .	146
<b>L</b> Equivalence between the <i>CDA-1</i> and PBMF	<b>151</b>
<b>M</b> Average case for the Conditional Dynamic Approximation	<b>153</b>
<b>N</b> Cavity Master Equation for the spin glass dynamics	<b>157</b>
<b>O</b> The average CME-2 in K-SAT	<b>161</b>
<b>P</b> From the average CDA-2 to the DINA in K-SAT	<b>165</b>
<b>V</b> Bibliography and acronyms	<b>169</b>
<b>Acronyms</b>	<b>180</b>



# **Part I**

## **Introduction**

## Chapter 1

# Discrete Variable Systems in Random Graphs

Until a little more than the middle of the last century, Statistical Physics mainly dealt with ordered systems. It was known how local interactions between identical variables located in a periodic lattice could give rise to collective phenomena such as ferromagnetism [1, 2]. However, motivated by the phenomenological richness and technical difficulties associated with spin glasses, a series of works in the third quarter of the century [3–5] paved the way for the study of disordered spin systems. One of the most studied models in those years was the Sherrington-Kirkpatrick (SK) model [4], whose Hamiltonian has the form  $H(\vec{\sigma}) = -\frac{1}{2} \sum_i \sum_j J_{ij} \sigma_i \sigma_j$ . In this case, there are  $N$  binary variables ( $\vec{\sigma} = \{\sigma_1, \dots, \sigma_N\}$ ) and all interactions are of infinite range (all variables interact with all others). Each pair of variables can interact ferromagnetically ( $J_{ij} > 0$ ) or antiferromagnetically ( $J_{ij} < 0$ ), and these couplings  $J_{ij}$  are chosen randomly.

The successful application of mean-field techniques such as the cavity method, the replica trick, and replica symmetry breaking (RSB) [5] to the SK model boosted the growth of the scientific community in the Physics of Disordered Systems. From its beginnings, this community was interested in topics that traditionally fell outside the scope of physics, such as combinatorial optimization [6–8] and neuroscience [9].

It was precisely the work with optimization tasks that most promoted the inclusion of a second type of disorder. The foundational article by Mézard and Parisi [10], from 2001, successfully extended the cavity method to models of discrete variables where the structure of the interactions is itself random. In contrast to infinite-range models like SK, they considered variables that interact only with a finite number of randomly chosen *neighbors*. If the variables are identified with nodes and the interactions with connections between nodes, it is natural to represent these models using graphs. Each graph is defined by a set of nodes  $V$  and a set of edges  $E$  between those nodes. The new type of Hamiltonians  $H(\vec{\sigma}) = -\sum_{(ij) \in E} J_{ij} \sigma_i \sigma_j$  contains only interactions between nodes that share an edge.

The existence of a notion of distance brought the approaches closer to realistic

physical situations, although instead of periodic lattices, random graphs that are locally trees were studied. This latter property, explained in detail later, is essential for the successful application of the cavity method. The paradigmatic examples, since [10], have been regular random graphs (RRG) and Erdős-Rényi (ER) random graphs [11].

One of the most relevant results of the new techniques [12] returned to the field of combinatorial optimization and was published just a year after Ref. [10], in 2002. This work analytically determined certain properties of typical instances of the well-known problem of random K-satisfiability (K-SAT) [13, 14], referenced again in Section 1.2. The authors of [12] also designed a specific algorithm, the *Survey Informed Decimation* (SID), whose current variants are the most effective for finding solutions to K-SAT.

The significant conceptual and practical advance represented by [12] did not exhaust the interest of the disordered systems community in this topic. Its success extended to other cases such as graph coloring [15]. However, the new techniques were not sufficient to close the topic of combinatorial optimization. Questions continued to arise, mainly focused on the behavior of algorithms used to solve instances of K-SAT and other problems in random graphs. A series of works [16, 17] proposed the existence of regimes where finding optimal solutions should be possible but difficult. However, these limits were questioned by empirical evidence obtained from some seemingly very efficient local search algorithms [18–20].

Beyond technical difficulties, it is hard to know with certainty whether the analytically obtained limits are violated or not. There is no appropriate theory to study the evolution of any given algorithm's run. The most relevant efforts have only managed to describe the behavior of some algorithms far from the regions where K-SAT is supposed to be most difficult to solve [21]. In general, there is a lack of methodology to precisely study problems that, like this one, are completely *out of thermodynamic equilibrium*.

The out-of-equilibrium dynamics of discrete variable systems in random graphs is an outstanding task in many other topics of interest [22–31]. However, the techniques to study them are scarce, complex to implement, and successful only in specific cases [21, 23, 25, 32–35]. This thesis precisely sets out, as a general objective:

- 1. To develop methods for studying the dynamics of discrete variable systems in random graphs.**

Mainly motivated by the same combinatorial optimization problems that drove the disordered systems community at the beginning of the century, the thesis also applies these tools to other phenomena such as the relaxation towards equilibrium of ferromagnetic models and the spread of epidemics in a network.

Although other methods available in the literature are mentioned in this chapter, the thesis has a very close relationship with the so-called Cavity Master Equation (CME) [35, 36]. This closure of the well-known master equation for the dynamics of discrete variables is a central topic in much of this document, and the first four specific objectives respond to it. The thesis proposes:

- 1) To establish relationships between the CME and other known techniques.
- 2) To obtain simpler equations that represent the average behavior of the CME over different graphs.
- 3) To correct some approximations of the CME related to the importance of spatial correlations in the system.
- 4) To apply the CME to the description of local search algorithms that solve instances of K-SAT.

However, some limitations of the CME are closely linked to the essence of obtaining the equations. Therefore, the last specific objective is:

- 5) To design new techniques for studying the dynamics of discrete variables in random graphs.

## Thesis structure

The thesis consists of seven chapters divided into three parts. The first part serves as an introduction and contains two chapters that describe the problems to be studied, review some of the techniques available in the literature, and introduce the CME method from first principles.

The second part, comprising three chapters, focuses on the CME and its results. Chapter 3 demonstrates the equivalence of the CME with the cavity method in equilibrium when the dynamics satisfy detailed balance. Additionally, simpler equations are derived to describe the average behavior of the CME in certain families of graphs, and comparisons are made with known results for the dynamics.

In Chapter 4, a hierarchical system of master equations is presented, in which the CME is only the first level of approximation. It is shown that the numerical results of higher levels are appreciably more accurate, and these are also compared with other techniques in the literature. Returning to the main motivation of the thesis, the last chapter of the second part (Chapter 5) applies the CME to describe two local search algorithms used to solve K-SAT.

The third and final part presents a new technique for the dynamics of discrete variables in random graphs: the Conditional Approximate Dynamics (CDA). The formulation of the CDA introduced in Chapter 6 includes several known methods in the literature as special cases. In some instances, the numerical results surpass those of the CME itself. In the same Chapter 6, some problems of the CDA are corrected to describe the spin glass dynamics of a known model at low temperatures. Chapter 7 returns to K-SAT one last time to compare the results of average equations for the CME and the CDA.

The thesis encompasses the content of six original articles corresponding to Refs. [37–42]. To facilitate understanding, a logical order of presentation has been designed that does not respect the chronological order in which the articles appeared. Due

to space constraints, several appendices are included to complement the reading. Additionally, unpublished results are presented.

Efforts have been made to maintain the original version of the acronyms used in the thesis. Therefore, most correspond to the English names of the theories, algorithms, or techniques. This is the case, for example, with CME and CDA. Readers are advised to refer to the list of acronyms at the end of the document.

In this chapter, the p-spin ferromagnetic model is first used to introduce the essential ideas about equilibrium phenomenology, which are fundamental throughout the thesis when interpreting the results. Then, the K-SAT problem, which as explained, is one of the fundamental motivations of the thesis, is addressed in more detail. To conclude, a brief overview of the literature on out-of-equilibrium dynamics is presented.

## 1.1 A Useful Model: The p-Spin Ferromagnetic Model

The p-spin model in random graphs motivated several works in the early 2000s due to its similarity to some combinatorial optimization problems like K-SAT [43–45]. In its ferromagnetic version, it is defined with the Hamiltonian  $H(\vec{\sigma}) = -J \sum_{a=1}^M \prod_{k \in a} \sigma_k$ , where  $J > 0$  and the binary variables  $\sigma_k = \pm 1$  are historically called spins. The system is composed of  $N$  spins ( $\vec{\sigma} = \{\sigma_k\}_{k=1}^N$ ) participating in  $M$  interactions or plaquettes ( $a = 1, \dots, M$ ). Each interaction involves exactly  $p$  spins, meaning each product in the Hamiltonian relates to exactly  $p$   $\sigma_k$  variables.

This type of system is usually represented with hypergraphs (see Fig. 1.1), which are graphs with two types of nodes. The squares in Fig. 1.1 are factor nodes and symbolize plaquettes. The circles correspond to variable nodes, or the system's spins. Hypergraphs are useful for representing systems where variables interact in groups. A variable can belong to several of these groups or plaquettes, as shown in Fig. 1.1. Traditionally, indices  $a, b, \dots$  are used for the factor nodes, and indices  $i, j, k, \dots$  are left for the variable nodes. In this thesis, the symbol  $k \in b$  indicates that spin  $k$  belongs to interaction or plaquette  $b$ , which in the hypergraph means that node  $k$  is in the neighborhood of node  $b$ .

The two most studied ways of constructing the random interaction hypergraph are:

- i) *p*-regular random hypergraph: each factor node is connected to exactly  $p$  variable nodes, and each variable to  $c$  factor nodes.
- ii) Random *p*-hypergraph with variable connectivity or Erdős-Rényi: each factor node is connected to exactly  $p$  variable nodes. The connectivity of the variable nodes follows a Poisson distribution  $P(c) = e^{-\kappa} \kappa^c / c!$ , where  $\kappa$  is the average connectivity.

In both families of hypergraphs, the typical cycle length diverges as  $N \rightarrow \infty$ . A cycle is merely a path in the graph that starts at a node and returns to it, always

moving from one node to one of its neighbors without passing through the same place twice. This will be relevant for later applying the cavity method.

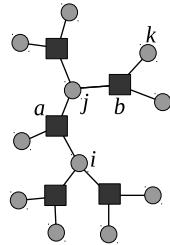


Figure 1.1: Illustration of a hypergraph with 11 variable nodes (circles) and 5 factor nodes (squares). Each square represents a plaquette that *contains* the circles neighboring that square. In this case, all plaquettes consist of exactly 3 variables, so  $p = 3$ .

According to Statistical Physics, in equilibrium the spin configurations follow the Boltzmann-Gibbs distribution  $P(\vec{\sigma}) = \frac{1}{Z} \exp\{-\beta H(\vec{\sigma})\}$ , where  $Z = \sum_{\vec{\sigma}} \exp\{-\beta H(\vec{\sigma})\}$  is the partition function and  $\beta = 1/T$  is the inverse of the temperature  $T$ . The classical formulation of the problem consists of studying a family of random graphs with determined connectivities and then obtaining the values of some observables of interest for each finite value of  $T$  [43]. The p-spin ferromagnetic model can also be posed as a combinatorial optimization problem. Its *solutions* would be the configurations of lowest energy, in this case those in which all the products of the Hamiltonian are equal to one ( $H = -M$ ). Thus, one can take the zero-temperature limit ( $\beta \rightarrow \infty$ ) and determine, given a family of graphs, the connectivity values for which it is possible to find solutions starting from a random configuration and letting the system relax [44].

In either case, the p-spin model displays a rich phenomenology. To introduce the basic elements for what follows without going into detail, here the finite-temperature problem is described in 3-regular random hypergraphs (i.e.,  $p = 3$  and  $c = 4$ ). This exhibits two types of transitions important for the thesis: the order-disorder transition and the spin glass transition.

In the p-spin ferromagnetic model, energy  $e = \langle H \rangle / N$  and magnetization  $m = \langle \sum_{k=1}^N \sigma_k \rangle / N$  can be defined, where  $\langle \cdot \rangle = \sum_{\vec{\sigma}} [\cdot] P(\vec{\sigma})$ . One of the configurations where the Hamiltonian  $H$  takes its minimum value is when  $\sigma_k = 1$  for all  $k$ . If the system moves through configurations very close to this ordered configuration, then  $m \lesssim 1$ .

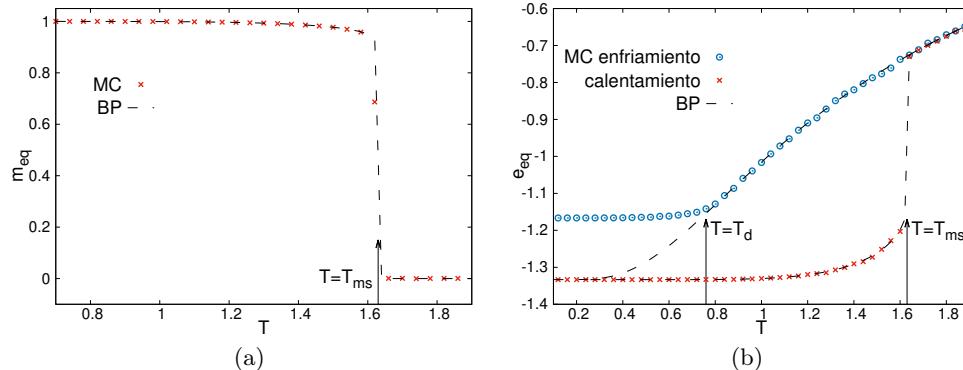


Figure 1.2: 3-spin ferromagnetic model defined on regular random hypergraphs with  $c = 4$ . Red and blue points represent Monte Carlo simulation results with system size  $N = 10000$ . Red points correspond to the slow heating of a system initially in the fully ordered configuration at low temperature ( $T = 0.12$ ). Blue points represent slow cooling, also known as Simulated Annealing, where the system starts from a random configuration at an initial temperature  $T = 2.0$ . In both cases, the temperature change magnitude is  $\Delta T = 0.02$ . The dashed lines result from the Belief Propagation (BP) algorithm. The panels indicate the temperature  $T_{ms} \approx 1.63$  where the ferromagnetic solution disappears. **(a)** Equilibrium magnetization for each temperature. **(b)** Equilibrium energy for each temperature. The temperature  $T_d \approx 0.753$  is marked where the point-to-set correlation length diverges and the spin glass transition occurs. *This figure was generated using data from Ref. [36] and is not part of the original thesis contents.*

At low temperatures, if the system is initially ordered, it is expected to remain more likely in configurations with  $m \lesssim 1$ . It is then worth asking what temperature  $T$  must be imposed to visit other configurations such that  $m(T) \approx 0$ . The red points in Fig. 1.2 result from Monte Carlo simulations [46, 47] where the initially ordered system evolves in contact with a thermal bath at a low temperature  $T = 0.12$ . After reaching equilibrium, the observables of interest are measured, the temperature is slightly increased, and the relaxation process is repeated. When this slow heating process is followed, the equilibrium value of the observables changes abruptly at a certain temperature.

To determine the temperature  $T_{ms}$  at which this discontinuous transition occurs, the well-known equilibrium cavity method [5] can be used. The corresponding Belief Propagation (BP) equations are presented in Section 3.1. Using them,  $T_{ms} \approx 1.63$  is obtained, the point where the system transitions from having  $m \approx 0$  to being magnetized. Figure 1.2a shows a very good quantitative correspondence between the Monte Carlo equilibrium magnetization values and the BP predictions. For  $T < T_{ms}$ , the system remains magnetized and in the ordered (ferromagnetic) phase, while for  $T > T_{ms}$ , it reaches an equilibrium state with  $m \approx 0$ , corresponding to the disordered

(paramagnetic) phase.

Figure 1.2b shows the results of Monte Carlo simulations different from the previous ones. The blue points were obtained by slowly cooling the system to gradually decrease the energy. This case is known as the Simulated Annealing (SA) algorithm [6, 36, 43]. Initially, a random configuration of spins is chosen, which then evolves to equilibrium at a high temperature. Each time equilibrium is reached, the temperature is slightly decreased, and the process is repeated. In this way, the system can reduce its energy more efficiently than if it had evolved from a random initial configuration at a constant low temperature.

In the low-temperature region, Monte Carlo relaxation is slow, and the energies obtained remain significantly higher than those predicted by BP. At low temperatures, an important assumption of this method becomes invalid. Long-range correlations emerge, preventing the hypergraph of interactions from being disconnected into several independent subgraphs. This is associated with the divergence of the system's relaxation time and the presence of a spin glass phase.

The works in Refs. [48–50] showed a simple way to detect when long-range correlations start to take effect. The relevant magnitude is the so-called point-to-set correlation between the spin at site  $i$  and the variables  $\bar{\sigma}_l$  at a distance  $l$  from said spin. Later, in Chapter 5, their definition and physical significance are discussed. The point-to-set correlation length diverges exactly at the spin glass transition temperature, which in the 3-spin ferromagnetic model is  $T_d \approx 0.753$ . Appendix A contains more details on determining the value of  $T_d$ . This topic will also be revisited in Chapters 5 and 6.

This section has described the order-disorder and spin glass transitions of a model on random hypergraphs: the p-spin ferromagnetic model. All this knowledge about the physics of the p-spin model, published in the early 2000s, paved the way for combinatorial optimization problems like K-SAT. Due to its importance for the rest of this thesis, the formulation of K-SAT is now addressed in more detail.

## 1.2 The random K-Satisfiability

Cook's pioneering work in 1971 [13] established two categories, P and NP, to classify computational problems according to their difficulty. This classification refers to the properties of the best-known algorithm for the decision version of the problem (where the answer is either Yes or No).

To make a fair comparison of problems, the parameter taken into account is the number  $n$  of input data that algorithms must receive. A less rigorous but straightforward definition of the set of "easier" problems would be: a problem is in P if there exists an algorithm that can provide an answer in a time bounded by a polynomial  $q_1(n)$ . In contrast, a problem is in NP if, after a possible solution is generated, there exists an algorithm that can determine whether it is correct or not in a time bounded by a polynomial  $q_2(n)$ . In other words, problems in P are solvable in polynomial time, while the solutions to NP problems can be verified in polynomial time.

It is also known that some members of NP are at least as hard as any other member of NP. These form the class of NP-complete problems, and their solution can be transformed into a solution of any other NP problem using a polynomial algorithm.

K-SAT was the first problem in history to be classified as NP-complete [13, 14] (for  $K \geq 3$ ). In its decision version, it poses the following:

**KSAT:** Given  $N$  Boolean variables  $\vec{x} = \{x_1, \dots, x_N\}$  grouped into  $M$  clauses. Each clause is the logical disjunction (logical OR) of  $K$  variables (which may appear negated). A Boolean formula  $f(\vec{x})$  is constructed, which is the conjunction (logical AND) of the  $M$  clauses.  
**Does there exist an assignment of the variables that satisfies  $f(\vec{x})$ ?**

An example of a  $K = 3$  variable clause is  $f_a = x_1 \vee \bar{x}_2 \vee x_3$ , where  $\bar{x}$  is the logical negation of  $x$ . The formula  $f = (x_1 \vee \bar{x}_2 \vee x_3) \wedge (\bar{x}_1 \vee x_2 \vee \bar{x}_3)$  is satisfied by several assignments, including  $x_1 = \text{True}$ ,  $x_2 = \text{False}$ , and  $x_3 = \text{False}$ . An instance of K-SAT is precisely a specific formula  $f$  where the variables are randomly grouped into clauses, and it is randomly chosen whether they will be negated or not. It is known that for low values of the parameter  $\alpha = M/N$ , K-SAT instances are typically satisfiable, while for high values of  $\alpha$ , a satisfactory assignment typically does not exist [12, 13]. This is known as the SAT/UNSAT transition and occurs for a certain value  $\alpha = \alpha_s$ .

Again, instances can be represented using hypergraphs like the one in Figure (1.1). In this case, clauses are symbolized by squares and Boolean variables by circles. In 3-SAT, the connectivity of variable nodes is not constant. An Erdős-Rényi hypergraph is used. The problem can be translated into finding the ground state of the Hamiltonian  $H[\vec{\sigma}] = \sum_{a=1}^M E_a$ . Here:

$$E_a = \prod_{i \in a} \frac{1 - J_i^a \sigma_i}{2}.$$

Boolean variables have been encoded in terms of other binary variables, the spins  $\sigma = \pm 1$ . The couplings  $J_i^a = \pm 1$  indicate whether the variable is negated ( $J_i^a = -1$ ) or not ( $J_i^a = 1$ ). Thus formulated, the problem can be addressed with tools from the physics of disordered systems. In the combinatorial optimization formulation, the statistical problem in the  $T \rightarrow 0$  limit is actually studied. Ref. [12] estimated that the SAT/UNSAT transition for 3-SAT occurs at  $\alpha_s(3) \approx 4.256$ . Later, a slightly higher value was obtained,  $\alpha_s(3) \approx 4.2667$  [16].

In this model, there is also a transition to the spin glass for  $\alpha = \alpha_d$  and  $T = 0$ . For  $K = 3$ ,  $\alpha_d$  is approximately 3.86 (details can be found in Ref. [51]). In the 3-spin ferromagnetic model, the transition to the spin glass directly causes the Monte Carlo algorithm to be stuck at energy values that are much higher than the minimum possible (see Fig. (1.2b)). Similarly, in the region  $\alpha_d < \alpha < \alpha_s$ , local search algorithms are expected to have great difficulty in solving K-SAT instances, although solutions generally exist because  $\alpha < \alpha_s$ . Each of these algorithms is defined by a rule that randomly selects a variable and flips it. This rule only uses information from the state of the variable and its neighborhood. As the divergence of the point-to-set

correlation length indicates that typically flipping one of the variables in the system has repercussions on unbounded sets of variables in the graph, it should be unlikely to find a solution for  $\alpha > \alpha_d$ .

However, several works after the initial estimation of  $\alpha_d$  showed empirical evidence questioning the implications of this prediction on the behavior of real algorithms [18–20, 52]. Especially important is the family of "focused" algorithms that only change the state of variables participating in unsatisfied clauses. Each solution would then be an attractor of the dynamics: once the algorithm finds a correct answer, it will remain in it with probability  $p = 1$ . The dynamics of this algorithm do not fulfill the detailed balance property, so they do not converge to a stationary state following the Boltzmann-Gibbs distribution. It is a dynamics *completely out of equilibrium*.

The behavior of focused algorithms is not entirely surprising and does not indicate that analytically calculated values for  $\alpha_d$  and other limits [16, 17] are incorrect. Theoretical results use equilibrium physics, that is, they employ the Boltzmann-Gibbs distribution as a measure in configuration space. It is not clear how significant this is for algorithms that do not respect detailed balance. It has been found that even algorithms like *Simulated Annealing* [6], which are inspired by physics and converge to the Boltzmann-Gibbs distribution for long times, can efficiently solve instances beyond the dynamic transition to the spin glass [53].

The way algorithms explore configuration space and how this could help them find solutions in difficult areas is unknown. Although the importance of studying their dynamics out of equilibrium is clear, not much progress has been made in this direction. The next section provides a brief review of the methods available in the literature and mentions the limited results in the study of algorithm behavior.

### 1.3 Dynamics

Up to this point, two problems defined on random graphs have been introduced: the p-spin ferromagnetic model and K-SAT. Their phenomenology has been described based on the results of equilibrium techniques. In both cases, they start with a Hamiltonian  $H(\vec{\sigma})$  that contains information about the random structure of interactions. The Boltzmann-Gibbs distribution  $P(\vec{\sigma}) = e^{-\beta H(\vec{\sigma})}/Z$  is used to calculate observables and determine the existence of phase transitions. However, in Section 1.2, the need to study the dynamics of variables that interact on the same random graphs, but for which the Boltzmann-Gibbs distribution is no longer valid, was explained. The system evolves according to *rules*  $r(\vec{\sigma} \rightarrow \vec{\sigma}')$  that are related to transition probabilities between different configurations, and  $P(\vec{\sigma})$  is a function of time  $P^t(\vec{\sigma})$ .

In general, finding the form of  $P^t(\vec{\sigma})$  for given rules  $r(\vec{\sigma} \rightarrow \vec{\sigma}')$  is difficult. It is even challenging to review the different methods used to find the solution exactly or approximately. The theory for dynamics is highly compartmentalized, lacking a global understanding of the phenomena of interest. Criteria such as the type of variables or the properties of interaction networks separate the available techniques into different groups very well. For example, one technique may be designed only for

discrete variables evolving continuously over random graphs.

The first major division can be made between models with continuous variables and those with discrete variables. Typically, systems with continuous variables are treated with Langevin equations [54]. Such models are especially relevant in the dynamics of disordered systems with a spin glass phase. In 1981, the pioneering work of Sompolinsky and Zippelius [55] studied the dynamics in a version of the Sherrington-Kirkpatrick (SK) model with *soft* spins where the variables are continuous. Later, Sompolinsky [56] attempted to construct a mean-field theory that, starting from dynamics, recovered the equilibrium properties of spin glasses.

Motivated by the richness of the phenomenology of these systems [57–59], a series of articles in the 1990s [60–63] pointed out some inconsistencies in Sompolinsky’s theory and exploited the idea of *weak ergodicity breaking* [64] to revisit this problem. In addition to the SK model [61], they included another simple but illustrative model with long-range interactions: the spherical  $p$ -spin [60]. This theory qualitatively reproduces relevant features of the dynamics of spin glasses such as aging.

However, this theory for continuous variables has limited impact on the study of discrete variables [65]. The necessary formalism must take a leap. Instead of the evolution of the variables themselves, the dynamics of the probability distribution of the system’s configurations is studied.

Techniques for discrete variables can be regrouped according to the properties of the graphs where interactions are defined. Some cases are less complex to handle. This happens, for example, for models with asymmetric interactions. The work of Derrida *et al.* presented, in 1987, the exact solution of a model for completely asymmetric neural networks on randomly connected graphs [66]. The same was done by other researchers for the asymmetric SK model (where all variables are connected to all others) [67, 68]. There are also approximations that use path integrals to successfully describe the parallel dynamics of asymmetric models on random graphs [69].

For the dynamics of discrete variables on symmetric graphs, even fewer results are available. Solutions have long been known for the relaxation of the Ising model on a chain (one dimension) [70] and for the fully connected Ising model. Outside of these two exact results for extreme cases regarding connectivity, there are only approximations.

An outstanding example is the Replica Dynamics Theory (DRT), which has been applied to both the symmetric SK model [71–73] and finite-connectivity random graphs [23, 24, 74]. The theory defines a finite set of order parameters and divides the configuration space into subspaces, each determined by a set of parameter values. For example, in the ferromagnetic  $p$ -spin model discussed in Section 1.1, one could choose magnetization  $m$  and energy  $e$  as observables. When the system size  $N$  is large, each pair  $(m, e)$  corresponds to a large number of configurations that would form a subspace. The DRT method assumes that if the parameters are appropriately chosen, the probability distribution of the system is constant within each subspace. The validity of this approximation has been questioned in some models [75, 76], and it is not expected to work for any dynamics out of equilibrium. However, DRT pro-

vides equations for the evolution of the observables of interest when averaging over all disorder realizations. It cannot be applied to a specific interaction graph, but its numerical results are very accurate and it is undoubtedly one of the best techniques available in the literature.

DRT approximations were introduced in the 1990s for the SK model and in the 2000s for random graphs. In parallel, other equations for the dynamics of local search algorithms in combinatorial optimization problems were introduced, again averaging over a family of random graphs [21]. The new technique was formalized in [22], showing how to construct a hierarchical system of approximations that progressively considers the effect of spatial correlations on the dynamics. This scheme turned out to be equivalent to DRT when the symmetry of the replicas is maintained, which is useful considering the greater simplicity of the equations in [22].

Given the practical difficulty of writing dynamic equations for a specific random graph, it is not surprising that attempts are made to describe the averaged behavior of an entire family of graphs. This is also consistent with a scientific tradition that seeks to capture the essence of phenomena in the simplest way possible. However, in some cases, it is relevant to study the evolution of a stochastic process on a specific graph. After the increase in data volume on real social interaction networks, several studies have focused on population dynamics such as epidemic spreading [26, 27, 77] or opinion dynamics [28–30].

Mean-field approximations for dynamics have been widely used in epidemic spreading models. Here, states or compartments in which an individual can be found at each time step are generally defined. For example, in the well-known Susceptible-Infectious-Recovered (SIR) model, there are three compartments: susceptible (S), infectious (I), and recovered (R). The system’s evolution is then governed by the master equation [54] for the probabilities of finding an individual in a specific state  $X$  at time  $t$ . In general, the equations for  $P^t(X)$  depend on probabilities defined over pairs of interacting individuals. When writing equations for pair probabilities, these involve probabilities of triplets of individuals. This forms a hierarchical system of equations that needs to be truncated at some point.

The Individual-Based Mean Field (IBMF) approximation disregards correlations between a variable and its neighbors to write differential equations that depend only on  $P^t(X)$  [27, 77]. As a correction to this scheme, the Pair-Based Mean Field (PBMF) approximation takes into account correlations between each pair of neighbors [32, 78], disregarding correlations between triplets of individuals this time. PBMF results in prototypical models such as Susceptible-Infectious-Susceptible (SIS) and the aforementioned SIR are accurate in random graphs. They are especially useful for predicting the steady state. Chapter 4 returns with more details on the topic of PBMF.

Despite the existence of the mentioned techniques, recent works have focused on another type of scheme that is a generalization of the cavity method, known by various names such as Dynamic Message Passing (DMP) or simply Dynamic Cavity Method (DCM). Its essential idea is to factorize the probability distribution characterizing the entire system dynamics, taking advantage of the properties of graphs that are locally

trees. The procedure, which will be discussed in subsection 2.2.1, is analogous to that used for equilibrium.

In 2009, the cavity method was first generalized to discrete-time dynamics of the Ising model on a random graph [33]. Initially, the computational complexity of the solution for reversible dynamics like the Ising model restricted results to short times and steady states for both parallel [33, 79] and sequential rules [80]. On the other hand, with DCM, the exact solution to the SIR model for epidemics on random graphs that are locally trees was quickly found [25, 81]. Other works [82] also took advantage of the fact that irreversible dynamics like SIR are simpler to deal with, both in continuous and discrete time.

Subsequently, methods for approximately solving the DCM equations in discrete time were proposed. In [83], a short memory closure was employed, which is restricted to dynamics with short-range temporal correlations. The discrete nature of time allows, in this case, to select exactly how many past steps are relevant to determine the probability of the system's instantaneous configuration. A more accurate but more complicated to implement and interpret closure was introduced in [34].

The pioneering work of [35] extended the Dynamic Cavity Method in 2017 to models of discrete variables in continuous time. The cavity master equation (CME) is a closure of the master equation for dynamics [54] written for auxiliary probabilities, called cavity probabilities. This technique obtains differential equations from the dynamic message passing equation, analogous to the equations presented in [83], where the variables are marginals of dynamic cavity messages. Although low spatial and temporal correlation approximations are made, this method can obtain numerical results for continuous-time dynamics on a single graph with low computational cost. The first results were obtained for Ising models on random graphs with pairwise interactions. A year later, it was shown how to extend the theory to many-body interactions (hypergraphs), and the ferromagnetic p-spin model on regular random graphs was numerically explored [36].

A significant part of this doctoral thesis is dedicated to the CME. Several original results are included starting from Chapter 3. In the second part of the thesis, stationary solutions of the CME are compared with solutions of the BP algorithm [37], equations averaged over families of random graphs are obtained [37, 38], a hierarchical system of cavity equations is presented [39], and the CME is applied to K-SAT [40].

To overcome some limitations of the CME, new techniques are also proposed in the thesis. In Appendix K, a new closure for the master equation is presented, which is inspired by the DCM itself and was recently published in 2023 [41]. It should be noted that the current interest in the DCM is not limited to the CME or similar approximations. In an even more recent article [84], an optimized version of the original technique from [34] was presented, called Message Passing using Matrix Products (MPBP). With DCM, it is also possible to change perspective and study the structure of the stationary probability distribution for long times, as demonstrated in [85].

Although it is evident that the DCM motivates many works today, Chapter 6

introduces a closure for the master equation that is simpler than the CME. The new Approximate Conditional Dynamics (CDA) provides better predictions for the dynamics of the ferromagnetic p-spin model at low temperatures [42]. The CDA is much closer to the equations of [21, 22], and PBMF [32] than to the dynamic cavity method. However, it provides a better description of the spin glass dynamics of this model.

To aid in the understanding of the rest of the thesis, the next chapter (Chapter 2) presents the Theory of Random Point Processes and its connection with master equations for dynamics, passing through the dynamic cavity method.

## Chapter 2

# Random Point Processes and Master Equations

To aid in the understanding of the following chapters, here's a brief introduction to relevant aspects of the Theory of Random Point Processes (TRPP), the equations of the Dynamic Cavity Method in continuous time, and its connection with the cavity master equation (CME) as presented in Ref. [35].

The master equation [54] for the dynamics of  $N$  binary variables  $\vec{\sigma} = \{\sigma_1, \dots, \sigma_N\}$  takes the following form:

$$\frac{dP^t(\vec{\sigma})}{dt} = - \sum_{i=1}^N [r_i(\vec{\sigma})P^t(\vec{\sigma}) - r_i(F_i[\vec{\sigma}])P^t(F_i[\vec{\sigma}])], \quad (2.1)$$

where  $F_i[\cdot]$  is an operator that transforms the state  $\vec{\sigma} = \{\sigma_1, \dots, \sigma_i, \dots, \sigma_N\}$  into the state  $F_i[\vec{\sigma}] = \{\sigma_1, \dots, -\sigma_i, \dots, \sigma_N\}$ , and  $r_i(\vec{\sigma})$  is the transition probability per unit time from configuration  $\vec{\sigma}$  to configuration  $F_i[\vec{\sigma}]$ . Note that Eq. (2.1) does not consider transitions where more than one variable changes at the same time. The states accessible from  $\vec{\sigma}$  are only the  $N$  possible  $F_i[\vec{\sigma}]$ , with  $i = 1, \dots, N$ .

Eq. (2.1) gives the temporal derivative of the probability distribution of the system configurations at time  $t$ . There are  $2^N$  differential equations like (2.1), forming a system impossible to solve numerically for large  $N$ . However, in many cases of interest, the functions  $r_i$  depend only on one variable and its neighbors  $\sigma_{\partial i}$ . If the above expression is summed over all configurations that have the same value of  $\sigma_i$ , for some site  $i \leq N$ , the local form of the master equation is obtained:

$$\frac{dP^t(\sigma_i)}{dt} = - \sum_{\sigma_{\partial i}} [r_i(\sigma_i, \sigma_{\partial i})P^t(\sigma_i, \sigma_{\partial i}) - r_i(-\sigma_i, \sigma_{\partial i})P^t(-\sigma_i, \sigma_{\partial i})]. \quad (2.2)$$

The left-hand side of Eq. (2.2) is the temporal derivative of the probability that the variable associated with node  $i$  takes the value  $\sigma_i$  at time  $t$ . The right-hand side, however, depends on the joint probabilities of the configurations of the same variable and all its neighbors. The system of equations obtained by considering (2.2) for all

values of  $i$  cannot be solved in the way it is formulated. One would need differential equations for all probabilities  $P(\sigma_i, \sigma_{\partial i})$ . It is then said that the system of equations is not *closed*.

The CME provides a way to close the system (2.2). As will be seen later, it introduces differential equations for auxiliary probabilities: the cavity probabilities. Its derivation starts from the TRPP to differentiate the dynamic message passing equations.

## 2.1 Random point processes and master equations

In this section, a brief introduction to TRPP [54, 86] is provided. After some basic definitions, subsection 2.1.1 addresses an independent binary variable that changes its state randomly according to transition probabilities. The formalism of Random Point Processes is used to obtain the master equation governing the dynamics. In Appendix B, the same procedure is extended to derive the master equation with  $N$  spins.

The most important mathematical objects of TRPP are the trajectories or histories of spins, symbolized by the letter  $X$ . For binary variables  $\sigma = \pm 1$ , a specific history starting at  $\sigma(t_0) = \sigma_0$  and ending at  $\sigma(t) = \sigma$  can be parameterized using the number of times the variable changes state and the times at which these jumps occur. The probability density of one of these histories given an initial value  $\sigma(t_0) = \sigma_0$  can be denoted as:

$$Q^t(X) = Q_s(t_0, t_1, \dots, t_s, t \mid \sigma(t_0) = \sigma_0). \quad (2.3)$$

$Q^t(X)$  represents the probability density of having a trajectory  $X$  that started at  $\sigma_0$  and then had jumps in the intervals  $(t_1, t_1 + dt_1), \dots, (t_s, t_s + dt_s)$ . To emphasize that the final time is  $t$ , one can explicitly write  $X \equiv X(t)$ . For simplicity, in most equations, the condition  $\sigma(t_0) = \sigma_0$  will be omitted when writing  $Q^t(X)$ .

The probability density  $Q^t(X)$  satisfies the following normalization relation:

$$1 = \sum_X^t Q^t(X) = \sum_{s=0}^{\infty} \int_{t_0}^t dt_1 \int_{t_1}^t dt_2 \dots \int_{t_{s-1}}^t dt_s Q_s(t_0, t_1, \dots, t_s, t \mid \sigma(t_0) = \sigma_0). \quad (2.4)$$

The sum  $\sum_X^t$  in Eq. (2.4) traverses all possible histories between  $t_0$  and  $t$  that start with  $\sigma(t_0) = \sigma_0$ . It is explicitly written in the right-hand side of Eq. (2.4) as a sum  $\sum_{s=0}^{\infty}$  of temporal integrals. Each term in this sum corresponds to histories that have exactly  $s$  jumps. The integrals take into account all possible times at which these jumps occur.

Instantaneous quantities can be written as marginals of  $Q^t(X)$ . For example:

$$P^t(\sigma) = \sum_{X \mid \sigma(t)=\sigma}^t Q^t(X), \quad (2.5)$$

where the sum  $\sum_{X|\sigma(t)=\sigma}^t$  traverses all histories  $X(t)$  such that  $\sigma(t) = \sigma$  and can be written similarly to  $\sum_X^t$ . In what follows, the symbol  $\sum_{X|\sigma}^t$  will be used to shorten the notation of  $\sum_{X|\sigma(t)=\sigma}^t$ .

The master equation is a common tool in the study of continuous-time dynamics. As a proof of the validity of TRPP, the following subsections demonstrate its equivalence with the formulation of master equations in two examples: the temporal evolution of a non-interacting variable and the dynamics of a system of  $N$  spins. The properties of  $P^t(\sigma)$  will be obtained by working directly with  $Q^t(X)$ .

### 2.1.1 One single variable

The master equation for the temporal evolution of a single variable without interaction is:

$$\frac{dP^t(\sigma)}{dt} = -r(\sigma)P^t(\sigma) + r(-\sigma)P^t(-\sigma), \quad (2.6)$$

where  $r(\sigma)$  is the transition probability per unit time between the states  $\sigma$  and  $-\sigma$ .

To obtain Eq. (2.6) using the TRPP, we will differentiate both sides of (2.5). This operation must be done carefully because the increment in time  $t$  implies a change in the very space where the probability density  $Q^t(X)$  is defined. Instead of using differentiation rules, we will employ the definition of the derivative as the limit of the incremental quotient:

$$\frac{dP^t(\sigma)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{P^{t+\Delta t}(\sigma) - P^t(\sigma)}{\Delta t}. \quad (2.7)$$

If we fix the final state of the trajectory as  $\sigma(t) = \sigma$  and consider a random point process with exponentially distributed jump times, then the probability density  $Q^t(X)$  takes the form [54, 86]:

$$Q^t(X) = r(\sigma_0)e^{-r(\sigma_0)(t_1-t_0)}r(-\sigma_0)e^{-r(-\sigma_0)(t_2-t_1)} \dots r(-\sigma)e^{-r(-\sigma)(t_s-t_{s-1})}e^{-r(\sigma)(t-t_s)}. \quad (2.8)$$

The right-hand side of Eq. (2.8) is the probability density of the waiting times of the history  $X : [t_0, t] \rightarrow \{1, -1\}$  whose jumps occur at  $(t_1, t_2, \dots, t_s)$ . These waiting times are exponentially distributed with parameter  $r(\sigma)$ .

Now we will write  $P^{t+\Delta t}(\sigma)$  as a marginal of  $Q^{t+\Delta t}(X)$  and then expand the result to the first order in  $\Delta t$ . Similarly to Eq. (2.4), we have:

$$P^{t+\Delta t}(\sigma) = \sum_s \int_{t_0}^{t+\Delta t} dt_1 \int_{t_1}^{t+\Delta t} dt_2 \dots \int_{t_{s-1}}^{t+\Delta t} dt_s Q^{t+\Delta t}(X). \quad (2.9)$$

Here, the sum on the right-hand side runs over all possible numbers of jumps (denoted by  $s$ ) such that  $\sigma(t) = \sigma$ .

To expand up to order  $\Delta t$ , we will first work with  $Q^{t+\Delta t}(X)$ , leaving the rest intact. Eq. (2.8) can be rearranged as follows:

$$Q^{t+\Delta t}(X) = \left[ \prod_{l=1}^s r(\sigma(t_l)) \right] \exp \left\{ - \int_{t_0}^{t+\Delta t} r(\sigma(\tau)) d\tau \right\}. \quad (2.10)$$

where the jumps in the history  $X$  occur at times  $t_l$ , with  $l = 1, 2, \dots, s$ .

In Eq. (2.10) the probability of having a jump in the interval  $[t, t + \Delta t]$  is  $r(-\sigma)\Delta t$ , because to ensure that  $\sigma(t + \Delta t) \equiv \sigma$  the last change can only be from  $-\sigma$  to  $\sigma$ . In general, the probability of  $n$  jumps occurring is proportional to  $(\Delta t)^n$ , meaning there will be no jumps in  $[t, t + \Delta t]$  with probability 1 as  $\Delta t \rightarrow 0$ . This implies that none of the times  $t_l$  in Eq. (2.10) belong to that interval.

We can then expand the exponential in (2.10):

$$\begin{aligned} Q^{t+\Delta t}(X) &= \left[ \prod_{l=1}^s r(\sigma(t_l)) \right] \exp \left\{ - \int_{t_0}^t r(\sigma(\tau)) d\tau \right\} [1 - r(\sigma)\Delta t] + o(\Delta t) \\ Q^{t+\Delta t}(X) &= Q^t(X)[1 - r(\sigma)\Delta t] + o(\Delta t). \end{aligned} \quad (2.11)$$

On the other hand, the sum of the iterated integrals in Eq. (2.9) is of order  $(\Delta t)^s$  by itself. For each integral, we have the property  $\int_{t_0}^{t+\Delta t} dt = \int_{t_0}^t dt + \int_t^{t+\Delta t} dt$  and we can write:

$$\begin{aligned} \sum_s \int_{t_0}^{t+\Delta t} dt_1 \int_{t_1}^{t+\Delta t} dt_2 \dots \int_{t_{s-1}}^{t+\Delta t} dt_s &= \sum_s \int_{t_0}^t dt_1 \int_{t_1}^t dt_2 \dots \int_{t_{s-1}}^t dt_s + \\ &\quad + \sum_s \int_{t_0}^t dt_1 \int_{t_1}^t dt_2 \dots \int_t^{t+\Delta t} dt_s + o(\Delta t). \end{aligned} \quad (2.12)$$

The first term on the right-hand side of Eq. (2.12) is an operator that acts on the space of histories that occur in the interval  $[t_0, t]$ . It can be straightforwardly applied to the expansion (2.11) of the probability density  $Q^{t+\Delta t}(X)$  to obtain a contribution of order  $O((\Delta t)^0) \equiv O(1)$ :

$$I_0 = \sum_s \int_{t_0}^t dt_1 \int_{t_1}^t dt_2 \dots \int_{t_{s-1}}^t dt_s Q^t(X) = P^t(\sigma). \quad (2.13)$$

and one of order  $O(\Delta t)$ :

$$I_1 = - \sum_s \int_{t_0}^t dt_1 \int_{t_1}^t dt_2 \dots \int_{t_{s-1}}^t dt_s Q^t(X) r(\sigma)\Delta t = -P^t(\sigma) r(\sigma)\Delta t. \quad (2.14)$$

Together,  $I_0 + I_1$  represent the probability of having  $\sigma(t) = \sigma$  and no jumps in the interval  $[t, t + \Delta t]$ .

In the second sum of (2.12), there is an operator that acts on the space of histories that have a single jump in the interval  $[t, t + \Delta t]$ . Therefore, a second contribution

of order  $O(\Delta t)$  can be obtained by applying this operator to the probability density  $Q^{t+\Delta t}(X)$  for a single jump occurring in  $[t, t + \Delta t]$ . Recalling the parametrization (2.10) and expanding in powers of  $\Delta t$ :

$$\begin{aligned} I_2 &= \sum_s \int_{t_0}^t dt_1 \dots \int_{t_{s-2}}^t dt_{s-1} \left[ \prod_{l=1}^{s-1} r(\sigma(t_l)) \right] e^{-\int_{t_0}^t r(\sigma(\tau)) d\tau} \times \\ &\quad \times \int_t^{t+\Delta t} dt_s r(\sigma(t_s)) e^{-\int_t^{t+\Delta t} r(\sigma(\tau)) d\tau} \\ I_2 &\approx \left( \sum_s \int_{t_0}^t dt_1 \dots \int_{t_{s-2}}^t dt_{s-1} \left[ \prod_{l=1}^{s-1} r(\sigma(t_l)) \right] e^{-\int_{t_0}^t r(\sigma(\tau)) d\tau} \right) (r(-\sigma) \Delta t). \end{aligned} \quad (2.15)$$

In Eq. (2.15), it is used that to satisfy the condition  $\sigma(t) = \sigma$ , the last jump must go from the value  $-\sigma$  to the value  $\sigma$ , with a probability per unit time  $r(-\sigma)$ . The sum inside the parentheses in the equation runs over all histories that end in  $\sigma(t) = -\sigma$ . Therefore, it is equal to the probability  $P^t(-\sigma)$ . It can be seen then that  $I_2 = P^t(-\sigma) r(-\sigma) \Delta t + o(\Delta t)$  represents the probability of having  $\sigma(t) = -\sigma$  and only one jump in  $[t, t + \Delta t]$ .

Putting all these expressions together:

$$\begin{aligned} P^{t+\Delta t}(\sigma) &= I_0 + I_1 + I_2 + o(\Delta t) \\ P^{t+\Delta t}(\sigma) &= P^t(\sigma) - P^t(\sigma) r(\sigma) \Delta t + P^t(-\sigma) r(-\sigma) \Delta t + o(\Delta t). \end{aligned} \quad (2.16)$$

If  $P^t(\sigma)$  is subtracted from both sides of Eq. (2.16) and divided by  $\Delta t$ , it is enough to take the limit  $\Delta t \rightarrow 0$  to obtain the desired result (2.6):

$$\frac{dP^t(\sigma)}{dt} = -r(\sigma)P^t(\sigma) + r(-\sigma)P^t(-\sigma).$$

## 2.2 Dynamic cavity method

This section introduces the Dynamic Cavity Method as presented in [35] for continuous time. The dynamic message-passing equations, which are the basis for deducing the cavity master equation, are written.

In subsection 2.2.1, the Dynamic Cavity Method (DCM) for graphs that are locally tree-like is introduced. The main objects of the DCM are the cavity messages, which in this case explicitly depend on the trajectories of the variables. In subsection 2.2.2, it is shown how to derive the equations for these messages with respect to time.

### 2.2.1 Trees and dynamic cavity messages

This subsection introduces the dynamic cavity messages for graphs that are locally tree-like. Now there are  $N$  spins, and each spin  $\sigma_i$  interacts with its neighbors:  $\sigma_{\partial i}$ . Therefore, the transition probabilities per unit time depend only on  $\sigma_i$  and  $\sigma_{\partial i}$ .

The probability density of the set of individual histories of the  $N$  spins,  $Q^t(\vec{X})$ , can be expressed as a product of the probability densities  $\Phi_i^t(X_i | \vec{X}_{\setminus i})$  of the occurrence of the history  $X_i$  when the trajectories of the rest of the variables  $\vec{X}_{\setminus i}$  are fixed [87]:

$$Q^t(X_1, X_2, \dots, X_N) = \prod_{i=1}^N \Phi_i^t(X_i | \vec{X}_{\setminus i}). \quad (2.17)$$

In this scenario, it is possible to select a spin  $\sigma_i$  and rewrite equation (2.17) using the fact that, in a neighborhood of  $i$ , the graph is indistinguishable from a tree:

$$Q^t(X_1, \dots, X_N) = \Phi_i^t(X_i | X_{\partial i}) \prod_{k \in \partial i} \left[ \Phi_k^t(X_k | X_{\partial k}) \prod_{m \in \partial k \setminus i} \left( \Phi_k^t(X_m | X_{\partial m}) \prod_{l \in \partial m \setminus k} \dots \right) \right]. \quad (2.18)$$

The symbol  $X_{\partial i}$  represents the set of histories of the nodes in the neighborhood of  $i$ . Let  $G_k^{(i)}$  be the subgraph that expands from site  $k$  after removing the connection  $(ik)$ . The set  $\{X\}_{ik}$  contains all the histories associated with the variables that belong to  $G_k^{(i)}$ , except for  $X_k$ . Then  $Q^t(\vec{X})$  can be expressed as:

$$Q^t(X_1, \dots, X_N) = \Phi_i^t(X_i | X_{\partial i}) \prod_{k \in \partial i} M_{ki}^t(X_i, X_k, \{X\}_{ik}). \quad (2.19)$$

Here,  $M_{ki}^t$  is a symbol for the expression in brackets in Eq. (2.18). Summing  $Q^t(\vec{X})$  over all histories except  $X_i$  and  $X_{\partial i}$  leads to:

$$Q^t(X_i, X_{\partial i}) = \Phi_i^t(X_i | X_{\partial i}) \prod_{k \in \partial i} \mu_{k \rightarrow (ki)}^t(X_k | X_i). \quad (2.20)$$

The new functions  $\mu_{k \rightarrow (ki)}^t(X_k | X_i)$ , called dynamic cavity messages, are the marginals:

$$\mu_{k \rightarrow (ki)}^t(X_k | X_i) = \sum_{\{X\}_{ik}}^t M_{ki}^t(X_i, X_k, \{X\}_{ik}). \quad (2.21)$$

Each dynamic cavity message is interpreted as the probability density of the history  $X_k$  given a fixed history  $X_i$ . If two neighboring nodes are now taken and similar reasoning is applied, it can be written as:

$$Q^t(X_i, X_j) = \mu_{i \rightarrow (ij)}^t(X_i | X_j) \mu_{j \rightarrow (ji)}^t(X_j | X_i). \quad (2.22)$$

The messages  $\mu$  can be parameterized similarly to how it was done in Eq. (2.8):

$$\begin{aligned} \mu_{i \rightarrow (ij)}^t(X_i | X_j) &= \lambda_{i \rightarrow (ij)}(X_i, X_j, t_0) e^{-\int_{t_0}^{t_1} \lambda_i^\tau d\tau} \times \lambda_{i \rightarrow (ij)}(X_i, X_j, t_1) e^{-\int_{t_1}^{t_2} \lambda_i^\tau d\tau} \times \\ &\dots \times \lambda_{i \rightarrow (ij)}(X_i, X_j, t_{s_i}) e^{-\int_{t_{s_i}}^t \lambda_i^\tau d\tau}, \end{aligned} \quad (2.23)$$

only now the jumps occur with unknown transition probabilities per unit time  $\lambda_{i \rightarrow (ij)}(X_i, X_j, t)$ , which are generally functions of the histories  $X_i$  and  $X_j$  taken from  $t_0$  to  $t$ .

The interaction terms  $\Phi_i^t(X_i|X_{\partial i})$  can be interpreted as the probability density of  $X_i$  conditioned on the histories of the spins in  $\partial i$ . They can be parameterized similarly to the cavity messages:

$$\Phi_i^t(X_i|X_{\partial i}) = \prod_{l_i=1}^{s_i} r_i(\sigma_i(t_{l_i}), \sigma_{\partial i}(t_{l_i})) \exp\left\{-\int_{t_0}^t r_i(\sigma_i(\tau), \sigma_{\partial i}(\tau)) d\tau\right\}. \quad (2.24)$$

If (2.20) is summed over  $X_{\partial i \setminus \{i,j\}}$  and the result is combined with (2.22), the dynamic message-passing equation is obtained:

$$\mu_{i \rightarrow (ij)}^t(X_i | X_j) = \sum_{X_{\partial i \setminus j}} \Phi_i^t(X_i | X_{\partial i}) \prod_{k \in \partial i \setminus j} \mu_{k \rightarrow (ki)}^t(X_k | X_i). \quad (2.25)$$

### 2.2.2 Differentiating the update rule

To simplify notation, in what follows the symbol  $\mu_{i \rightarrow (ij)}^t$  will also be used for the dynamic cavity messages, omitting the dependence on  $X_i$  and  $X_j$ . The sum on the right-hand side of (2.25) can be written in more detail. Let  $F$  be the argument of the sum:

$$\sum_{\{X_k\}, k \in \partial i \setminus j}^{[t_0, t]} F(X_i, X_{\partial i}, t) = \sum_{\{s_k\}, k \in \partial i \setminus j} \left[ \prod_{k=1}^d \int_{t_0}^t dt_1^k \int_{t_1^k}^t dt_2^k \dots \int_{t_{s_k-1}^k}^t dt_{s_k}^k \right] F(X_i, X_{\partial i}, t). \quad (2.26)$$

Differentiating Eq. (2.25) must be handled as carefully as in Section 2.1. We will calculate the limit:

$$\lim_{\Delta t \rightarrow 0} \frac{\mu_{i \rightarrow (ij)}^{t+\Delta t}(X_i | X_j) - \mu_{i \rightarrow (ij)}^t(X_i | X_j)}{\Delta t}. \quad (2.27)$$

To expand  $\mu_{i \rightarrow j}(t + \Delta t)$  to the first order in  $\Delta t$ , we will consider, as in Section 2.1, that no jumps occur in the interval  $[t, t + \Delta t]$ . Thus:

$$\begin{aligned} \mu_{i \rightarrow (ij)}^{t+\Delta t} &= \prod_{l_i=1}^{s_i} \lambda_{i \rightarrow (ij)}(t_{l_i}) \exp\left\{-\int_{t_0}^{t+\Delta t} \lambda_{i \rightarrow (ij)}(\tau) d\tau\right\} \\ \mu_{i \rightarrow (ij)}^{t+\Delta t} &= [1 - \lambda_{i \rightarrow (ij)}(t)\Delta t] \prod_{l_i=1}^{s_i} \lambda_{i \rightarrow (ij)}(t_{l_i}) \exp\left\{-\int_{t_0}^t \lambda_{i \rightarrow (ij)}(\tau) d\tau\right\} + o(\Delta t) \\ \mu_{i \rightarrow (ij)}^{t+\Delta t} &= [1 - \lambda_{i \rightarrow (ij)}(t)\Delta t] \mu_{i \rightarrow (ij)}^t + o(\Delta t). \end{aligned} \quad (2.28)$$

With probability 1 when  $\Delta t \rightarrow 0$ , the derivative of the left-hand side of Eq. (2.25) is equal to  $-\lambda_{i \rightarrow (ij)}(t) \mu_{i \rightarrow (ij)}^t$ .

The expansion to the order of  $\Delta t$  of the probability densities  $\Phi_i^{t+\Delta t}$  gives an analogous result, where  $r_i$  replaces  $\lambda_{i \rightarrow j}^t$ :

$$\Phi_i^{t+\Delta t}(X_i | X_{\partial i}) = \Phi_i^t(X_i | X_{\partial i}) [1 - r_i(\sigma_i(t), \sigma_{\partial i}(t)) \Delta t] + o(\Delta t). \quad (2.29)$$

The expansion of the right-hand side has, as in subsection 2.1.1, two contributions of order  $\Delta t$ . One of them is obtained by considering that none of the histories  $X_k$ , with  $k \in \partial i \setminus j$ , have a jump in the interval  $[t, t + \Delta t]$ :

$$I_0 + I_1 = \sum_{X_{\partial i \setminus j}}^{[t_0, t]} F(X_i, X_{\partial i}, t) \left\{ 1 - \left[ \sum_k \lambda_{k \rightarrow (ki)}(t) + r_i(t) \right] \Delta t \right\}. \quad (2.30)$$

The other one considers all the sets of histories  $X_{\partial i \setminus j}$  where one and only one of the histories  $X_k$  has a jump in the interval  $[t, \Delta t]$ :

$$I_2 = \sum_k \sum_{X_{\partial i \setminus j}}^{[t_0, t]} F(X_i, X_{\partial i}, t) \lambda_{k \rightarrow (ki)}(t) \Delta t. \quad (2.31)$$

Then:

$$\sum_{X_{\partial i \setminus j}}^{[t_0, t+\Delta t]} F(X_i, X_{\partial i}, t + \Delta t) = I_0 + I_1 + I_2 + o(\Delta t). \quad (2.32)$$

In Eq. (2.32), the terms that have  $\lambda$  factors cancel out. By combining Eq. (2.28) with Eq. (2.32):

$$\lambda_{i \rightarrow (ij)}[X_i, X_j, t] \mu_{i \rightarrow (ij)}^t(X_i | X_j) = \sum_{X_{\partial i \setminus j}}^{[t_0, t]} r_i[\sigma_i(t), \sigma_j(t), \sigma_{\partial i \setminus j}(t)] F(X_i, X_{\partial i}, t). \quad (2.33)$$

The right-hand side of the previous equation can be summed over all the histories of the spins in  $k \in \partial i \setminus j$  while keeping their final configuration  $\sigma_{\partial i \setminus j}$  fixed. The result is:

$$\lambda_{i \rightarrow (ij)}(X_i, X_j, t) \mu_{i \rightarrow (ij)}^t(X_i | X_j) = \sum_{\sigma_{\partial i \setminus j}} r_i(\sigma_i(t), \sigma_j(t), \sigma_{\partial i \setminus j}) p^t(\sigma_{\partial i \setminus j}, X_i \| X_j), \quad (2.34)$$

where  $p$  was introduced as a marginal of  $F$ . Note that the quantity  $p^t(\sigma_{\partial i \setminus j}, X_i \| X_j)$  is a conditional probability where the history  $X_j$  is fixed, corresponding to a marginal of  $F$ . Therefore, it is a cavity conditional probability. To mark the difference from a usual conditional probability, the separator “ $\|$ ” is used instead of simply “ $|$ ”.

The relation (2.34) is fundamental in obtaining the CME [35]. The following section briefly indicates how to proceed from here.

## 2.3 Cavity Master Equation

The contents of Sections 2.1 and 2.2 are sufficient to understand the essential aspects of the CME derivation as presented in Ref. [35]. This allows us to focus more on discussing the involved approximations in this section.

The CME introduces differential equations for the auxiliary probabilities:

$$p^t(\sigma_i \parallel X_j) = \sum_{X_i | \sigma_i}^{[t_0, t]} \mu_{i \rightarrow (ij)}^t(X_i | X_j). \quad (2.35)$$

Each  $p^t(\sigma_i \parallel X_j)$  is interpreted as the probability of finding variable  $i$  with the value  $\sigma_i$  at time  $t$ , given that the history of  $j$  is fixed and equal to  $X_j$ .

Both sides of Eq. (2.35) can be differentiated again, applying Eq. (2.34), to obtain an exact differential equation for graphs that are locally trees:

$$\begin{aligned} \frac{dp^t(\sigma_i \parallel X_j)}{dt} = & - \sum_{\sigma_{\partial i \setminus j}} r_i(\sigma_i, \sigma_j(t), \sigma_{\partial i \setminus j}) p^t(\sigma_{\partial i \setminus j}, \sigma_i \parallel X_j) + \\ & + \sum_{\sigma_{\partial i \setminus j}} r_i(-\sigma_i, \sigma_j(t), \sigma_{\partial i \setminus j}) p^t(\sigma_{\partial i \setminus j}, -\sigma_i \parallel X_j). \end{aligned} \quad (2.36)$$

This equation has two problems: i) it is not closed, as we would need equations for the probabilities appearing on the right-hand side, and ii) it contains functions of the histories  $X_j$ , which are highly dimensional objects. To resolve both issues, Ref. [35] applies a set of approximations discussed below.

First, we use the definition of conditional probability in the same space where  $X_j$  is fixed to write  $p^t(\sigma_{\partial i \setminus j}, \sigma_i \parallel X_j) = p^t(\sigma_{\partial i \setminus j} \mid \sigma_i \parallel X_j) p^t(\sigma_i \parallel X_j)$ . Assuming that spatial correlations are low, in the first conditional probability  $p^t(\sigma_{\partial i \setminus j} \mid \sigma_i \parallel X_j)$  we can neglect the influence of the trajectory  $X_j$ . In effect, node  $j$  is at a distance  $d = 2$  from the nodes in  $\partial i \setminus j$ . It is expected to be more important to maintain the information about node  $i$ , which is a neighbor of all nodes in  $\partial i \setminus j$ . The expression then becomes:

$$p^t(\sigma_{\partial i \setminus j}, \sigma_i \parallel X_j) \approx p^t(\sigma_{\partial i \setminus j} \mid \sigma_i) p^t(\sigma_i \parallel X_j). \quad (2.37)$$

In graphs that are locally trees, if we remove node  $i$ , the variables in  $\partial i \setminus j$  will be virtually disconnected. We can then propose the following factorization:

$$p^t(\sigma_{\partial i \setminus j}, \sigma_i \parallel X_j) \approx \left[ \prod_{k \in \partial i \setminus j} p^t(\sigma_k \mid \sigma_i) \right] p^t(\sigma_i \parallel X_j). \quad (2.38)$$

However, this relationship, which is exact in equilibrium, is only an approximation for the dynamics. As the probability  $P^t(\sigma_i)$  changes over time, it is not sufficient to specify the state  $\sigma_i$  to disconnect the neighbors of  $i$ . Even if the value of  $\sigma_i$  is fixed at time  $t$ , the variables in  $\partial i \setminus j$  will be non-trivially correlated through their interaction

with  $\sigma_i$  at earlier times. The factorization of Eq. (2.38) will only be valid if temporal and spatial correlations are weak.

To conclude, a Markovian approximation of the dynamics will be made by substituting the trajectory  $X_j$  with its final state  $\sigma_j(t) = \sigma_j$ . Again, it is assumed that spatial correlations are not significant. It is sufficient to know the final state of  $X_j$  to determine the conditional probability distribution  $p^t(\sigma_i \parallel \sigma_j)$ . Thus:

$$p^t(\sigma_{\partial i \setminus j}, \sigma_i \parallel X_j) \approx \left[ \prod_{k \in \partial i \setminus j} p^t(\sigma_k \mid \sigma_i) \right] p^t(\sigma_i \parallel \sigma_j). \quad (2.39)$$

Assuming further that  $p^t(\sigma_k \mid \sigma_i) \equiv p^t(\sigma_k \parallel \sigma_i)$ , the cavity master equation is finally written as:

$$\begin{aligned} \frac{dp^t(\sigma_i \parallel \sigma_j)}{dt} = & - \sum_{\sigma_{\partial i \setminus j}} r_i(\sigma_i, \sigma_j, \sigma_{\partial i \setminus j}) \left[ \prod_{k \in \partial i \setminus j} p^t(\sigma_k \mid \sigma_i) \right] p^t(\sigma_i \parallel \sigma_j) + \\ & + \sum_{\sigma_{\partial i \setminus j}} r_i(-\sigma_i, \sigma_j, \sigma_{\partial i \setminus j}) \left[ \prod_{k \in \partial i \setminus j} p^t(\sigma_k \mid -\sigma_i) \right] p^t(-\sigma_i \parallel \sigma_j). \end{aligned} \quad (2.40)$$

By making a similar factorization (see Fig. (2.1)) in the local master equation (2.2), a closed system of equations for the continuous-time dynamics of discrete variables in random graphs can finally be obtained. The joint probability  $P^t(\sigma_i, \sigma_{\partial i})$  should be written in terms of  $P^t(\sigma_i)$  and the auxiliary quantities  $p^t(\sigma_k \parallel \sigma_i)$ . Under the same assumptions used in the path to Eq. (2.39) and for graphs that are locally trees, the following is proposed:

$$P(\sigma_i, \sigma_{\partial i}) \approx \left[ \prod_{k \in \partial i} p^t(\sigma_k \mid \sigma_i) \right] P^t(\sigma_i). \quad (2.41)$$

Eq. (2.40) complements the differential equation for the local probability  $P^t(\sigma_i)$ :

$$\begin{aligned} \frac{dP^t(\sigma_i)}{dt} = & - \sum_{\sigma_{\partial i \setminus j}} r_i(\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i} p^t(\sigma_k \mid \sigma_i) \right] P^t(\sigma_i) + \\ & + \sum_{\sigma_{\partial i \setminus j}} r_i(-\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i} p^t(\sigma_k \mid -\sigma_i) \right] P^t(-\sigma_i). \end{aligned} \quad (2.42)$$

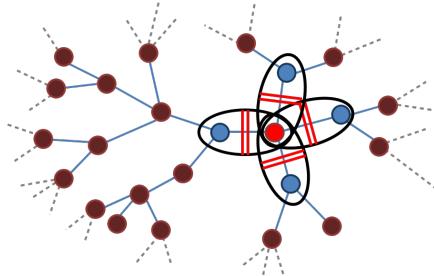


Figure 2.1: Illustration of the factorization done in Eq. (2.42)

Numerical integration of Eqs. (2.40) and (2.42) allows us to obtain the temporal evolution of observables such as magnetization and energy [35].

This chapter concludes with the presentation of the mathematical methodology to obtain master equations from TRPP. This will be particularly important for understanding Chapter 4. Additionally, a specific closure of the master equation: the CME was introduced.

In the second part of this thesis, the CME is applied to different problems, including K-SAT. The equilibrium probability distribution of the CME is compared with the corresponding one from the equilibrium cavity method. Moreover, average equations for the CME over families of random graphs are developed, and a hierarchical system of cavity master equations is constructed.

## Part II

# Cavity Master Equation

## Chapter 3

# Equilibrium and average case

As presented in Chapter 2, the Cavity Master Equation (CME) is formulated to predict the behavior of systems on specific graphs. Once the structure of the interactions and the form of the transition probabilities per unit time  $r_i$  are defined, the differential equations can be integrated numerically to obtain the time dependence of the observables. Therefore, it is possible to study dynamics completely out of equilibrium, such as those of the focused algorithms described in Chapter 1.

However, before venturing into largely unexplored terrain, it is useful to return to models over which there is greater control. Describing the relaxation towards equilibrium of dynamics that fulfill detailed balance allows connecting the results of the new technique with previous knowledge. The first part of this chapter demonstrates that the stationary state of the CME coincides with the solution of the Belief Propagation (BP) equations in models where thermodynamic equilibrium exists. Sections 3.1 and 3.2 introduce the equilibrium cavity method and show its relationship with the CME, respectively.

The second part of the chapter presents different methods for averaging the CME over an entire family of graphs. The average equations simplify the description of the dynamics, reduce the computational cost of the numerical solution, and are easily comparable with results from the literature. Following the idea of confronting known results, it is shown that the CME reduces to the exact solution of the ferromagnetic Ising model with asymmetric interactions. This is a particular case of the model for asymmetric neural networks solved in Ref. [66].

### 3.1 Equilibrium cavity method

When random graphs are locally trees and there are no long-range correlations in the system, it is possible to write the distribution of global probabilities as a product of local probabilities. In this section, the Belief Propagation (BP) equations involved in this factorization of the global probability are written.

It is useful to present the equations in their most general case, when working with an interaction hypergraph with variable nodes  $i, j, k, \dots$  and factor nodes  $a, b, \dots$  (see

Section 1.1). Two auxiliary probabilities are defined,  $\nu_{i \rightarrow a}(\sigma_i)$  and  $\eta_{a \rightarrow i}(\sigma_i)$ , which satisfy the equations:

$$\nu_{i \rightarrow a}(\sigma_i) = \frac{1}{Z_{i \rightarrow a}} \prod_{b \in \partial i \setminus a} \eta_{b \rightarrow i}(\sigma_i) \quad (3.1)$$

$$\eta_{a \rightarrow i}(\sigma_i) = \frac{1}{Z_{a \rightarrow i}} \sum_{\sigma_{a \setminus i}} e^{\beta J \sigma_i \prod_{k \in a \setminus i} \sigma_k} \prod_{k \in a \setminus i} \nu_{k \rightarrow a}(\sigma_k), \quad (3.2)$$

where

- $Z_{i \rightarrow a}$  and  $Z_{a \rightarrow i}$  are normalization constants.
- $\partial i$  symbolizes the set of variable nodes that are in plaquettes to which  $i$  belongs.
- $b \subset \partial i \setminus a$  are all the plaquettes to which  $i$  belongs, except for  $a$ .
- $a \setminus i$  represents all the variable nodes that are in plaque  $a$ , excluding node  $i$ .

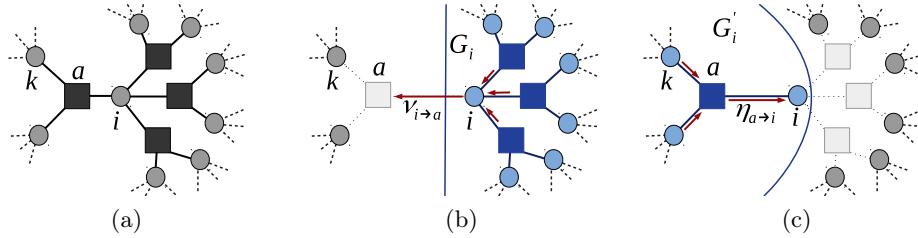


Figure 3.1: Illustration of the meaning of messages in the cavity method. (a) Original graph (b) Auxiliary graph for the message  $\nu_{i \rightarrow a}(\sigma_i)$  (c) Auxiliary graph for the message  $\eta_{a \rightarrow i}(\sigma_i)$ .

The interpretation of cavity probabilities  $\nu_{i \rightarrow a}(\sigma_i)$  and  $\eta_{a \rightarrow i}(\sigma_i)$  is closely related to the local properties of the mentioned random graphs. When the system size  $N$  is large, the surroundings of site  $i$  will look like in Fig. (3.1a). If factor node  $a$  is removed, variable nodes  $i$  and  $k$  can only be connected by a cycle which, with high probability, is long. Assuming that the correlation length in the graph is finite, from the viewpoint of  $i$ , the graph would be separated into  $p$  disconnected subgraphs, only one of which fulfills  $i \in G_i$  (see Fig. (3.1b)). The magnitude  $\nu_{i \rightarrow a}(\sigma_i)$  is precisely the probability associated with site  $i$  in  $G_i$ . Fig. (3.1c) shows a different modification where the cavity probability  $\eta_{a \rightarrow i}(\sigma_i)$  is associated with site  $i$  in the new graph  $G'_i$ . The ability to reconstruct the probability distribution of the entire random hypergraph by moving through variable nodes and factor nodes allows writing equations (3.1) and (3.2).

Following the same logic as in Fig. (3.1c), it is possible to reconstruct the local probability of site  $i$  and the probability of nodes in a plaquette:

$$p(\sigma_i) = \frac{1}{Z_i} \prod_{a \in \partial i} \eta_{a \rightarrow i}(\sigma_i) \quad (3.3)$$

$$p(\sigma_a) = \frac{1}{Z_a} e^{\beta J \prod_{i \in a} \sigma_i} \prod_{i \in a} \nu_{i \rightarrow a}(\sigma_i), \quad (3.4)$$

where  $Z_i$  and  $Z_a$  are normalization constants. With (3.3) and (3.4), the magnetization and energy of the entire system can be calculated.

Eqs. (3.1) and (3.2) can be iteratively solved in what is called the Belief Propagation (BP) algorithm. The quantities  $\nu_{i \rightarrow a}(\sigma_i)$  and  $\eta_{a \rightarrow i}(\sigma_i)$  are then called *messages*. For example, the dashed lines in Fig. (1.2a) result from the convergence of the BP algorithm if the initial messages are  $\nu_{i \rightarrow a}(\sigma_i) = \eta_{a \rightarrow i}(\sigma_i) = \delta_{\sigma_i, 1}$  for all pairs  $(i, a)$ . For  $T < T_{ms} \approx 1.63$ , there exists an ordered BP solution, i.e., a ferromagnetic solution of the cavity method. For  $T > T_{ms}$ , the algorithm converges towards a different solution where  $m \approx 0$ , always in agreement with the results of Monte Carlo simulations.

## 3.2 Equivalence with message passing

Although the Master Equation of Continuous Time (CME) results from differentiating the equations of dynamic message passing in continuous time, it is not clear how its steady state relates to the results of the original cavity method. In this section, it is proven that the solution of the CME in graphs that are locally trees corresponds exactly to the probability distribution determined by the BP algorithm. This establishes an important connection with a large number of results in the literature, given the relevance of BP in various fields such as Statistical Physics [2], error correction codes [88], and artificial intelligence [89].

The mathematical treatment is simpler when referring to the case of interactions between pairs of variables. In this case, the Hamiltonian is simply  $H(\vec{\sigma}) = \sum_{\langle ij \rangle} J_{ij} \sigma_i \sigma_j$ , where the sum goes over all pairs of variables connected in a random graph. Extending the procedures to the case of random hypergraphs is not difficult.

To ensure that the system reaches equilibrium in the limit for long times, the transition probabilities per unit time cannot take any form. Imposing the following detailed balance relations in the master equation leads the system to the Boltzmann-Gibbs distribution:

$$\frac{r_i(-\sigma_i, \sigma_{\partial i})}{r_i(\sigma_i, \sigma_{\partial i})} = \frac{P_{eq}(\sigma_i, \sigma_{\partial i})}{P_{eq}(-\sigma_i, \sigma_{\partial i})} = e^{2\beta \sigma_i \sum_{k \in \partial i} J_{ik} \sigma_k}. \quad (3.5)$$

In Eq. (3.5),  $P_{eq}(\sigma_i, \sigma_{\partial i})$  is the equilibrium probability distribution. That is, it is the limit of  $P^t(\sigma_i, \sigma_{\partial i})$  as  $t \rightarrow \infty$ .

### 3.2.1 CME's stationary state and equilibrium cavity method

It's important to remember that the Master Equation of Continuous Time (CME) deals with cavity conditional probabilities  $p(\sigma_i \parallel \sigma_j)$ , while in BP, messages  $\nu_{i \rightarrow (ij)}(\sigma_i)$  with a different nature appear. The magnitude  $p(\sigma_i \parallel \sigma_j)$  is the probability of finding site  $i$  in state  $\sigma_i$  if the spin at  $j$  is fixed to  $\sigma_j$  at time  $t$ . The message  $\nu_{i \rightarrow (ij)}(\sigma_i)$ , on the other hand, represents the local probability distribution for site  $i$  when the interaction with  $j$  is removed from the graph.

The BP equations for pair interactions in the absence of external fields can be written in a condensed form as:

$$\nu_{i \rightarrow (ij)}(\sigma_i) = \frac{1}{Z_{i \rightarrow (ij)}} \prod_{k \in \partial i \setminus j} \eta_{(ki) \rightarrow i}(\sigma_i) \quad (3.6)$$

$$\eta_{(ki) \rightarrow i}(J_{ik}, \sigma_i) = \frac{1}{Z_{(ki) \rightarrow i}} \sum_{\sigma_k} e^{\beta J_{ik} \sigma_i \sigma_k} \nu_{k \rightarrow (ki)}(\sigma_k). \quad (3.7)$$

For graphs that are locally trees, Eqs. (3.6) and (3.7) are asymptotically exact as  $N \rightarrow \infty$ .

One can also write the equilibrium distribution for the probabilities associated with a pair of variables as:

$$P(\sigma_i, \sigma_j) = \frac{1}{Z_{ij}} e^{\beta J_{ij} \sigma_i \sigma_j} \nu_{i \rightarrow (ij)}(\sigma_i) \nu_{j \rightarrow (ji)}(\sigma_j) \quad (3.8)$$

from which the exact conditional distributions can be immediately obtained:

$$P(\sigma_i | \sigma_j) = \frac{1}{\eta_{(ij) \rightarrow j}(J_{ij}, \sigma_j)} e^{\beta J_{ij} \sigma_i \sigma_j} \nu_{i \rightarrow (ij)}(\sigma_i). \quad (3.9)$$

To prove the equivalence of these BP expressions with the solution of the CME for long times, it's convenient to parameterize the CME distribution as <sup>1</sup>:

$$p(\sigma_i \parallel \sigma_j) = \frac{1}{\hat{\eta}_{(ij) \rightarrow j}(U_{ij}, \sigma_j)} e^{\beta U_{ij} \sigma_i \sigma_j} \hat{\nu}_{i \rightarrow (ij)}(\sigma_i). \quad (3.10)$$

Now we must show that  $U_{ij} = J_{ij}$  and that  $\hat{\nu}_{i \rightarrow (ij)}(\sigma_i)$  satisfies Eq. (3.6). This will imply simultaneously that the solution of the CME for conditional probabilities is exactly Eq. (3.9) and that it is equivalent to the cavity method in equilibrium.

From the detailed balance relation (3.5) and applying the joint probability factorization  $P(\sigma_i, \sigma_{\partial i}) = P(\sigma_i, \sigma_j) \prod_{k \in \partial i \setminus j} p(\sigma_k \parallel \sigma_i)$ , which is analogous to (2.41), we obtain:

$$\frac{p(\sigma_i \parallel \sigma_j)}{p(-\sigma_i \parallel \sigma_j)} = \sum_{\sigma_{\partial i \setminus j}} \frac{r_i(-\sigma_i, \sigma_{\partial i})}{r_i(\sigma_i, \sigma_{\partial i})} \prod_{k \in \partial i \setminus j} p(\sigma_k \parallel -\sigma_i). \quad (3.11)$$

---

<sup>1</sup>This is a general expression if  $U_{ij}$  and  $\hat{\nu}_{i \rightarrow j}(\sigma_i)$  are free parameters.

The next step is to insert Eq. (3.5) and the parameterization (3.10) for all conditional distributions into Eq. (3.11):

$$\begin{aligned} e^{2\beta U_{ij}\sigma_i\sigma_j} \frac{\hat{\nu}_{i\rightarrow(ij)}(\sigma_i)}{\hat{\nu}_{i\rightarrow(ij)}(-\sigma_i)} &= e^{2\beta J_{ij}\sigma_i\sigma_j} \prod_{k\in\partial i\setminus j} \frac{\hat{\eta}_{(ki)\rightarrow i}(2J_{ik}-U_{ik}, \sigma_i)}{\hat{\eta}_{(ki)\rightarrow i}(U_{ik}, -\sigma_i)} \\ e^{2\beta(U_{ij}-J_{ij})\sigma_i\sigma_j} &= \frac{\hat{\nu}_{i\rightarrow(ij)}(-\sigma_i)}{\hat{\nu}_{i\rightarrow(ij)}(\sigma_i)} \prod_{k\in\partial i\setminus j} \frac{\hat{\eta}_{(ki)\rightarrow i}(2J_{ik}-U_{ik}, \sigma_i)}{\hat{\eta}_{(ki)\rightarrow i}(U_{ik}, -\sigma_i)}. \end{aligned} \quad (3.12)$$

The first thing to notice in Eq. (3.12) is that the right-hand side does not depend on  $\sigma_j$ , so the left-hand side must satisfy  $U_{ij} = J_{ij}$ . Naturally, this means that  $U_{ik} = J_{ik}$  can be written for any other pair  $(ik)$ . Then:

$$\frac{\hat{\nu}_{i\rightarrow(ij)}(\sigma_i)}{\hat{\nu}_{i\rightarrow(ij)}(-\sigma_i)} = \prod_{k\in\partial i\setminus j} \frac{\hat{\eta}_{(ki)\rightarrow i}(J_{ik}, \sigma_i)}{\hat{\eta}_{(ki)\rightarrow i}(J_{ik}, -\sigma_i)}. \quad (3.13)$$

$$\hat{\nu}_{i\rightarrow(ij)}(\sigma_i) \propto \prod_{k\in\partial i\setminus j} \hat{\eta}_{(ki)\rightarrow i}(J_{ik}, \sigma_i). \quad (3.14)$$

Eq. (3.14) directly implies that the trial distributions  $\hat{\nu}$  introduced in the parameterization (3.10) satisfy the equilibrium equations of the cavity method (3.6). This proves that the stationary distributions of the CME are equivalent to the solution of the BP equations. Thus, in graphs that are locally trees, the CME reproduces the exact equilibrium solution. In Appendix C, it's shown how fixed-point equations can be directly obtained for the cavity conditional probabilities  $p(\sigma_i \mid \sigma_j)$  and how their numerical solution yields the same results as the BP algorithm.

### 3.3 Average case

In this section, equations representing the average behavior of the CME are introduced for three different models on random graphs: the asymmetric ferromagnetic model, the symmetric ferromagnetic model, and the Susceptible-Infectious-Susceptible (SIS) model for epidemic propagation. Among them, only in the symmetric ferromagnetic model is there relaxation towards a stationary state where the Boltzmann-Gibbs distribution is fulfilled. In the other two, the dynamics are completely out of equilibrium.

#### 3.3.1 Fully asymmetric ferromagnet

In this section, equations representing the average behavior of the CME are introduced for three different models on random graphs: the asymmetric ferromagnetic model, the symmetric ferromagnetic model, and the Susceptible-Infectious-Susceptible (SIS) model for epidemic propagation. Among them, only in the symmetric ferromagnetic model is there relaxation towards a stationary state where the Boltzmann-Gibbs distribution is fulfilled. In the other two, the dynamics are completely out of equilibrium.

Similarly to Section 3.2, where a connection with the well-known cavity method in equilibrium was established, it's useful to show simplified versions of the CME that can be compared with other known results. In the next subsection, the average behavior of the CME is obtained for the dynamics of a system with ferromagnetic interactions that are completely asymmetric. The equations for the temporal evolution of magnetization coincide with the exact result from Ref. [66].

In this case, there's no equilibrium description of the system. Instead, the following transition probabilities per unit time are defined:

$$r_i(\sigma_i, \sigma_{\partial i}) = \frac{1}{2}(1 - \sigma_i \tanh(\beta \sum_{k \in \partial i} J_{ki} \sigma_k)), \quad (3.15)$$

where it's possible to have  $J_{ik} \neq J_{ki}$ . That is, if the spin at node  $i$  takes the value  $\sigma_i = 1$ , then the spin  $k$  will be under the influence of a local field  $h_{ik} = J_{ik}$  due to its interaction with  $i$ . At the same time, if  $\sigma_k = 1$ , the spin at site  $i$  will perceive a field  $h_{ki} = J_{ki}$  which is generally different. The interactions are thus *asymmetric*.

In the model to be studied, the couplings  $J_{ki}$  follow the distribution:

$$Q(J_{ki}) = \frac{\kappa}{N-1} \delta(J_{ki} - 1) + \left(1 - \frac{\kappa}{N-1}\right) \delta(J_{ki}), \quad (3.16)$$

where  $\kappa$  is the average connectivity of a node in the interaction graph. That is, on average there will be  $\kappa$  variables for which  $J_{ki} = 1$ . The choice of  $J_{ki}$  is independent of the choice of  $J_{ik}$ . Therefore, for finite  $\kappa$  and in the limit  $N \gg 1$ , the probability of having  $J_{ik} = J_{ki} = 1$  tends to zero.

The key point in the following deductions is that, due to the asymmetric interactions, the total local field  $h_i = \sum_{k \in \partial i} J_{ki} \sigma_k$  follows a distribution that is independent of the value of the spin  $\sigma_i$ . This means that in Eq. (2.42), since  $J_{ki} = 1$ , it is almost certain that  $J_{ik} = 0$ . Therefore, the cavity conditional probabilities  $p(\sigma_i \parallel \sigma_j)$  of the CME actually do not depend on  $\sigma_j$ .

Applying the operator  $\sum_{\sigma_i} \sigma_i [\cdot]$  to Eq. (2.42) yields:

$$\frac{dm_i(t)}{dt} = -m_i(t) + \sum_{\sigma'_i} P(\sigma'_i) \sum_{\sigma_{\partial i^-}} \tanh(\beta \sum_{k \in \partial i^-} J_{ki} \sigma_k) \prod_{k \in \partial i^-} p(\sigma_k), \quad (3.17)$$

where each  $p(\sigma_k \parallel \sigma'_i)$  has been replaced by  $p(\sigma_k)$  since there is no real dependence on the variable  $\sigma'_i$ . Only nodes  $k$  such that  $J_{ki} \neq 0$  appear in the equation. This new neighborhood is represented by the symbol  $\partial i^-$ . Additionally, the local magnetization  $m_i(t) \equiv \sum_{\sigma_i} \sigma_i P(\sigma_i)$  is defined. If we now introduce the *cavity magnetization*  $\nu_i(t) \equiv \sum_{\sigma_i} \sigma_i p(\sigma_i)$  and explicitly sum over  $\sigma_i$ , we arrive at:

$$\frac{dm_i(t)}{dt} = -m_i(t) + \sum_{\sigma_{\partial i^-}} \tanh(\beta \sum_{k \in \partial i^-} J_{ki} \sigma_k) \prod_{k \in \partial i^-} \left( \frac{1 + \sigma_k \nu_k}{2} \right). \quad (3.18)$$

Similarly, from the CME (see Eq. (2.40)), differential equations for all  $\nu_i(t)$  can be obtained:

$$\frac{d\nu_i(t)}{dt} = -\nu_i(t) + \sum_{\sigma_{\partial i^-}} \tanh(\beta \sum_{k \in \partial i^-} J_{ki} \sigma_k) \prod_{k \in \partial i^-} \left( \frac{1 + \sigma_k \nu_k}{2} \right). \quad (3.19)$$

It can be seen that Eqs. (3.18) and (3.19) are equal. Now  $p(\sigma_i \| \sigma_j)$  no longer depends on  $\sigma_j$  and only the spins in the set  $\partial i^-$  matter. One can simply write the following equation for magnetization:

$$\frac{dm_i(t)}{dt} = -m_i(t) + \sum_{\sigma_{\partial i^-}} \tanh(\beta \sum_{k \in \partial i^-} J_{ki} \sigma_k) \prod_{k \in \partial i^-} \left( \frac{1 + \sigma_k m_k}{2} \right). \quad (3.20)$$

At this point, we still have an equation written for the node  $i$  of a specific random graph. To average over graphs belonging to a specific family (such as ER), some concepts need to be clarified. First, instead of describing the temporal evolution of a single  $m_i$ , the average dynamics of all sites that have exactly  $c^-$  neighbors such that  $J_{ki} \neq 0$  will be described. The parameter  $c^-$  is called the *in-degree* connectivity, and from now on, the term *in-connectivity* will be used. This yields:

$$\frac{dm_{c^-}}{dt} = -m_{c^-} + \sum_{\{\sigma_k\}} \tanh(\beta \sum_{k=1}^{c^-} J_{ki} \sigma_k) \prod_{k=1}^{c^-} \sum_{c_k^- = 0}^{\infty} Q(c_k^-) \left( \frac{1 + \sigma_k m_{c_k^-}}{2} \right). \quad (3.21)$$

Here, a new quantity  $m_{c^-}$  was introduced, which is the average magnetization corresponding to sites with in-connectivity  $c^-$ . To obtain its dynamics, one must sum over all possible equations for a node with these characteristics, which reduces to summing over all possible  $\{c_k^-\}$  for the node's neighbors. In Eq. (3.21),  $Q(c_k^-)$  was written to symbolize the probability distribution of the in-connectivities.

If Eq. (3.21) is averaged again over the family of graphs given by the distribution (3.16) while keeping  $\kappa$  fixed, we have:

$$\frac{d\hat{m}(t)}{dt} = -\hat{m}(t) + \sum_{c=0}^{\infty} Q(c) \sum_{n=0}^c \binom{c}{n} \left( \frac{1 + \hat{m}}{2} \right)^n \left( \frac{1 - \hat{m}}{2} \right)^{c-n} \tanh(\beta(2n - c)), \quad (3.22)$$

where  $Q(c) = e^{-\kappa} \kappa^c / c!$  is the Poisson distribution followed by the node connectivities in ER graphs. The magnetization  $\hat{m}$  is the average of  $m_{c^-}$  taken over all possible in-connectivities of site  $i$ . It should be noted that obtaining Eq. (3.22) has been considerably facilitated by the independence of the local field  $h_i$  and the spin  $\sigma_i$ . It was possible to smoothly average first over the neighbors' connectivities and then over the node's connectivity. Averaging over the family of graphs thus became an average of independent variables.

Equation (3.22) is a special case of the solution obtained by Derrida *et al.* for an asymmetric neural network in Ref. [66]. Fortunately, we can proceed further with the calculation (see Appendix D) to write:

$$\frac{d\hat{m}(t)}{dt} = -\hat{m}(t) + e^{-\kappa} S(\hat{m}, \beta, \kappa, 0), \quad (3.23)$$

where

$$S(\hat{m}, \beta, q, \sigma) = \sum_{n=-\infty}^{\infty} \tanh(\beta J(n + \sigma)) \left( \frac{1 + \hat{m}}{1 - \hat{m}} \right)^{n/2} I_n(q \sqrt{1 - \hat{m}^2}). \quad (3.24)$$

In this point the modified Bessel function of the first kind, or Infeld function  $I_\nu(x)$ , has been used.

This means that the probability of finding a local field  $\hat{h} = n$  is:

$$D(\hat{h}=n) = \left( \frac{1 + \hat{m}}{1 - \hat{m}} \right)^{n/2} I_n(q \sqrt{1 - \hat{m}^2}). \quad (3.25)$$

Then, the temporal dependence of the magnetization can be obtained by integrating Eq. (3.23). In Fig. (3.2), the results of the average equations are compared with the numerical average of the CME in several ER graphs with  $\kappa = 3$  and  $N = 1000$ . Also shown are the averages of the magnetization obtained with Monte Carlo simulations on the same graphs. The average equation (3.23) provides an accurate description of the dynamics of this observable. Differences with simulations are likely due to finite size effects rather than difficulties with the equations.

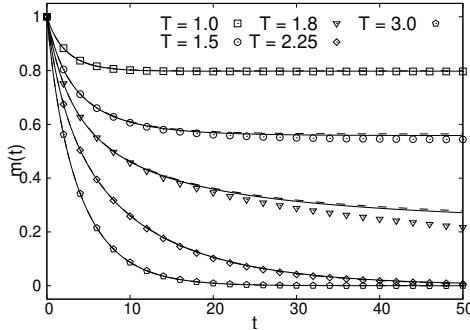


Figure 3.2: Temporal evolution of the magnetization obtained with average equations (dashed lines) and with numerical average of the CME (solid lines) and Monte Carlo (points) for the ferromagnetic model with asymmetric interactions. Averages were taken over  $s = 300$  Erdős-Rényi graphs with average connectivity  $\kappa = 3$  and size  $N = 1000$ . For each graph,  $n = 10000$  Monte Carlo histories were run. In all cases, a completely magnetized system was brought into contact with a thermal bath at fixed temperature  $T$  to study its relaxation to equilibrium.

### 3.3.2 Symmetric ferromagnet

Then, it's natural to proceed with the symmetric ferromagnetic model, also defined on ER graphs. Now,  $J_{ik} = J_{ki}$  for all node pairs  $(ik)$  in the system. This adds complexity to the calculation since it's no longer possible to say that the local field on a spin is independent of the value of that spin. The system does have an associated

Hamiltonian and an equilibrium state following the Boltzmann-Gibbs distribution. The Hamiltonian is  $H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j$ , where we take  $J = 1$  for simplicity.

As mentioned earlier, in an ER graph, the probability of finding a node with exactly  $c$  neighbors is  $Q(c) = e^{-\kappa} \kappa^c / c!$ , where  $\kappa$  is the average connectivity. To average the CME in the symmetric ferromagnetic model, it is necessary to remember that the cavity conditional probabilities  $p(\sigma_i \parallel \sigma_j)$  are defined over interacting spin pairs. It is convenient to introduce the probability  $q(\gamma)$  that  $\sigma_i$  has exactly a number  $\gamma + 1$  of neighbors, counting  $j$ .

$$q(\gamma) = \frac{(\gamma + 1)Q(c = \gamma + 1)}{\sum_{l=0}^{\infty} l Q(c = l)} = \frac{e^{-\kappa} \kappa^{\gamma}}{\gamma!} = Q(\gamma). \quad (3.26)$$

This means that  $\langle \gamma \rangle = \langle c \rangle \equiv \kappa$ . Then, Eqs. (2.40) and (2.42) can be averaged using these probabilities for each site. This procedure can be reproduced numerically with a population dynamics, and the results justify continuing with the analytical calculation (see Appendix E).

If we write, as in subsection 3.3.1, the average dynamics of nodes that have the same connectivity  $c$ , we arrive at:

$$\begin{aligned} \dot{P}_c(\sigma) = & -\frac{1}{2} \{ P_c(\sigma) - P_c(-\sigma) \} + \frac{1}{2} \sigma \sum_{\{\sigma_k\}} \tanh(\beta J \sum_{k=1}^c \sigma_k) \times \\ & \times \left\{ \left[ \prod_{k=1}^c \hat{p}(\sigma_k \parallel \sigma) \right] P_c(\sigma) + \left[ \prod_{k=1}^c \hat{p}(\sigma_k \parallel -\sigma) \right] P_c(-\sigma) \right\}. \end{aligned} \quad (3.27)$$

It should be noted that upon averaging, Eq. (3.27) no longer depends on the site  $i$ . The parameter  $c$  is a natural number, and  $\sigma = \pm 1$  is simply a name for a binary variable. The local probabilities  $P_c(\sigma)$  precisely represent the average of those appearing in the master equation (2.42) taken over all nodes with the same connectivity. Similarly, we can define the probability  $p_{\gamma_k}(\sigma_k \parallel \sigma)$ . In Eq. (3.27), we introduced the average cavity conditional probability from the probabilities defined for a fixed parameter  $\gamma_k$ :  $\hat{p}(\sigma_k \parallel \sigma) = \sum_{\gamma_k=0}^{\infty} q(\gamma_k) p_{\gamma_k}(\sigma_k \parallel \sigma)$ .

Taking the Glauber rule for dynamics (see Eq. 3.15), multiplying Eq. (3.27) by  $Q(c_i)$ , and summing over all values of  $c_i$  yields:

$$\begin{aligned} \dot{\hat{P}}(\sigma) = & -\frac{1}{2} \{ \hat{P}(\sigma) - \hat{P}(-\sigma) \} + \frac{1}{2} \sigma \sum_{c=0}^{\infty} Q(c) \sum_{\{\sigma_k\}} \tanh(\beta J \sum_{k=1}^c \sigma_k) \times \\ & \times \left\{ \left[ \prod_{k=1}^c \hat{p}(\sigma_k \parallel \sigma) \right] P_c(\sigma) + \left[ \prod_{k=1}^c \hat{p}(\sigma_k \parallel -\sigma) \right] P_c(-\sigma) \right\}. \end{aligned} \quad (3.28)$$

In this equation, we emphasize the difference between the local probabilities  $P_c(\sigma)$  and their averages  $\hat{P}(\sigma) = \sum_c Q(c) P_c(\sigma)$ . The magnitude  $\hat{P}(\sigma)$  represents the probability that any spin has value  $\sigma$ , regardless of the site and connectivity.

Similarly, starting from Eq. (2.40) and averaging over all values of  $\{\gamma_i, \gamma_{k \in \partial i \setminus j}\}$ , we get:

$$\begin{aligned} \dot{\hat{p}}(\sigma' \parallel \sigma) &= -\frac{1}{2}\{\hat{p}(\sigma' \parallel \sigma) - \hat{p}(-\sigma' \parallel \sigma)\} + \frac{\sigma'}{2} \sum_{\gamma=0}^{\infty} q(\gamma) \times \\ &\quad \times \sum_{\{\sigma_k\}} \tanh(\beta J \sum_{k=1}^{\gamma} \sigma_k + \beta J \sigma) \sum_{\sigma''} \left[ \prod_{k=1}^{\gamma} \hat{p}(\sigma_k \parallel \sigma'') p_{\gamma}(\sigma'' \parallel \sigma) \right]. \end{aligned} \quad (3.29)$$

In Appendix E, a numerical scheme is presented to test the validity of Eqs. (3.28) and (3.29). Next, we must work with the infinite sums in these equations. As in subsection 3.3.1, we will rewrite Eq. (3.28) in terms of the average magnetization  $\hat{m} = \sum_{\sigma'} \sigma' \hat{P}(\sigma')$ :

$$\dot{\hat{m}} = -\hat{m} + \sum_{\sigma'} \sum_{c=0}^{\infty} Q(c) P_c(\sigma') \sum_{\{\sigma_k\}} \tanh(\beta J \sum_{k=1}^c \sigma_k) \prod_{k=1}^c \hat{p}(\sigma_k \parallel \sigma'). \quad (3.30)$$

The infinite sum in Eq. (3.30) can be approximated using a relationship analogous to  $\int_{m=-1}^1 \tanh(m) P(m) dm \approx \tanh\left(\int_{m=-1}^1 m P(m) dm\right)$  to obtain:

$$\begin{aligned} \dot{\hat{m}} &= -\hat{m} + \sum_{\sigma'} \sum_{c=0}^{\infty} Q(c) P_c(\sigma') \tanh\left(\beta J \sum_{\{\sigma_k\}} \sum_{k=1}^c \sigma_k \prod_{l=1}^k \hat{p}(\sigma_l \parallel \sigma')\right) \\ \dot{\hat{m}} &= -\hat{m} + \sum_{\sigma'} \sum_{c=0}^{\infty} Q(c) P_c(\sigma') \tanh\left(\beta J \sum_{k=1}^c \tilde{m}(\sigma')\right) \\ \dot{\hat{m}} &= -\hat{m} + \sum_{\sigma'} \sum_{c=0}^{\infty} Q(c) P_c(\sigma') \tanh(\beta J c \tilde{m}(\sigma')), \end{aligned} \quad (3.31)$$

where we define the cavity magnetization  $\tilde{m}(\sigma') \equiv \sum_{\sigma_k} \hat{p}(\sigma_k \parallel \sigma') \sigma_k$ . For this magnitude, we can also write a differential equation starting from Eq. (3.29):

$$\frac{d\tilde{m}(\sigma)}{dt} = -\tilde{m}(\sigma) + \sum_{\sigma'} \sum_{\gamma=0}^{\infty} q(\gamma) p_{\gamma}(\sigma' \parallel \sigma) \tanh[\beta J \gamma \tilde{m}(\sigma') + \beta J \sigma]. \quad (3.32)$$

Eqs. (3.31) and (3.32) depend on two types of variables: on one hand, the averages  $\hat{m}$  and  $\tilde{m}(\sigma)$ , and on the other hand, the probabilities  $P_c(\sigma)$  and  $p_{\gamma}(\sigma' \parallel \sigma)$  defined for a given connectivity. To write the equations only in terms of the average magnitudes, one last approximation must be made. In both equations,  $P_c(\sigma)$  and  $p_{\gamma}(\sigma' \parallel \sigma)$  will be replaced by their respective averages  $\hat{P}(\sigma)$  and  $\hat{p}(\sigma' \parallel \sigma)$  to obtain:

$$\begin{aligned} \dot{\hat{m}} &= -\hat{m} + \sum_{\sigma'} \hat{P}(\sigma') \sum_{c=0}^{\infty} Q(c) \tanh[\beta J c \tilde{m}(\sigma')] \\ \dot{\hat{m}} &= -\hat{m} + \sum_{\sigma'} \frac{1 + \sigma' \hat{m}}{2} \sum_{c=0}^{\infty} Q(c) \tanh[\beta J c \tilde{m}(\sigma')] \end{aligned} \quad (3.33)$$

and

$$\frac{d\tilde{m}(\sigma)}{dt} = -\tilde{m}(\sigma) + \sum_{\sigma'} \frac{1 + \sigma' \tilde{m}(\sigma)}{2} \sum_{\gamma=0}^{\infty} q(\gamma) \tanh [\beta J \gamma \tilde{m}(\sigma') + \beta J \sigma]. \quad (3.34)$$

In the Eqs. (3.33) and (3.34), the relationships between probabilities and their corresponding magnetizations  $\hat{P}(\sigma') = (1 + \sigma' \hat{m})/2$  and  $\hat{p}(\sigma' \parallel \sigma) = (1 + \sigma' \tilde{m}(\sigma))/2$  have been employed. As can be observed, the information about the graph family is encoded in the distributions  $Q(c)$  and  $q(\gamma)$  whose form has not been specified. The same Eqs. (3.33) and (3.34) can be used for all graph families where the CME is valid, i.e., graphs that are locally tree-like.

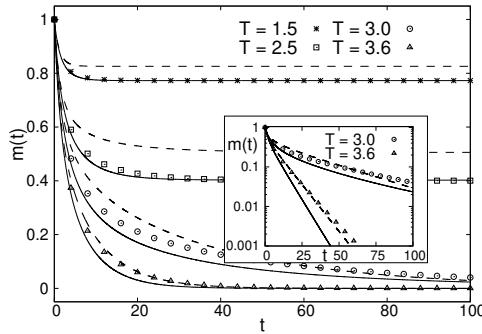


Figure 3.3: Result of the average equations for the CME (dashed lines) in Erdős-Rényi graphs with average connectivity  $\kappa = 3$  ( $T_c \approx 2.89$ ). It is compared with the numerical average of the CME (solid lines) and Monte Carlo simulations (points) on  $s = 350$  graphs of size  $N = 4000$ . The sums of the Eqs. (3.33) and (3.34) were calculated approximately using the first 50 terms. In all cases, a completely magnetized system in contact with a thermal bath at fixed temperature  $T$  was used as a starting point. For each graph,  $n = 10^5$  Monte Carlo histories were run. The inset graph represents the results for  $T = 3.0$  and  $T = 3.6$  in logarithmic scale for the  $y$ -axis.

To numerically solve the average equations, it is necessary to truncate the infinite sums for some  $c_{\max}$  and  $\gamma_{\max}$ . Fig. (3.3) compares the results with the average of the CME and Monte Carlo simulations on several ER graphs with average connectivity  $\kappa = 3$ . As expected from the works in Refs. [35,36], the numerical average of the CME accurately describes the average behavior of the Monte Carlo simulations, even near the order-disorder transition temperature  $T_c(\kappa = 3) \approx 2.89$  [90]. This technique can predict the relaxation of magnetization and its steady state.

Fig. (3.3) shows that the average equations accurately describe the dynamics in the high-temperature region ( $T > T_c$ ). On the other hand, for low temperatures, the equations overestimate the value of magnetization obtained from the simulations. The results improve as the temperature moves away from  $T_c$ . The approximation responsible for these problems is likely the one taken in Eq. (3.31). This relation

of the type  $\int_{m=-1}^1 \tanh(m) P(m) dm \approx \tanh\left(\int_{m=-1}^1 m P(m) dm\right)$  neglects long-range correlations that are important near  $T_c$ . A different method to obtain the average equations, which allows for more accurate results at the cost of losing simplicity in the equations, is presented in Appendix F.

### 3.3.3 SIS model for epidemics

---

The last of the models for which average equations will be found is the Susceptible-Infectious-Susceptible (SIS) model for the spread of epidemics, which is a relevant issue in various scenarios. Problems such as the dissemination of a disease in a population [27] or a computer virus in a network [26] interest numerous scientists. It is unnecessary to explain the attention brought by the COVID-19 pandemic towards the development of theoretical tools to understand the dynamics of epidemics [91].

Since the pioneering work of Kermack and McKendrick [92] on the Susceptible-Infectious-Recovered (SIR) model, various compartmental models have been generated. The idea is to divide the population into groups that are assumed to be homogeneous. All individuals in the same group interact in the same way with the rest. However, this does not mean that all individuals in a group are identical. The dynamics can be defined on a specific interaction network.

Again, the interactions are defined on a graph, and each node or individual has two possible states: susceptible  $S$  or infectious  $I$ . The rules that define the system's dynamics are:

- A node in the infectious state  $I$  transitions to the susceptible state with probability  $\mu$  per unit of time.
- A node in the infectious state  $I$  can pass the disease to a neighboring susceptible node  $S$  with probability  $\beta$  per unit of time.

It should be noted that the configuration where there are no infectious nodes is an absorbing state, so the dynamics do not satisfy detailed balance. The possible steady states do not follow the Boltzmann-Gibbs distribution, and the system is entirely out of equilibrium. From this perspective, this model is more interesting than the symmetric ferromagnetic model in subsection 3.3.2. On the other hand, it is more challenging to treat than the asymmetric ferromagnetic model in subsection 3.3.1. However, it is shown below how to obtain average equations in this case as well.

Returning to the previous formulations, if a binary variable  $\sigma_i = \{S, I\}$  is assigned to each node, the master equation for the evolution of the local probability  $P(\sigma_i)$  takes the form:

$$\frac{dP^t(\sigma_i)}{dt} = - \sum_{\sigma_{\partial i}} [r_i(\sigma_i, \sigma_{\partial i}) P^t(\sigma_i, \sigma_{\partial i}) - r_i(-\sigma_i, \sigma_{\partial i}) P^t(-\sigma_i, \sigma_{\partial i})]. \quad (3.35)$$

Here,  $r(\sigma_i = I, \sigma_{\partial i}) = \mu$  for all configurations of the neighbors, and  $r(\sigma_i = S, \sigma_{\partial i}) =$

$\beta \sum_{k \in \partial i} \delta_{\sigma_k, I}$ , where  $\delta_{\sigma_k, I}$  is a Kronecker delta that is 1 only if  $\sigma_k = I$ . It is not difficult to then write the corresponding CME for this model and thus find a closure for the master equation [38]. The equations can be considerably simplified:

$$\frac{dP_i}{dt} = -\mu P_i + \beta(1 - P_i) \sum_{k \in \partial i} p(k \rightarrow i) \quad (3.36)$$

$$\frac{dp(i \rightarrow j)}{dt} = -\mu p(i \rightarrow j) + (1 - p(i \rightarrow j))\beta \sum_{k \in \partial i \setminus j} p(k \rightarrow i), \quad (3.37)$$

where  $P_i \equiv P(\sigma_i = I)$  y  $p(i \rightarrow j) \equiv p(\sigma_i = I \parallel \sigma_j = S)$ .

---

As in the previous subsections, the average dynamics for nodes with the same connectivity can be written as:

$$\frac{dP_c}{dt} = -\mu P_c + \beta(1 - P_c) \sum_{k=1}^c \tilde{p}_\rightarrow \quad (3.38)$$

$$\frac{dp_\gamma(\rightarrow)}{dt} = -\mu p_\gamma(\rightarrow) + (1 - p_\gamma(\rightarrow))\beta \sum_{k=1}^\gamma \tilde{p}_\rightarrow, \quad (3.39)$$

where  $P_c$  is the probability of finding an infectious node with connectivity  $c$ . The quantity  $p_\gamma(\rightarrow)$  is the average of all conditional cavity probabilities with parameter  $\gamma$ .  $\tilde{p}_\rightarrow$  has also been defined as  $\sum_\gamma q(\gamma)p_\gamma(\rightarrow)$ .

By averaging Eqs. (3.38) and (3.39) over the distributions  $Q(c)$  and  $q(\gamma)$  of connectivities, we obtain:

$$\dot{\hat{P}} = -\mu \hat{P} + \beta \tilde{p}_\rightarrow \sum_{c=0}^{\infty} c Q(c)(1 - P_c) \quad (3.40)$$

$$\dot{\tilde{p}}_\rightarrow = -\mu \tilde{p}_\rightarrow + \beta \tilde{p}_\rightarrow \sum \gamma = 0^\infty \gamma q(\gamma)(1 - p_\gamma(\rightarrow)). \quad (3.41)$$

These equations are still not closed as the right-hand side members still depend on the probabilities  $P_c$  and  $p_\gamma(\rightarrow)$  which depend on the connectivities. It will be shown, however, that it is possible to obtain closed equations in ER graphs using an *ansatz* for the local probabilities.

It should be remembered that the distributions  $Q(c)$  and  $q(\gamma)$  depend on the average connectivity of the ER graphs:  $\kappa$ . The terms inside the sums of Eqs. (3.40) and (3.41) can be related to the derivatives of  $\hat{P}$  and  $\tilde{p}_\rightarrow$  with respect to the parameter  $\kappa$ :

$$\frac{\partial \hat{P}}{\partial \kappa} = \frac{\partial}{\partial \kappa} \left( \sum_{c=0}^{\infty} Q(c) P_c \right) = \frac{\partial}{\partial \kappa} \left( \sum c = 0^\infty \frac{e^{-\kappa} \kappa^c}{c!} P_c \right) \quad (3.42)$$

$$\frac{\partial \tilde{p}_\rightarrow}{\partial \kappa} = \frac{\partial}{\partial \kappa} \left( \sum_{\gamma=0}^{\infty} q(\gamma) p_\gamma(\rightarrow) \right) = \frac{\partial}{\partial \kappa} \left( \sum \gamma = 0^\infty \frac{e^{-\kappa} \kappa^\gamma}{\gamma!} p_\gamma(\rightarrow) \right). \quad (3.43)$$

Calculating the derivatives leads to:

$$\frac{\partial \hat{P}}{\partial \kappa} = -\hat{P} + \frac{1}{\kappa} \sum_{c=0}^{\infty} c Q(c) P_c \quad (3.44)$$

$$\frac{\partial \tilde{p}_{\rightarrow}}{\partial \kappa} = -\tilde{p}_{\rightarrow} + \frac{1}{\kappa} \sum_{\gamma=0}^{\infty} \gamma q(\gamma) p_{\gamma}(\rightarrow). \quad (3.45)$$

The sums on the right-hand side of Eqs. (3.44) and (3.45) also appear in Eqs. (3.40) and (3.41) for the time derivatives. Recalling that  $\kappa = \sum_c c Q(c)$ , Eq. (3.40) can be rewritten as follows:

$$\begin{aligned} \dot{\hat{P}} &= -\mu \hat{P} + \beta \tilde{p}_{\rightarrow} \sum_{c=0}^{\infty} c Q(c) (1 - P_c) \\ \dot{\hat{P}} &= -\mu \hat{P} + \beta \tilde{p}_{\rightarrow} \sum_{c=0}^{\infty} c Q(c) - \beta \tilde{p}_{\rightarrow} \sum_{c=0}^{\infty} c Q(c) P_c \\ \dot{\hat{P}} &= -\mu \hat{P} + \beta \kappa \tilde{p}_{\rightarrow} - \beta \kappa \tilde{p}_{\rightarrow} \left( \hat{P} + \frac{\partial \hat{P}}{\partial \kappa} \right). \end{aligned} \quad (3.46)$$

The equation for the probabilities  $\tilde{p}_{\rightarrow}$  can be deduced analogously:

$$\dot{\tilde{p}}_{\rightarrow} = -\mu \tilde{p}_{\rightarrow} + \beta \kappa \tilde{p}_{\rightarrow} - \beta \kappa \tilde{p}_{\rightarrow} \left( \tilde{p}_{\rightarrow} + \frac{\partial \tilde{p}_{\rightarrow}}{\partial \kappa} \right). \quad (3.47)$$

To numerically solve these equations, an *ansatz* is needed for the form of the averages  $\hat{P}$  and  $\tilde{p}_{\rightarrow}$  in terms of the average connectivity  $\kappa$  of the graph. The following proposal is inspired by the mean-field solutions for the dynamics of spin systems:

$$\hat{P}(t) = \frac{1}{2} [1 + \tanh(\beta \kappa \chi(t) - \mu)] \quad (3.48)$$

$$\tilde{p}_{\rightarrow}(t) = \frac{1}{2} [1 + \tanh(\beta \kappa \epsilon(t) - \mu)], \quad (3.49)$$

where  $\chi(t)$  and  $\epsilon(t)$  are time-dependent parameters that will be obtained self-consistently. For this, the derivative with respect to  $\kappa$  is calculated on both sides of Eqs. (3.48) and (3.49):

$$\frac{\partial \hat{P}}{\partial \kappa} = \frac{1}{2} [1 - \tanh^2(\beta \kappa \chi(t) - \mu)] \beta \chi(t) \quad (3.50)$$

$$\frac{\partial \tilde{p}_{\rightarrow}}{\partial \kappa} = \frac{1}{2} [1 - \tanh^2(\beta \kappa \epsilon(t) - \mu)] \beta \epsilon(t). \quad (3.51)$$

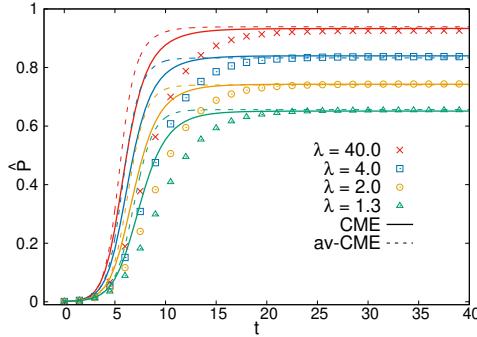


Figure 3.4: Comparison between the average equations (dashed lines) for the SIS model and the numerical average of the CME (solid lines) and Monte Carlo simulations (points) for several Erdős Rényi graphs with average connectivity  $\kappa = 3$  and size  $N = 1000$ . The figure shows the time evolution of the average probability of finding an infectious node  $\hat{P}$ . Different  $s = 20$  graphs were taken to calculate the numerical averages. For each graph,  $N = 10000$  epidemics were run, always starting with a single random infectious node. For comparison with the equations, Monte Carlo simulations where no infectious nodes exist for some time  $t < 40$  were discarded [93]. A fixed value for  $\beta$  ( $\beta = 0.4$ ) and different values for  $\mu$  were chosen. The legend shows the values of the fraction  $\lambda = \beta/\mu$ . As  $\lambda$  decreases, the infection probability  $\hat{P}$  of the steady state decreases.

Now the Eqs. (3.48) and (3.49) can be reused to eliminate  $\chi(t)$  and  $\epsilon(t)$  and obtain closed expressions for the derivatives:

$$\frac{\partial \hat{P}}{\partial \kappa} = \frac{1}{2\kappa} \left[ 1 - (2\hat{P} - 1)^2 \right] [\tanh^{-1}(2\hat{P} - 1) + \mu] \quad (3.52)$$

$$\frac{\partial \tilde{p}_\rightarrow}{\partial \kappa} = \frac{1}{2\kappa} \left[ 1 - (2\tilde{p}_\rightarrow - 1)^2 \right] [\tanh^{-1}(2\tilde{p}_\rightarrow - 1) + \mu]. \quad (3.53)$$

This allows for the numerical integration of Eqs. (3.46) and (3.47). Fig. (3.4) shows the results for the average infection probability ( $\hat{P}$ ). Also shown are the numerical averages of the original CME (see Eqs. (3.36) and (3.37)) and Monte Carlo simulations on several ER graphs with average connectivity  $\kappa = 3$ . Despite the simplicity of the chosen *ansatz*, the average equations describe the average epidemic dynamics with acceptable accuracy and predict the steady states. It is possible to write simple equations to reproduce the temporal evolution of the system with mean-field approximations that do not explicitly consider the local structure of the graphs.

## Partial conclusions

In Section 3.2, it was shown that the steady state of the CME corresponds to the cavity method in equilibrium if the transition probabilities that define the dynamics

respect detailed balance. Therefore, the long-time limit of the CME is the exact solution in graphs that are locally trees in a wide variety of cases. This result is significant for several scientific communities in which BP and the cavity method are part of the usual working techniques.

On the other hand, in Section 3.3, three systems of average equations were derived from the CME in random graphs. The exact solution for the asymmetric ferromagnetic model was recovered, and the dynamics of two more complex models were described: the symmetric ferromagnetic model and the SIS model for epidemics. In these cases, it is possible to make accurate predictions using only the average values of the probabilities.

In summary, the CME has been connected with results from the literature, and simpler equations have been derived that represent an average of the CME over families of random graphs. Both advances contribute to drawing attention to this new technique for dynamics and facilitate the path for other researchers interested in using it.

In the next chapter, corrections are introduced for some of the approximations made in the deduction of the CME. Numerical results of the new equations will be presented and compared with other theoretical methods.

## Chapter 4

# Hierarchical system of cavity master equations

In Chapters 2 and 3, the CME was introduced and this technique was connected with important results from the literature. However, the approximations leading to the CME cause deviations of the numerical results from the actual behavior of the system's dynamics. The procedure employed in the deduction of the equations does not allow for controlling the impact of these approximations or estimating their importance.

The problems of the CME are mainly related to the role of correlations in the system's dynamics. As explained in Section 2.3, to close the local master equation, the factorization  $P^t(\sigma_i, \sigma_{\partial i}) \approx P(\sigma_i) \prod_{k \in \partial i} p^t(\sigma_k \| \sigma_i)$  is applied, which neglects the possible presence of spatial correlations among the neighbors of node  $i$ . Even if the graph is locally a tree and conditioned on the value of  $\sigma_i$ , the temporal dependence of  $P^t(\sigma_{\partial i} | \sigma_i)$  introduces spatial correlations that the previous factorization does not capture.

This chapter presents a hierarchical system of cavity master equations where the CME is only the first level of approximation. Higher levels include spatial correlations at increasing distances in the graph, thus correcting one of the approximations from Section 2.3. The first section is dedicated to the deduction of the new equations, for which everything discussed in Chapter 2 is relevant. In subsections 4.2.1, 4.2.2, and 4.2.3, numerical results are shown and compared with those of other techniques from the literature: the Dynamic Independent Neighborhood Approximation (DINA) [21, 22] and the Pair-Based Mean-Field (PBMF) approximation for epidemics [32, 78].

### 4.1 Derivation

The original formulation of the CME is written for the cavity conditional probabilities defined over a pair of interacting variables:  $p^t(\sigma_i \| \sigma_j)$ . The procedure described in Section 2.3 is then extended to obtain a system of cavity master equations that can

be defined over any set of connected nodes in a random graph.

If a group of connected nodes is removed from a tree-like graph, the tree is divided into several subgraphs. In Fig. (4.1a), the subgraphs resulting from removing the nodes depicted in white are shaded in gray. As shown in subsection 2.2.1, the probability density corresponding to these latter nodes is:

$$Q^t(X_i, X_{\partial i}) = \Phi_i^t(X_i | X_{\partial i}) \prod_{k \in \partial i} \mu_{k \rightarrow (ki)}^t(X_k | X_i). \quad (4.1)$$

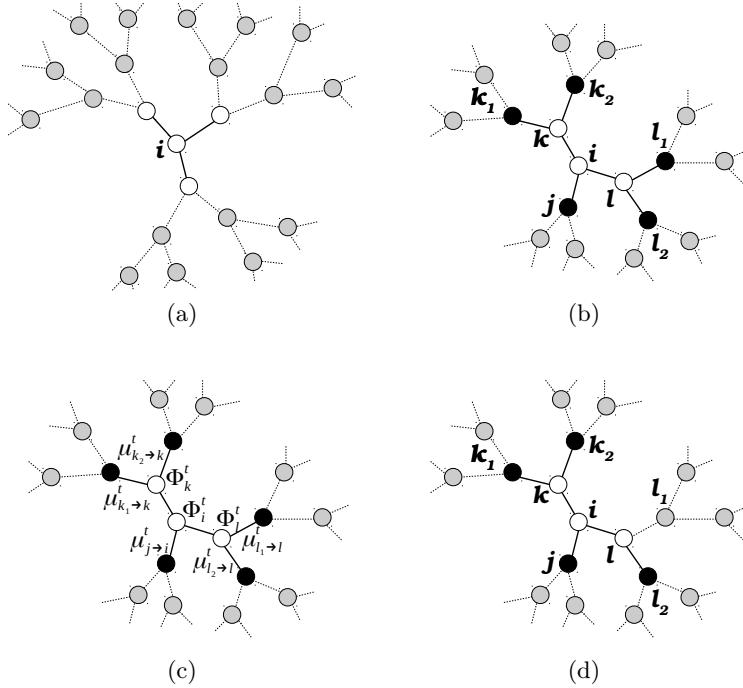


Figure 4.1: **a)** If the set of 4 connected nodes (in white) is removed, the tree-like graph is divided into 6 subgraphs (gray nodes). **b)** Example of internal (white) and external (black) nodes. **c)** The local weights and cavity messages can be used to write the probability density corresponding to the same set (see Eq. (4.2)). **d)** In this case, summing over the history  $X_{l_1}$  is enough to obtain the corresponding probability density (see Eq. (4.3)).

Similarly, the probability density  $Q$  can be written for any set  $S$  of connected nodes. It is useful to first define the concepts of *internal* node and *external* node. While the latter have at least one neighbor outside  $S$ , the former have all their neighbors within  $S$ . Fig. (4.1b) shows an example where the nodes of  $S$  are colored white or black. The nodes external to  $S$  are colored black and the internal nodes are white. If each external node has only one neighbor within  $S$ , it is straightforward to write the corresponding probability density.

For example, in Fig. (4.1c) it is shown that for the same connected set of nodes it is sufficient to write the product:

$$\begin{aligned} Q^t(X_i, X_{\partial i}, X_{\partial k \setminus i}, X_{\partial l \setminus i}) &= \Phi_i^t(X_i | X_{\partial i}) \left[ \prod_{m \in \partial i \setminus j} \Phi_m^t(X_m | X_{\partial m}) \right] \times \\ &\quad \times \mu_{j \rightarrow (ji)}^t(X_j | X_i) \left[ \prod_{n \in \partial k \setminus i} \mu_{n \rightarrow (nk)}^t(X_n | X_k) \right] \left[ \prod_{n \in \partial l \setminus i} \mu_{n \rightarrow (nl)}^t(X_n | X_l) \right]. \end{aligned} \quad (4.2)$$

For each internal node, a weight  $\Phi$  is multiplied, and for each external node, a weight  $\mu$  is multiplied. Both probabilities were defined in subsection 2.2.1. Connected sets like the one in Fig. (4.1d) do not require a different treatment. To obtain the corresponding probability density, it is sufficient to sum Eq. (4.2) over all histories of the node  $l_1$  in the figure.

$$Q^t(X_i, X_{\partial i}, X_{\partial k \setminus i}, X_{l_1}, X_{l_2}) = \sum_{X_{l_1}}^t Q^t(X_i, X_{\partial i}, X_{\partial k \setminus i}, X_{l_1}, X_{l_2}). \quad (4.3)$$

Following these rules, the probability density associated with the histories of any set of connected nodes can be written. Once an expression like (4.2) is obtained, the derivative of the instantaneous probability can be calculated:

$$\begin{aligned} P^t(\sigma_i, \sigma_{\partial i}, \sigma_{\partial k \setminus i}, \sigma_{\partial l \setminus i}) &= \\ &\sum_{X_i | \sigma_i} \sum_{X_{\partial i} | \sigma_{\partial i}} \sum_{X_{\partial k \setminus i} | \sigma_{\partial k \setminus i}} \sum_{X_{\partial l \setminus i} | \sigma_{\partial l \setminus i}} Q^t(X_i, X_{\partial i}, X_{\partial k \setminus i}, X_{\partial l \setminus i}). \end{aligned} \quad (4.4)$$

The procedure is similar to that followed in subsection 2.2.2 and can be applied to any probability  $P$  defined over a set of connected nodes such that  $\mathcal{I} = \{i_1, i_2, \dots, i_n\}$  are the internal nodes and  $\mathcal{O} = \{o_1, o_2, \dots, o_m\}$  are the external nodes. The differential equation obtained is:

$$\begin{aligned} \frac{dP^t(\vec{\sigma}_{\mathcal{O}}, \vec{\sigma}_{\mathcal{I}})}{dt} &= - \sum_{l=1}^n \left[ r_{i_l}(\sigma_{i_l}, \sigma_{\partial i_l}) P^t(\vec{\sigma}_{\mathcal{O}}, \vec{\sigma}_{\mathcal{I}}) - r_{i_l}(-\sigma_{i_l}, \sigma_{\partial i_l}) P^t(\vec{\sigma}_{\mathcal{O}}, F_{i_l}[\vec{\sigma}_{\mathcal{I}}]) \right] - \\ &- \sum_{l=1}^m \sum_{\sigma_{\partial o_l \setminus \{\mathcal{I}, \mathcal{O}\}}} \left[ r_{o_l}(\sigma_{o_l}, \sigma_{\partial o_l}) P^t(\sigma_{\partial o_l \setminus \{\mathcal{I}, \mathcal{O}\}}, \vec{\sigma}_{\mathcal{O}}, \vec{\sigma}_{\mathcal{I}}) - \right. \\ &\quad \left. - r_{o_l}(-\sigma_{o_l}, \sigma_{\partial o_l}) P^t(\sigma_{\partial o_l \setminus \{\mathcal{I}, \mathcal{O}\}}, F_{o_l}[\vec{\sigma}_{\mathcal{O}}], \vec{\sigma}_{\mathcal{I}}) \right]. \end{aligned} \quad (4.5)$$

The first sum of the master equation (4.5) represents the probability per unit of time that one of the spins in the internal nodes changes its value between  $t$  and  $t + dt$ . The second and third lines contain the contribution corresponding to the external nodes. It is precisely in these latter lines that one can notice that the time derivative of  $P^t(\vec{\sigma}_{\mathcal{O}}, \vec{\sigma}_{\mathcal{I}})$  depends on probability distributions defined over larger sets

of connected nodes:  $P^t(\sigma_{\partial o_l \setminus \mathcal{I}}, \vec{\sigma}_{\mathcal{O}}, \vec{\sigma}_{\mathcal{I}})$ . This means that Eq. (4.5) does not provide a closed system of differential equations. Similarly to what was done for the original CME, auxiliary equations must be deduced to complement Eq. (4.5).

#### 4.1.1 Equations for the cavity probabilities

To close the master equation (4.5), new equations will be introduced that allow writing  $P^t(\sigma_{\partial o_l \setminus \mathcal{I}}, \vec{\sigma}_{\mathcal{O}}, \vec{\sigma}_{\mathcal{I}})$  in terms of  $P^t(\vec{\sigma}_{\mathcal{O}}, \vec{\sigma}_{\mathcal{I}})$ . As in subsection (2.3), the cavity conditional probabilities will be used, but in this case defined over sets with more than two variables.

It is convenient to define a probability density  $q$  in the space of histories, but only over a reduced space in which the value of one of the histories  $X_j$  is completely fixed. This means that the densities  $Q$  and  $q$  are fundamentally different, existing in distinct stochastic processes. For  $q$ , the history  $X_j$  is just an external parameter equivalent to a time-dependent field at site  $j$ .

For example, the probability density  $Q$  for the connected set of nodes in Figs. (4.1b) and (4.1c) is given by Eq. (4.2). If the history of the node  $X_j$  is fixed as an external parameter, it is no longer necessary to include the weight expressed by the dynamic cavity message  $\mu_{j \rightarrow (ji)}^t(X_j | X_i)$ . The corresponding probability density  $q$  is simply:

$$\begin{aligned} q^t(X_i, X_{\partial i \setminus j}, X_{\partial k \setminus i}, X_{\partial l \setminus i} \parallel X_j) = & \Phi_i^t(X_i | X_{\partial i}) \left[ \prod_{m \in \partial i \setminus j} \Phi_m^t(X_m | X_{\partial m}) \right] \times \\ & \times \left[ \prod_{n \in \partial k \setminus i} \mu_{n \rightarrow (nk)}^t(X_n | X_k) \right] \left[ \prod_{l \in \partial l \setminus i} \mu_{n \rightarrow (nl)}^t(X_n | X_l) \right]. \end{aligned} \quad (4.6)$$

The symbol  $\parallel$  is used to create a contrast with the usual symbol for conditional probabilities  $|$ , because, as seen,  $q$  is not defined as a usual conditional probability. Similarly to how it was explained previously for  $Q$ , the cavity probability density  $q$  can be written for any set of connected nodes if the history of one of its external nodes is fixed.

With this in mind, one can proceed with the deduction of a differential equation for the following marginal probabilities:

$$p^t(\vec{\sigma}_{\mathcal{O} \setminus o_j}, \vec{\sigma}_{\mathcal{I}} \parallel X_{o_j}) = \sum_{\vec{X}_{\mathcal{O} \setminus o_j} \mid \vec{\sigma}_{\mathcal{O} \setminus o_j}}^t \sum_{\vec{X}_{\mathcal{I}} \mid \vec{\sigma}_{\mathcal{I}}}^t q^t(\vec{X}_{\mathcal{O} \setminus o_j}, \vec{X}_{\mathcal{I}} \parallel X_{o_j}). \quad (4.7)$$

These are also cavity conditional probabilities, only now written generically for any set  $\mathcal{I}$  of internal nodes and  $\mathcal{O}$  of external nodes. With a differentiation process

analogous to the one used to obtain Eq. (4.5), we arrive at:

$$\begin{aligned} \frac{dp^t(\vec{\sigma}_{\mathcal{O} \setminus o_j}, \vec{\sigma}_{\mathcal{I}} \| X_{o_j})}{dt} = & - \sum_{l=1}^n [r_{i_l}(\sigma_{i_l}, \sigma_{\partial i_l}) p^t(\vec{\sigma}_{\mathcal{O} \setminus o_j}, \vec{\sigma}_{\mathcal{I}} \| X_{o_j}) - \\ & - r_{i_l}(-\sigma_{i_l}, \sigma_{\partial i_l}) p^t(\vec{\sigma}_{\mathcal{O} \setminus o_j}, F_{i_l}[\vec{\sigma}_{\mathcal{I}}] \| X_{o_j})] - \\ & - \sum_{\substack{l=1 \\ l \neq j}}^m \sum_{\sigma_{\partial o_l \setminus \mathcal{I}}} [r_{o_l}(\sigma_{o_l}, \sigma_{\partial o_l}) p^t(\sigma_{\partial o_l \setminus \mathcal{I}}, \vec{\sigma}_{\mathcal{O} \setminus o_j}, \vec{\sigma}_{\mathcal{I}} \| X_{o_j}) - \\ & - r_{o_l}(-\sigma_{o_l}, \sigma_{\partial o_l}) p^t(\sigma_{\partial o_l \setminus \mathcal{I}}, F_{o_l}[\vec{\sigma}_{\mathcal{O} \setminus o_j}], \vec{\sigma}_{\mathcal{I}} \| X_{o_j})]. \end{aligned} \quad (4.8)$$

Like Eq. (4.5), Eq. (4.8) is exact in graphs that are locally trees. However, neither of the two equations can be solved analytically or numerically. The following subsection shows how to close the equations.

#### 4.1.2 Closure and hierarchical approximations

First, it will be assumed that to determine the cavity probability distribution  $p^t(\vec{\sigma}_{\mathcal{O} \setminus o_j}, \vec{\sigma}_{\mathcal{I}} \| X_{o_j})$  of the configuration at time  $t$ , it is sufficient to know the final state of the history  $X_{o_j}$ . Thus, the complete trajectory  $X_{o_j}$  will be replaced by its value  $\sigma_{o_j}$  at the final time  $t$  throughout Eq. (4.8).

Even so, the equation is not closed because the time derivative of the probabilities  $p^t(\vec{\sigma}_{\mathcal{O} \setminus o_j}, \vec{\sigma}_{\mathcal{I}} \| \sigma_{o_j})$  depends on other probabilities defined over larger sets of nodes:  $p^t(\sigma_{\partial o_l \setminus \mathcal{I}}, \vec{\sigma}_{\mathcal{O} \setminus o_j}, \vec{\sigma}_{\mathcal{I}} \| \sigma_{o_j})$ .

At this point, it is useful to recall what was done for the CME in Section 2.3. Through some approximations, the probability  $p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)$  was written entirely in terms of the probabilities of two nodes  $p^t(\sigma_i \| \sigma_j)$ . From the definition of conditional probability, we have:

$$p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) = p^t(\sigma_{\partial i \setminus j} | \sigma_i \| \sigma_j) p^t(\sigma_i \| \sigma_j). \quad (4.9)$$

It must be remembered that the spins in  $\sigma_{\partial i \setminus j}$  do not interact directly with  $\sigma_j$ . When the random graph is sufficiently large, these variables become virtually disconnected upon removing node  $i$  from the graph. Since the probability  $p^t(\sigma_{\partial i \setminus j} | \sigma_i \| \sigma_j)$  is already conditioned on the value of  $\sigma_i$ , it will be assumed that the information about  $\sigma_j$  is superfluous and thus  $p^t(\sigma_{\partial i \setminus j} | \sigma_i \| \sigma_j) \approx p^t(\sigma_{\partial i \setminus j} | \sigma_i)$ .

Similarly, we will say that the spins  $\sigma_{\partial i \setminus j}$  themselves are independent once the value of  $\sigma_i$  at time  $t$  is known, so we have the factorization:

$$p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) = p^t(\sigma_{\partial i \setminus j} | \sigma_i \| \sigma_j) p^t(\sigma_i \| \sigma_j) \approx \left[ \prod_{k \in \partial i \setminus j} p^t(\sigma_k \| \sigma_i) \right] p^t(\sigma_i \| \sigma_j), \quad (4.10)$$

where the conditional probabilities  $p^t(\sigma_i | \sigma_j)$  have also been equated with the cavity conditional probabilities  $p^t(\sigma_i \| \sigma_j)$ .

The previous assumptions are correct in equilibrium, as explained when introducing the cavity method in Section 3.1. However, in a dynamic process, there are temporal correlations between the neighbors of node  $i$  with the value of  $\sigma_i$  at times prior to  $t$ . In this case, conditioning on  $\sigma_i$  at time  $t$  is not enough for its neighbors to be independent. Eq. (4.10) is an approximation that should work only if those temporal correlations are low.

In Section 2.3, it was shown how Eq. (4.10) leads to the CME. Although this technique has been useful for describing the dynamics of different models in previous works [35, 36], we go a step further here to obtain equations that are appreciably more accurate in several cases of interest. The new differential equations are written for the time evolution of the cavity conditional probabilities  $p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)$ . According to the general equation (4.8), we have:

$$\begin{aligned} \frac{dp^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)}{dt} = & -r_i(\sigma_i, \sigma_{\partial i}) p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) + r_i(-\sigma_i, \sigma_{\partial i}) p^t(\sigma_{\partial i \setminus j}, -\sigma_i \| \sigma_j) \\ & - \sum_{l \in \partial i \setminus j} \sum_{\sigma_{\partial l \setminus i}} \left\{ r_l(\sigma_l, \sigma_{\partial l}) p^t(\sigma_{\partial l \setminus i}, \sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) - \right. \\ & \left. - r_l(-\sigma_l, \sigma_{\partial l}) p^t(\sigma_{\partial l \setminus i}, F_l[\sigma_{\partial i \setminus j}], \sigma_i \| \sigma_j) \right\}, \end{aligned} \quad (4.11)$$

where  $X_j$  has already been replaced by its final value  $\sigma_j$ .

To close these equations, we must introduce some relation between the probabilities  $p^t(\sigma_{\partial l \setminus i}, \sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)$  and  $p^t(\sigma_{\partial i \setminus j}, \sigma_i | \text{big} | \sigma_j)$ . Analogous to Eq. (4.9), we have:

$$p^t(\sigma_{\partial l \setminus i}, \sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) = p^t(\sigma_{\partial l \setminus i} | \sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j). \quad (4.12)$$

In the considered set, the node closest to the set of nodes  $\partial l \setminus i$  is precisely  $l$ . It is said that  $l$  is at distance  $d = 1$  from  $\partial l \setminus i$  because there are graph edges that connect them directly (Fig. (4.1b) illustrates this example). On the other hand, node  $i$  is at distance  $d = 2$  from all nodes in  $\partial l \setminus i$ . The shortest path connecting them contains exactly two graph edges. The rest of the nodes appearing in Eq. (4.12) are at distance  $d = 3$  from  $\partial l \setminus i$ .

As before, it will be considered that conditioning  $p^t(\sigma_{\partial l \setminus i} | \sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)$  on the nodes farthest from  $\sigma_{\partial l \setminus i}$  is superfluous once  $\sigma_i$  and  $\sigma_l$  are known. Therefore:

$$p^t(\sigma_{\partial l \setminus i} | \sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) \approx p^t(\sigma_{\partial l \setminus i} | \sigma_l \| \sigma_i) \approx \frac{p^t(\sigma_{\partial l \setminus i}, \sigma_l \| \sigma_i)}{\sum_{\sigma'_{\partial l \setminus i}} p^t(\sigma'_{\partial l \setminus i}, \sigma_l \| \sigma_i)}, \quad (4.13)$$

where  $p^t(\sigma_{\partial l \setminus i}, \sigma_l | \sigma_i)$  has again been changed to  $p^t(\sigma_{\partial l \setminus i}, \sigma_l \| \sigma_i)$ , which is exact only if the correlations between the spins are low [35].

Note that with Eq. (4.13) we have succeeded in writing  $p^t(\sigma_{\partial l \setminus i} | \sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)$  entirely in terms of  $p^t(\sigma_{\partial l \setminus i}, \sigma_l | \text{big} | \sigma_i)$ . The new cavity master equation is:

$$\begin{aligned} \frac{dp^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)}{dt} = & -r_i(\sigma_i, \sigma_{\partial i}) p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) + r_i(-\sigma_i, \sigma_{\partial i}) p^t(\sigma_{\partial i \setminus j}, -\sigma_i \| \sigma_j) \\ & - \sum_{l \in \partial i \setminus j} \sum_{\sigma_{\partial l \setminus i}} \left\{ r_l(\sigma_l, \sigma_{\partial l}) p^t(\sigma_{\partial l \setminus i} | \sigma_l \| \sigma_i) p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) - \right. \\ & \left. - r_l(-\sigma_l, \sigma_{\partial l}) p^t(\sigma_{\partial l \setminus i} | -\sigma_l \| \sigma_i) p^t(F_l[\sigma_{\partial i \setminus j}], \sigma_i \| \sigma_j) \right\}. \end{aligned} \quad (4.14)$$

Similarly, the equations for the probabilities  $P^t$  associated with a node and all its neighbors are written as follows:

$$\begin{aligned} \frac{dP^t(\sigma_{\partial i}, \sigma_i)}{dt} = & -r_i(\sigma_i, \sigma_{\partial i}) P^t(\sigma_{\partial i}, \sigma_i) + r_i(-\sigma_i, \sigma_{\partial i}) P^t(\sigma_{\partial i}, -\sigma_i) \\ & - \sum_{l \in \partial i} \sum_{\sigma_{\partial l \setminus i}} \left\{ r_l(\sigma_l, \sigma_{\partial l}) p^t(\sigma_{\partial l \setminus i} | \sigma_l \| \sigma_i) P^t(\sigma_{\partial i}, \sigma_i) - \right. \\ & \left. - r_l(-\sigma_l, \sigma_{\partial l}) p^t(\sigma_{\partial l \setminus i} | -\sigma_l \| \sigma_i) P^t(F_l[\sigma_{\partial i}], \sigma_i) \right\}. \end{aligned} \quad (4.15)$$

In contrast to the CME, which is written for cavity conditional probabilities  $p^t(\sigma_i \| \sigma_j)$  where the maximum distance between the involved nodes is  $d_{\max} = 1$ , the new equations (4.14), (4.13), and (4.15) consider the probabilities  $p^t(\sigma_{\partial l \setminus i}, \sigma_l \| \sigma_i)$  where  $d_{\max} = 2$ . Thus, they allow for non-trivial correlations between the spins  $\sigma_{\partial l \setminus i}$ , which are no longer necessarily independent when conditioned on  $\sigma_i$ .

The same procedure can be repeated to include  $d_{\max} = 3, 4, 5, \dots$  (see Appendix G). In this way, a hierarchical system of approximations is constructed, where the equations of the original CME are just the first level, and Eqs. (4.14), (4.13), and (4.15) represent the second level. In the rest of this chapter, the former will be referred to as *CME-1* and the latter as *CME-2*.

The next section presents numerical results of CME-2 and compares them with CME-1 and other methods from the literature.

## 4.2 Numerical results

To numerically integrate the equations, the dynamics rules  $r_i(\sigma_i, \sigma_{\partial i})$  and the structure of the interaction graph must be defined. In what follows, random graphs from two families are considered: Erdős-Rényi (ER) and random regular graphs (RRG). Section 1.1 explains how to construct these graphs when the interactions are between more than two variables. Here, the interactions are only between pairs of variables, but the procedure for their construction is analogous.

Subsection 4.2.1 shows the numerical differences between *CME-1* and *CME-2* in two well-known models of disordered systems physics on random graphs: the ferromagnetic Ising model and the Viana-Bray model for spin glasses. Both are defined on ER graphs. In 4.2.2, *CME-2* and *CME-3* are compared with the equations from

Ref. [22] in the dynamics of the ferromagnetic Ising model on RRG. Finally, subsection 4.2.3 applies *CME-2* to the Susceptible-Infectious-Susceptible model for epidemics.

### 4.2.1 First and second levels of approximation

Here again, the Glauber transition probabilities per unit time are used [70] (see Eq. 3.15). Using Kronecker deltas to introduce the number  $u_i \equiv u_i(\sigma_i, \sigma_{\partial i}) = \sum_{k \in \partial i} \delta_{J_{ki}, -\sigma_k \sigma_i}$  of unsatisfied interactions between  $i$  and its neighbors, the form of the rule  $r_i$  simplifies to:

$$r_i(\sigma_i, \sigma_{\partial i}) \equiv r_i(c_i, u_i) = \frac{1}{2} [1 - \tanh(\beta(c_i - 2u_i))], \quad (4.16)$$

where  $c_i$  is the connectivity of spin  $i$ . When the number of unsatisfied interactions  $u_i$  is large, there is a high probability that the spin will flip from  $\sigma_i$  to  $-\sigma_i$ .

As seen in subsection 3.3.2, by setting  $J_{ij} = 1$  for all pairs  $(ij)$ , the ferromagnetic Ising model defined on a random graph is obtained, which is taken as an ER graph. Figure (4.2) compares the results of Monte Carlo simulations of the system dynamics with the numerical integration of the equations at the first (*CME-1*) and second (*CME-2*) levels of approximation. In all cases, the initial condition is the configuration where all spins point in the same direction. Panels (4.2b) and (4.2d) plot the local errors between the theoretical techniques and Monte Carlo simulations:

$$\delta m^{\text{TH}}(t) = \sqrt{\frac{1}{N} \sum_i (m_i^{\text{TH}}(t) - m_i^{\text{MC}}(t))^2} \quad (4.17)$$

$$\delta e^{\text{TH}}(t) = \sqrt{\frac{1}{2M} \sum_{i \neq j} (e_{ij}^{\text{TH}}(t) - e_{ij}^{\text{MC}}(t))^2}, \quad (4.18)$$

where  $M$  is the number of connected pairs, or edges, in the interaction graph. The quantities  $x^{\text{TH}}$  are obtained from (*CME-1*) or (*CME-2*), while  $x^{\text{MC}}$  is the result of Monte Carlo simulations.

The order-disorder transition occurs at the temperature  $T_c(\kappa) = (\text{arctanh}[1/\kappa])^{-1}$  [90] (in this case  $T_c(\kappa = 3) \approx 2.89$ ). Above this temperature, the equilibrium magnetizations are zero, and below, they are non-zero. Both approximation levels, *CME-1* and *CME-2*, accurately describe the behavior of magnetization, especially in the stationary state. This is not surprising since Section 3.2 demonstrated that the stationary state of *CME-1* is equivalent to the cavity method in equilibrium. Both *CME-1* and *CME-2* provide the exact solution in this case. The discrepancies observed in the figure are due to finite-size effects of the system.

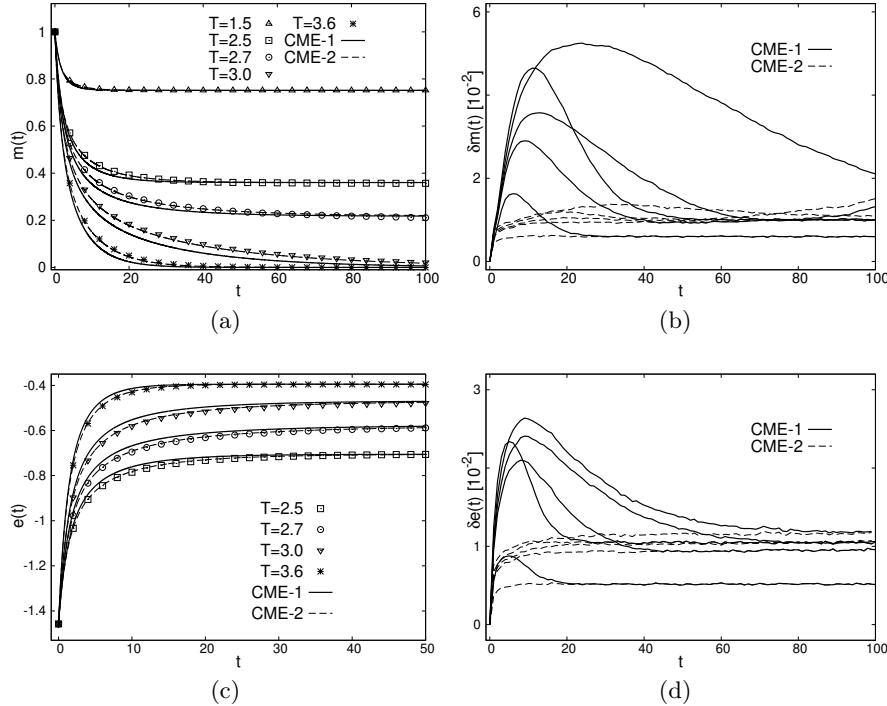


Figure 4.2: Dynamics of the ferromagnetic Ising model according to Monte Carlo simulations (points) and the integration of the *CME-1* equations (solid lines) and *CME-2* equations (dashed lines). In all cases, a completely magnetized system evolves in contact with a thermal bath at a fixed temperature  $T$ . Interactions were defined on Erdős-Rényi graphs of size  $N = 5000$  and average connectivity  $\kappa = 3$ . **a)** Temporal evolution of magnetization. **b)** Local error corresponding to magnetization (see Eq. (4.17)). **c)** Temporal evolution of energy density. **d)** Local error corresponding to energy (see Eq. (4.18)).

The results of *CME-2* are closer to the simulations than those of *CME-1* in the transient regime (relaxation to equilibrium). The local errors of Eqs. (4.17) and (4.18) are also significantly smaller.

Another way to select the couplings is to choose each one from the bimodal distribution  $d(J_{ij}) = 1/2\delta(J_{ij} - 1) + 1/2\delta(J_{ij} + 1)$ . In this model, also defined for ER graphs, there is a transition to a spin glass phase at temperature  $T_{SG}$ . Since correlations play an important role when temperatures drop below  $T_{SG}$ , it is known that the approximations made in subsection 4.1.2 cannot provide results as good as those for the ferromagnetic model.

Figure (4.3) shows results for the Viana-Bray model on ER graphs with  $\kappa = 3$ . The transition occurs in this case at  $T_{SG} = (\text{arctanh}[\sqrt{1/c}])^{-1} \approx 1.52$  [94]. As expected, if  $T < T_{SG}$ , the description provided by *CME-1* and *CME-2* of the dynamics is not

accurate. While the local errors corresponding to Viana-Bray, plotted in panels (4.3b) and (4.3d), are on the order of  $10^{-1}$ , the graphs corresponding to the ferromagnetic model (panels (4.2b) and (4.2d)) are on a scale that is one order of magnitude lower ( $10^{-2}$ ).

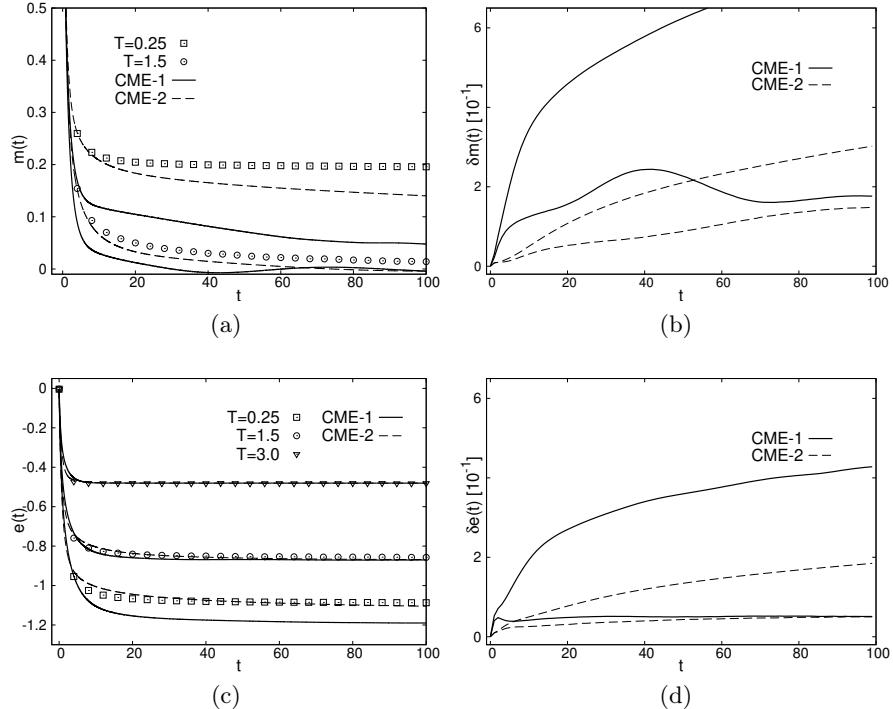


Figure 4.3: Dynamics of the Viana-Bray model for spin glass according to Monte Carlo simulations (points) and the integration of the *CME-1* equations (solid lines) and *CME-2* equations (dashed lines). In all cases, a completely magnetized system evolves in contact with a thermal bath at a fixed temperature  $T$ . Interactions were defined on Erdős-Rényi graphs of size  $N = 1000$  and average connectivity  $\kappa = 3$ . **a)** Temporal evolution of magnetization. **b)** Local error corresponding to magnetization (see Eq. (4.17)). **c)** Temporal evolution of energy density. **d)** Local error corresponding to energy (see Eq. (4.18)).

However, in all cases, *CME-2* is closer to the Monte Carlo results, and its local errors are smaller. As observed in the lower panels of Fig. (4.3), this second level of approximation predicts the average energy density  $e(t)$  even at low temperatures such as  $T = 0.25 \ll T_{SG}$ . One should be cautious with this result since the corresponding local errors are large and continue to grow over time.

The results from Figs. (4.2) and (4.3) indicate that the accuracy of the original CME (here *CME-1*) is considerably improved by moving to the second level of the

proposed hierarchical approximation system. It is important to note that integrating this *CME-2* for models defined on random graphs with low average connectivity does not imply high computational efforts considering the capabilities of modern personal computers. However, as connectivity increases, the number of equations grows and the number of terms in the sums of each equation increases (see Eq. (4.14)), which makes its implementation for a given graph more challenging.

In the next subsection, *CME-2* is compared with a technique that also considers the differential equations of the probabilities associated with a spin and its entire neighborhood. In this sense, they are approximations of the same order, though different.

#### 4.2.2 Comparison with the Dynamical Independent Neighbors Approximation

The hierarchical approximations introduced to study stochastic local search algorithms [21, 95] and systematized in Ref. [22] combine theoretical simplicity with numerical precision in various traditional models of disordered systems physics. In this subsection, the previously proposed equations are compared with the Dynamic Approximation of Independent Neighbors (DINA), which is the second level of the approximations presented in Ref. [22].

In the case of the ferromagnetic Ising model with Glauber dynamic rules (4.16) defined on random regular graphs (RRG), where each node has exactly  $c$  neighbors, DINA reduces to the following equations:

$$\begin{aligned} \frac{d\hat{P}^t(\sigma, u)}{dt} = & -r(c, u)\hat{P}^t(\sigma, u) + r(c, c-u)\hat{P}^t(-\sigma, c-u) - \\ & - \sum_{u'=0}^{c-1} (c-u') r(c, u') \hat{P}^t(\sigma, u') \left[ \hat{P}^t(u | \sigma, \sigma, u') - \hat{P}^t(u-1 | \sigma, \sigma, u') \right] - \\ & - \sum_{u'=0}^{c-1} (u'+1) r(c, u'+1) \hat{P}^t(\sigma, u'+1) \left[ \hat{P}^t(u-1 | \sigma, -\sigma, u') - \hat{P}^t(u | \sigma, -\sigma, u') \right]. \end{aligned} \quad (4.19)$$

The equations for  $u=0, c$  are easily obtained from (4.19) by removing the second or third lines, respectively. Closure relations are also needed [22]:

$$\hat{P}^t(\hat{u} | \sigma, \sigma, \hat{u}') \approx \frac{(c-\hat{u})\hat{P}^t(\sigma, \hat{u})}{\sum_{u'}(c-u')\hat{P}^t(\sigma, u')} \quad (4.20)$$

$$\hat{P}^t(\hat{u} | \sigma, -\sigma, \hat{u}') \approx \frac{(\hat{u}+1)\hat{P}^t(\sigma, \hat{u}+1)}{\sum_{u'}u'\hat{P}^t(\sigma, u')}. \quad (4.21)$$

Here,  $r(c, u) \equiv r_i(c_i = c, u_i = u)$  (see Eq. (4.16)) and  $\hat{P}^t(\sigma, u)$  is the probability of finding a node with spin  $\sigma$  and  $u$  unsatisfied links with its neighbors. This can be written as an average over the family of random regular graphs of intermediate magnitude:

$$P^t(\sigma, u) = \frac{1}{N} \sum_{i=1}^N \sum_{\sigma_i} \sum_{\sigma_{\partial i}} P^t(\sigma_i, \sigma_{\partial i}) \delta_{\sigma, \sigma_i} \delta_{(\sum_{k \in \partial i} \sigma_k), c-2u}. \quad (4.22)$$

Note that Eq. (4.21) is an average equation similar to that in Section 3.3. In other words, DINA is an averaging technique; it does not provide information about the dynamics on a specific graph of interactions but rather about the average dynamics.

Similarly to  $\hat{P}^t(\sigma, u)$ , one can define the probability  $\hat{P}^t(\sigma, \sigma', \hat{u}, \hat{u}')$  of finding a pair of connected variables  $(\sigma, \sigma')$  such that  $\hat{u}$  and  $\hat{u}'$  are their corresponding numbers of unsatisfied interactions with their other neighbors. In this case,  $\hat{u}, \hat{u}'$  are integers between zero and  $c - 1$ . The conditional probability  $\hat{P}^t(\hat{u} | \sigma, \sigma', \hat{u}')$  from Eq. (4.19) is written as:

$$\hat{P}^t(\hat{u} | \sigma, \sigma', \hat{u}') = \frac{\hat{P}^t(\sigma, \sigma', \hat{u}, \hat{u}')}{\sum_{\hat{u}} \hat{P}^t(\sigma, \sigma', \hat{u}, \hat{u}')}. \quad (4.23)$$

The closure relation used in DINA can be deduced from the Dynamic Replica Theory (DRT) with an *ansatz* that maintains the symmetry of replicas [22]. This explains why DINA provides very accurate results in models where detailed balance is satisfied, where the assumptions of DRT are more likely to be fulfilled.

The inserted plot in Fig. (4.4a) shows a very good match between the average magnetization obtained from DINA and Monte Carlo simulations for the ferromagnetic Ising model on random regular graphs. The inserted plot in Fig. (4.4b) quantifies the numerical differences between Monte Carlo and DINA. Since DINA does not provide local information and no local error formula like Eq. (4.17) can be applied in this case, the absolute difference between the average magnetizations predicted by DINA and Monte Carlo was taken. This parameter  $\Delta m$  is on the order of  $10^{-3}$  for all times studied.

Like DINA, *CME-2* is written for the probabilities associated with a node and its neighbors  $P^t(\sigma_i, \sigma_{\partial i})$ . In fact, in both cases, the closure relations take conditional probabilities such as  $\hat{P}^t(\hat{u} | \sigma, \sigma', \hat{u}')$  and discard information that is more than  $d = 2$  distance away from the variables in the probability argument. An important difference is that *CME-2* is written for dynamics on a specific graph, while DINA describes the average dynamics.

However, the equations of *CME-1*, *CME-2*, and *CME-3* can be reduced to a single equation in each case if random regular graphs and homogeneous initial conditions are considered. Information about the graph structure reduces to knowing the number  $c$  of neighbors per node. Appendix H explains how to obtain the different average equations needed to reproduce Fig. (4.4).

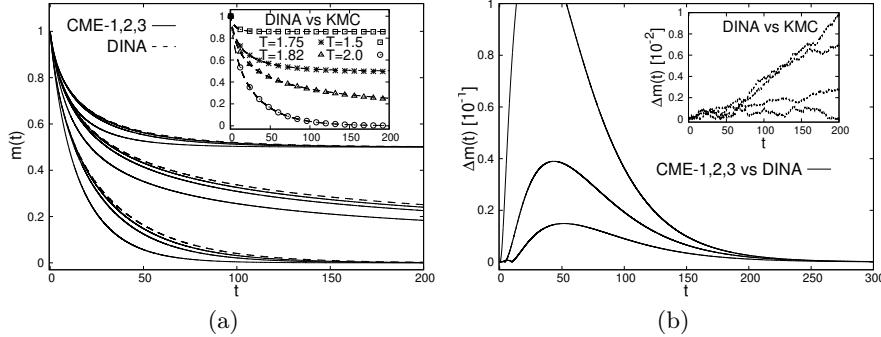


Figure 4.4: Dynamics of the ferromagnetic Ising model defined on random regular graphs with connectivity  $c = 3$ . The figure presents results from Monte Carlo simulations (points), as well as the numerical integration of DINA (dashed lines) and the equations of *CME-1*, *CME-2*, and *CME-3* (solid lines). In all cases, a completely magnetized system evolves in contact with a thermal bath at a fixed temperature  $T$ . The Monte Carlo simulations were conducted for graphs of size  $N = 10^4$ . **a)** Evolution of the average magnetization. The solid lines in the main plot represent the results from the first three levels *CME-1,2,3*. These approach the dashed DINA curve as the approximation level increases. From top to bottom, the temperatures are  $T = 1.75, 1.82, 2.0$ . The inserted plot compares DINA with Monte Carlo simulations. From top to bottom, the temperatures are  $T = 1.5, 1.75, 1.82, 2.0$ . **b)** The main plot shows the absolute difference  $\Delta m$  between the magnetizations predicted by DINA and the different approximation levels presented for the temperature  $T = 2.0$ . As this level increases, the solid curves approach  $\Delta m = 0$ . The inserted plot similarly compares DINA with Monte Carlo simulations. From top to bottom, the temperatures are  $T = 1.82, 1.75, 1.5, 2.0$ .

For example, *CME-2* transforms into:

$$\begin{aligned} \frac{d\hat{P}^t(\sigma, u)}{dt} = & -r(c, u)\hat{P}^t(\sigma, u) + r(c, c-u)\hat{P}^t(-\sigma, c-u) - \\ & -(c-u)\langle r(c, u') \rangle_{\sigma, \sigma'}\hat{P}^t(\sigma, u) + (u+1)\langle r(c, u') \rangle_{-\sigma, \sigma}\hat{P}^t(\sigma, u+1) - \\ & -u\langle r(c, u') \rangle_{-\sigma, \sigma}\hat{P}^t(\sigma, u) + (c-u+1)\langle r(c, u') \rangle_{\sigma, \sigma}\hat{P}^t(\sigma, u-1), \end{aligned} \quad (4.24)$$

where the averages  $\langle r(c, u') \rangle_{\sigma, \sigma'} = \sum_{u'=0}^{c-1} r(c, u') \hat{p}^t(u' | \sigma \| \sigma')$  are defined. The equations for  $u=0, c$  are easily obtained from (4.24) by removing the second or third lines, respectively.

Equation (4.24) is analogous to Equation (4.19) not only because both express the temporal derivative of the same quantity but also because they have similar structures. The main difference lies in the approximation:

$$\hat{P}^t(\sigma, u) \hat{P}^t(u' | \sigma', \sigma, u) \approx \hat{P}^t(\sigma, u) \hat{p}^t(u' | \sigma', \sigma) \approx \hat{P}^t(\sigma, u) \hat{p}^t(u' | \sigma' \| \sigma). \quad (4.25)$$

This is the expression for the conditional probabilities used in *CME-2*. In contrast, DINA uses the relations (4.20) and (4.21).

Fig. (4.4) includes a comparison between the results from DINA and those from the *CME-1*, *CME-2*, and *CME-3* equations. As the approximation level increases, the magnetization predicted by the equations presented in this chapter approaches that obtained from DINA. At the same time, this means that the dynamics of the Monte Carlo simulations are described increasingly better.

In the next subsection, the new equations are applied to a different model, which does not fulfill detailed balance and is therefore completely out of equilibrium. In this case, the results are compared with one of the best mean-field techniques available in the literature.

#### 4.2.3 SIS model for epidemics

In this subsection, we work with the Susceptible-Infectious-Susceptible (SIS) model for epidemics and define interactions on a random regular graph. The SIS model was defined in subsection 3.3.3. Here, we present numerical results for the different approximations of the hierarchical system introduced in Section 4.1.

The dynamics of the SIS on random graphs has motivated abundant literature. One of the most successful methodologies in this scenario is the mean-field approach [96, 97]. It complements the master equation with some useful factorization and calculates epidemic thresholds. The most accurate version is the already mentioned pair-based mean-field approximation (PBMF), which accounts for pair correlations in the system and has been extensively studied in the literature [98, 99] with very good results.

The PBMF equations are:

$$\frac{d\rho_i}{dt} = -\mu\rho_i + \beta \sum_{j \in \partial i} (\rho_j - \psi_{ij}) \quad (4.26)$$

$$\begin{aligned} \frac{d\psi_{ij}}{dt} = & -2(\mu + \beta)\psi_{ij} + \beta(\rho_i + \rho_j) + \\ & + \beta \frac{\rho_j - \psi_{ij}}{1 - \rho_i} \sum_{k \in \partial i \setminus j} (\rho_k - \psi_{ik}) + \beta \frac{\rho_i - \psi_{ij}}{1 - \rho_j} \sum_{k \in \partial j \setminus i} (\rho_k - \psi_{jk}), \end{aligned} \quad (4.27)$$

where  $\rho_i(t) = \sum_{x_i} x_i P^t(x_i)$  and  $\psi_{ij}(t) = \sum_{x_i} \sum_{x_j} x_i x_j P^t(x_i, x_j)$ .

Reference [38] provides an extensive comparison between PBMF and the *CME-1* equations. It shows that PBMF generally makes more accurate predictions about dynamic observables. Figs. (4.5a) and (4.5c) illustrate this for the specific case of an epidemic starting with a single infectious node on a random regular graph of size

$N = 1000$  and connectivity  $c = 3$ . Although the stationary state matches, PBMF is undeniably a better description of the transient regime. The peaks of the local errors (see Fig. (4.5c)) are approximately four times larger for *CME-1*.

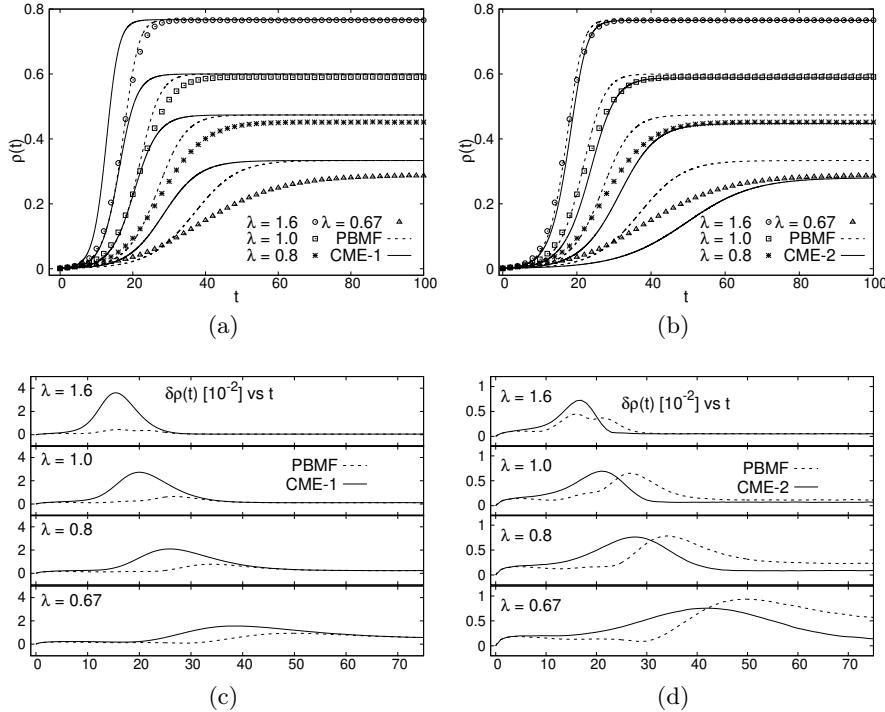


Figure 4.5: Dynamics of the SIS model for epidemics according to Monte Carlo simulations (points), as well as the integration of PBMF (dashed lines) and the equations in the first (*CME-1*, thick solid lines) and second (*CME-2*, thin solid lines) levels of approximation presented in this chapter. A single value for the parameter  $\beta = 0.4$  (see subsection 3.3.3) was used with different values of  $\mu = 0.25, 0.4, 0.5, 0.6$ . These correspond to different infectiousness levels  $\lambda = 1.6, 1.0, 0.8, 0.67$ . In all cases, the infection started from a single node on a random regular graph of size  $N = 1000$  and connectivity  $c = 3$ . To compare with the equations, Monte Carlo simulations where infectious nodes disappeared for some time  $t < 100$  were discarded [93]. **a)** and **b)** Evolution of the average infection probability  $\rho(t)$ . **c)** and **d)** Local error analogous to Equation (4.17) between theoretical techniques and Monte Carlo simulations.

Figs. (4.5b) and (4.5d) contain a comparison with the second level of approximation, *CME-2*. These equations accurately describe the behavior of the simulated epidemic. Fig. (4.5d) shows that the maximum local errors of *CME-2* are of the same order as those of PBMF. For low infectiousness, the stationary state of the simulations aligns better with the predictions of *CME-2*. It can also be observed that for the transient regime, both methods exhibit errors of the same order.

## Partial conclusions

The new hierarchical approximation system improves upon the original CME (*CME-1*) results in all the studied models. Each level of approximation introduces corrections to the CME that account for spatial correlations at larger distances in the interaction graph. Subsection 4.2.1 explores two traditional models from the physics of disordered systems where *CME-2* outperforms *CME-1*. Subsection 4.2.2 demonstrates that the accuracy of the results improves monotonically as the level of approximation increases. The predictions reach a quality comparable to that of successful methods in the literature, such as DINA.

In subsection 4.2.3, it was shown that the second level of approximations of the hierarchical system introduced in Section 4.1 improves the results of PBMF in describing an epidemic spreading over a regular random graph. The same will be true for other families of random graphs, as *CME-2* more accurately considers the system's correlations. This is a significant result given the importance of PBMF among theoretical techniques for epidemic dynamics.

The methodology introduced here is easily generalizable, and this is leveraged to develop another hierarchical system in Chapter 6. The introduction of *CME-2* is also useful for the final chapter of the thesis.

It is particularly attractive that the second level of approximation, *CME-2*, is already appreciably more accurate than the original CME (*CME-1*). However, in the next chapter, *CME-1* rather than *CME-2* will be applied to a completely out-of-equilibrium system. This returns to one of the main motivations of this thesis: the dynamics of stochastic local search algorithms used to solve instances of K-SAT. In this case, the high connectivities involved make the numerical implementation of *CME-2* considerably more challenging.

## Chapter 5

# Local search algorithms in the K-SAT problem

The random K-Satisfiability (K-SAT) problem has been a recurrent topic for the scientific community dedicated to the physics of disordered systems [12, 16–21]. Although this was already discussed in Section 1.2 of the first chapter, it is worth remembering that there is little understanding of the behavior of local search algorithms used to solve instances of K-SAT. There is a lack of tools to describe their dynamics, which in many cases are intrinsically out of equilibrium.

In this chapter, the CME is used to study two algorithms that do not satisfy detailed balance: Focused Metropolis Search (FMS) and the well-known *WalkSAT*. The former is very efficient and can solve instances of K-SAT near the transition between the satisfiable (SAT) zone and the unsatisfiable (UNSAT) zone. The latter algorithm is also interesting because some techniques have been previously developed [21, 95] to study its dynamics completely out of equilibrium. This allows comparing the results of the CME with those of these equations from the literature.

### 5.1 Difficult instances and focused algorithms

One of the relevant questions about combinatorial optimization problems concerns the generation of difficult instances [12]. In the case of K-SAT, Boolean formulas for which there is a solution but it is difficult to find become a valuable test set for new algorithms. In Section 1.2, it was explained that the first theoretical predictions of Disordered Systems Physics about the *hard zone* for K-SAT [12, 16, 51] introduced the parameter  $\alpha_d$ . It was expected that in the region  $\alpha_d < \alpha < \alpha_s$ , where  $\alpha_s$  marks the SAT-UNSAT transition, local search algorithms could not find solutions to sufficiently large formulas.

These ideas were supported by the previous knowledge the community had about the p-spin model [43–45]. As shown in Section 1.1, in the ferromagnetic p-spin model, there is a temperature  $T_d$  of transition to the spin glass phase. Below  $T_d$ , Monte

Carlo simulations, which is a local search algorithm, are simply unable to find the system's minimum energy state if starting from a random configuration. Similarly,  $\alpha_d$  marks the limit above which the zero-temperature Monte Carlo algorithm cannot find solutions to K-SAT if its dynamics start from a random configuration of the variables.

In both cases, the dynamic transition to the spin glass is related to the appearance of long-range correlations. A series of works [48–51] demonstrated that the relevant magnitude is the point-to-set correlation:

$$C_{PS}(i, l) = \sum_{\vec{\sigma}_l} P(\vec{\sigma}_l) \left( \sum_{\sigma_i} P(\sigma_i | \vec{\sigma}_l) \sigma_i \right)^2 - \left( \sum_{\sigma_i} P(\sigma_i) \sigma_i \right)^2. \quad (5.1)$$

If the variable at site  $i$  is  $\sigma_i$  with probability  $P(\sigma_i)$ , the number  $C_{PS}(i, l)$  is a way to express how different the average value of  $\sigma_i$  is once the configuration of the variables at distance  $l$ :  $\vec{\sigma}_l$  is known. When  $C_{PS}(i, l) = 0$ , it is irrelevant to know  $\vec{\sigma}_l$  and the point-to-set correlation is zero. The probabilities in Eq. (5.1) are calculated using the Boltzmann-Gibbs distribution at temperature  $T = 0$  for the boundary  $P(\vec{\sigma}_l)$ . There is a region  $\alpha < \alpha_d$  where  $C_{PS}(i, l) \rightarrow 0$  when  $l \rightarrow \infty$ . On the other hand, for  $\alpha > \alpha_d$  this correlation maintains a value different from zero for any distance  $l$  [51]. To determine the value of  $\alpha$  for which  $C_{PS}(l) \neq 0$  when  $l \rightarrow \infty$ , the correlation distance is defined as  $l^*(\epsilon) = \min\{l : C_{PS}(l) < \epsilon\}$ .

Appendix A contains more details about this calculation. The details for K-SAT can be found in Ref. [51]. Fig.(5.1a) shows  $C_{PS}(l)$  for  $K = 3$ . From  $\alpha_d \approx 3.86$  there is a non-zero correlation at any distance  $l$ .

As already mentioned, it is expected that in the zone  $\alpha_d < \alpha < \alpha_s$ , local search algorithms will have great difficulty solving instances of K-SAT, although solutions generally exist. Each of these algorithms is defined by a rule that allows randomly selecting a variable and flipping it. This rule only uses information about the state of the variable and its neighborhood. Since the divergence of the point-to-set correlation indicates that typically flipping one of the system's variables has repercussions on unbounded sets of variables in the graph, it should be unlikely to find a solution for  $\alpha > \alpha_d$ .

However, Fig. (5.1b) shows that one of the local search algorithms studied in this chapter, called Focused Metropolis Search (FMS) [52], is capable of solving most instances of 3-SAT for values of  $\alpha$  appreciably greater than  $\alpha_d$ . FMS belongs to the family of *focused* algorithms that only change the state of variables involved in unsatisfied clauses (see Section 1.2). Each solution would then be an attractor of the dynamics: once FMS finds a correct answer, it will remain there with probability  $p = 1$ . The dynamics of this algorithm do not satisfy the detailed balance property, so FMS does not converge to a stationary state that follows the Boltzmann-Gibbs distribution. It is a *completely out-of-equilibrium* dynamics.

Of course, it could be argued that the results in Fig.(5.1b) are not sufficient. The statistical theory for  $\alpha_d$  is valid in the limit where the number of variables  $N$  goes to infinity. Another technical difficulty is determining precisely that the solution times are not exponential in  $N$ , meaning that the algorithm is not as bad as an

exhaustive search. Although this topic is not explored here due to space and simplicity constraints, it is worth mentioning that several works [18–20] have presented more evidence suggesting that some local search algorithms can solve instances in zones where the task is supposed to be difficult.

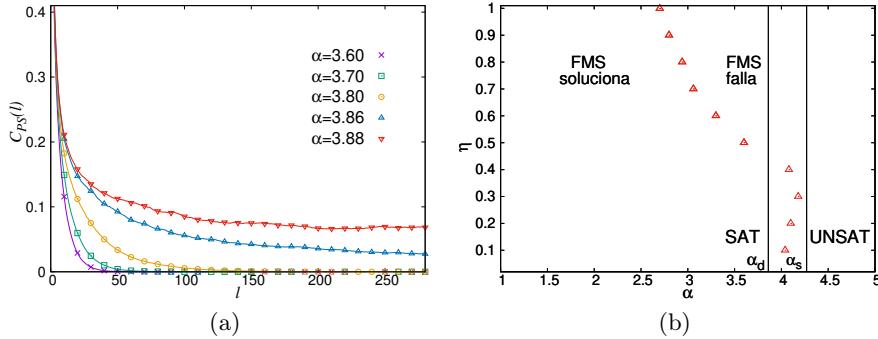


Figure 5.1: Transition to spin glass in K-SAT and local search algorithms. **a)** Point-to-set correlation for 3-SAT. Dependence on distance for various values of  $\alpha$ . **b)** Results of the local search algorithm Focused Metropolis Search (FMS). Instances of 3-SAT with  $N = 50000$  were used. For each value of the algorithmic parameter  $\eta$  (see Section 5.3), the value of  $\alpha$  was determined where half of the algorithm runs find a solution before time  $t/N = 10^5$ . 100 instances were tested for each  $\alpha$ . To the right of the red points, most FMS runs fail. To the left, most succeed. The left vertical line marks the value  $\alpha_d(3) \approx 3.86$ . The right vertical line marks the SAT/UNSAT transition ( $\alpha = 4.2667$ ).

The rest of the sections of this chapter are dedicated to describing two focused local search algorithms: FMS itself and WalkSAT. Few theoretical techniques are known that can be applied in these cases. Hence the relevance of applying the CME to make predictions about their dynamics and their algorithmic transitions.

## 5.2 Adaptation of the Cavity Master Equation

In the K-SAT setup, there is a Boolean formula  $f(\vec{x})$  with  $M$  clauses and  $K$  variables  $x_i$  per clause (see Section 1.2). In total, there are  $N$  variables  $x_i$ . Finding a configuration of the variables that satisfies  $f$  is equivalent to finding the ground state of the Hamiltonian:

$$H(\vec{\sigma}) = \sum_{a=1}^M \prod_{i \in a} \frac{1 - J_i^a \sigma_i}{2}, \quad (5.2)$$

where a variable  $\sigma_i = \pm 1$  has been introduced for each Boolean variable  $x_i$  in the formula. The index  $a$  runs over all clauses of the Boolean formula. The couplings

$J_i^a = \pm 1$  indicate whether the variable  $\sigma_i$  appears negated ( $J_i^a = -1$ ) or not ( $J_i^a = 1$ ) in clause  $a$ .

Thus, K-SAT is transformed into a problem that can be represented using random hypergraphs. In Fig. (5.2a), an example of a Boolean formula is illustrated. The variables  $\sigma_i$  are represented by circles and correspond to the variable nodes of the hypergraph, while the clauses are represented by squares and correspond to the factor nodes. When a variable  $\sigma_i$  is connected to a clause  $a$  with a blue edge, this means that  $J_i^a = 1$ , while for a red edge,  $J_i^a = -1$ .

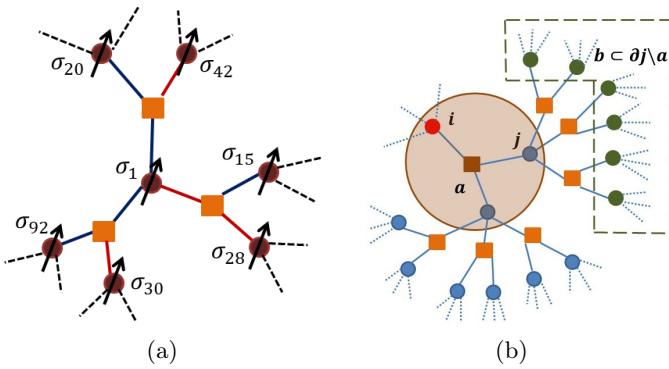


Figure 5.2: Schematic representation of K-SAT instances using hypergraphs. Squares symbolize factor nodes and circles variable nodes. **a)** An example of a Boolean formula is illustrated. Blue edges correspond to positive  $J$  couplings and red edges to negative couplings. The formula is:  $f = (x_1 \vee x_{92} \vee \bar{x}_{30}) \wedge (\bar{x}_1 \vee x_{15} \vee \bar{x}_{28}) \wedge (x_1 \vee x_{20} \vee \bar{x}_{42})$ . **b)** Nodes involved in the CME for hypergraphs are illustrated (see Eq. (5.3)). Nodes marked in green correspond to variable nodes in the set  $\partial j \setminus a$ .

In Chapters 2, 3, and 4, the CME for graphs with pairwise interactions was presented. With minor changes, it is possible to extend it to models defined on hypergraphs [35]. The form of the equations is very similar to Eq. (2.40):

$$\frac{dp(\sigma_{a \setminus i} \| \sigma_i)}{dt} = - \sum_{j \in a \setminus i} \sum_{\sigma_{\partial j \setminus a}} \left\{ r_j(\sigma_j, \sigma_{\partial j}) \left[ \prod_{b \in \partial j \setminus a} p(\sigma_{b \setminus j} \| \sigma_j) \right] p(\sigma_{a \setminus i} \| \sigma_i) \right. \\ \left. - r_j(-\sigma_j, \sigma_{\partial j}) \left[ \prod_{b \in \partial j \setminus a} p(\sigma_{b \setminus j} \| -\sigma_j) \right] p(F_j[\sigma_{a \setminus i}] \| \sigma_i) \right\}. \quad (5.3)$$

Recall that the operator  $F_j[\cdot]$  inverts the value of the spin  $\sigma_j$ , the letters  $a$  and  $b$  represent factor nodes, and the letters  $i$ ,  $j$ , and  $k$  represent variable nodes. The set  $a \setminus i$  contains all variable nodes connected to  $a$  except node  $i$ .

Eq. (5.3) is complemented by a differential equation for the local probabilities:

$$\frac{dP(\sigma_a)}{dt} = - \sum_{j \in a} \sum_{\sigma_{\partial j \setminus a}} \left\{ r_j(\sigma_j, \sigma_{\partial j}) \left[ \prod_{b \in \partial j \setminus a} p(\sigma_{b \setminus j} \| \sigma_j) \right] P(\sigma_a) \right. \\ \left. - r_j(-\sigma_j, \sigma_{\partial j}) \left[ \prod_{b \in \partial j \setminus a} p(\sigma_{b \setminus j} \| -\sigma_j) \right] P(F_j[\sigma_a]) \right\}, \quad (5.4)$$

where  $\sigma_a$  represents the configuration of all variables in clause  $a$ .

The structure of Eq. (5.3) is illustrated in Fig. (5.2b). The sums in the equation run over all configurations of nodes that are neighbors of  $j$  but not in  $a$ . In the figure, these nodes are colored green.

The CME can then be used to describe the dynamics of local search algorithms by specifying the form of the rules  $r_i(\sigma_i, \sigma_{\partial i})$ . Here, nothing needs to be assumed about the structure of the probability distributions that will be the solution of the equations. Nothing prevents using transition probabilities per unit time  $r_i$  that do not satisfy detailed balance and lead the system to be completely out of equilibrium. Appendix I shows how to simplify the equations under some quite general assumptions about the form of  $r_i$ .

In the following sections, the equations are applied to the FMS and WalkSAT algorithms, explaining each algorithm's operation and their respective  $r_i$  rules.

### 5.3 Focused Metropolis Search (FMS)

This section presents the results of the CME in describing the dynamics of the FMS [52], part of the family of *focused* algorithms. The FMS explores the configuration space by deciding in each iteration whether to flip the value of a variable, with the peculiarity of only visiting unsatisfied clauses. Below is a pseudo-code of the FMS to facilitate understanding:

```

1: Let  $\bar{\sigma}$  be a random assignment of variables
2: while  $\bar{\sigma}$  is not a solution do
3:   Select a clause  $C$  uniformly among the unsatisfied clauses
4:   Select a variable  $\sigma_i \in C$  uniformly among the  $K$  possible variables
5:   Calculate the change  $\Delta E$  in the number of unsatisfied clauses if  $\sigma_i$  is flipped
6:   if  $\Delta E \leq 0$  then
7:     Flip  $\sigma_i$ 
8:   else
9:     With probability  $\eta^{\Delta E}$ , flip  $\sigma_i$ 
10:  end if
11: end while
```

The algorithmic parameter  $\eta$  is defined in the interval  $[0, 1]$  and relates to the stochasticity or noise magnitude in the system. When  $\eta = 1$ , all proposed changes are accepted with probability one. On the other hand, when  $\eta = 0$ , only changes that decrease the total number of unsatisfied clauses are accepted. It is observed that

once within an unsatisfied clause and identifying  $\eta$  with a function of temperature ( $\eta = e^{-\beta}$ ), the rules of the dynamics are none other than the Metropolis rules.

Fig. (5.3a) shows the temporal evolution of the average value  $e(t)$  of the number of unsatisfied clauses per variable when running the FMS on different K-SAT instances. The time  $t$  is simply the number of iterations divided by the system size  $N$ . The parameter  $\alpha = M/N$  measures the difficulty of the Boolean formula. The larger  $\alpha$  is, the more constraints per variable and the longer the FMS takes to solve the formula.

Eventually, for each  $\eta$ , some  $\alpha_{\text{FMS}}(\eta)$  appears, beyond which the FMS typically does not solve the formulas, and the final value of  $e(t)$  is strictly positive. This phenomenon is known as an algorithmic transition and is a direct measure of the FMS's efficiency in this case. The larger  $\alpha_{\text{FMS}}(\eta)$  is, the more effective the algorithm is at solving problem instances. For simplicity,  $\alpha_{\text{FMS}}(\eta)$  is defined as the  $\alpha$  value where, on average, the FMS finds a solution for exactly half of the Boolean formulas it faces.

To compare with the CME, it is necessary to determine if the system size used (the number of variables in the formula) is relevant. Fig. (5.3b) presents results for various  $N$  values and the same  $\alpha$  and  $\eta$  in all cases. The algorithm's behavior does not change significantly with  $N$ . This avoids the need to integrate Eqs. (5.3) and (5.4) for large  $N$ , which would have been computationally expensive.

Figure (5.1b) in Section 1.2 shows that FMS can efficiently solve instances of 3-SAT very close to the SAT-UNSAT transition ( $\alpha_s = 4.267$ ). For  $\eta \approx 0.3$ , an algorithmic transition occurs at  $\alpha_{\text{FMS}} \approx 4.23$  [52]. Hence, there is interest in describing its dynamics. To do this, the  $r_i(\sigma_i, \sigma_{\partial i})$  rules appearing in the CME (see Eq. (5.3)) must be appropriately defined.

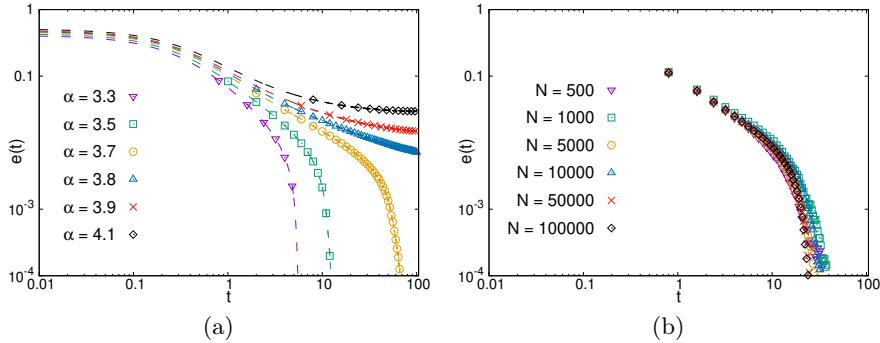


Figure 5.3: FMS results on 3-SAT instances. Both panels show the average number of unsatisfied clauses per variable as a function of time. A logarithmic scale is used for both axes. Each point is the average of 500 runs with  $\eta = 0.45$  in all cases. **a)** Dependence of FMS behavior on the value of  $\alpha = M/N$ . A transition is observed between a region where FMS solves the instances and another where it fails to find solutions. The chosen system size was  $N = 10^5$ . The dashed lines are included only as a visual aid. **b)** Dependence on the system size  $N$  for  $\eta = 0.45$  and  $\alpha = 3.6$ .

First, note that the probability of choosing a specific unsatisfied clause in a step of FMS is  $1/E(\vec{\sigma})$ , where  $E(\vec{\sigma}) = H(\vec{\sigma})$  is simply the number of unsatisfied clauses corresponding to the global configuration  $\vec{\sigma}$ . Then, the probability of choosing a variable randomly within the clause is  $1/K$  and the probability of flipping it is given by the Metropolis rule. Finally, the probability  $p_i$  of flipping the variable  $\sigma_i$  is:

$$p_i = \frac{E_i(\sigma_i, \sigma_{\partial i})}{K E(\vec{\sigma})} \min\{\eta^{\Delta E_i(\sigma_i, \sigma_{\partial i})}, 1\} \quad (5.5)$$

where  $E_i(\sigma_i, \sigma_{\partial i})$  is the number of unsatisfied clauses involving the variable  $\sigma_i$  and  $\Delta E_i(\sigma_i, \sigma_{\partial i})$  is the change in  $E_i$  when  $\sigma_i$  is flipped:

$$E_i(\sigma_i, \sigma_{\partial i}) = \sum_{a \in \partial i} \prod_{i \in a} \frac{1 - J_i^a \sigma_i}{2} \quad (5.6)$$

$$\Delta E_i(\sigma_i, \sigma_{\partial i}) = E_i(-\sigma_i, \sigma_{\partial i}) - E_i(\sigma_i, \sigma_{\partial i}) \quad (5.7)$$

It is important to note that the probability  $p_i$  depends on the entire system's configuration through  $E(\vec{\sigma})$ . To write the  $r_i$  rules of the CME, some approximation must be made. By estimating the average number of unsatisfied clauses from local probabilities:

$$\langle E \rangle = \sum_{a=1}^M \left[ \prod_{i \in a} \frac{1 - J_i^a \sigma_i}{2} \right] P(\sigma_a) \quad (5.8)$$

a local version of the probabilities  $p_i$  can be written, and these values are assigned to the dynamic rules used by the CME:

$$r_i(\sigma_i, \sigma_{\partial i}) = \frac{1}{K e(t)} E_i(\sigma_i, \sigma_{\partial i}) \min\{\eta^{\Delta E_i(\sigma_i, \sigma_{\partial i})}, 1\}. \quad (5.9)$$

Here, the observable  $e = \langle E \rangle / N$  has been introduced, which is equivalent to the energy density worked with in previous chapters. Since the unit of time involves precisely  $N$  iterations of the algorithm, the factor  $1/N$  that would appear in the denominator of Eq. (5.9) is not included. This rule defines a self-consistent dynamic where local transition probabilities depend on the system's global energy.

With Eq. (5.9), we are ready to numerically integrate the CME and compare it with the FMS runs. Fig. (5.4) shows the results for two different values of  $\eta$ . It can be observed that the number of unsatisfied clauses in the FMS runs (points) rapidly decreases to zero for all the values of  $\alpha$  included in the figure. Similarly, the lines representing the CME go to zero quickly for small values of  $\alpha$ . However, for larger  $\alpha$ , the CME converges to a non-zero value of  $e$ , diverging from the algorithm.

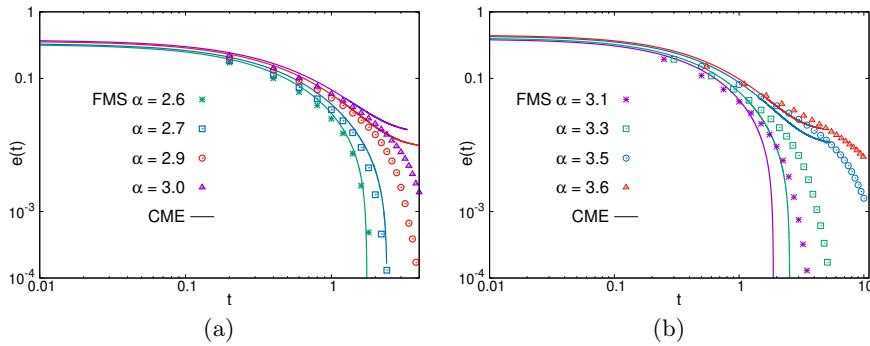


Figure 5.4: Comparison between CME (lines) and FMS (points) in 3-SAT instances. Both panels show the average number of unsatisfied clauses per variable as a function of time. A logarithmic scale is used for both axes. Each point is the average of 500 FMS algorithm runs. **a)**  $\eta = 0.65$  and system size  $N = 5000$  were used. **b)**  $\eta = 0.4$  and  $N = 2000$ .

This qualitative change in the CME predicts an algorithmic transition occurring for  $\alpha_{\text{CME}}(\eta)$  less than the correct  $\alpha_{\text{FMS}}(\eta)$ . Fig. (5.4) indicates that while  $\alpha < \alpha_{\text{CME}}(\eta)$ , the CME lines are close to the algorithm points. Near the transition, high correlations appear that cannot be captured by the theoretical technique used.

Therefore, Fig. (5.5) presents results in the two zones  $\alpha < \alpha_{\text{CME}}(\eta)$  and  $\alpha > \alpha_{\text{FMS}}(\eta)$ . The CME's description is much more accurate in these regions, especially compared to simpler closures of the master equation, such as the following:

$$\frac{dP(\sigma_i)}{dt} = - \sum_{\sigma_{\partial i}} \left[ \prod_{k \in \partial i} P(\sigma_k) \right] \left\{ r_i(\sigma_i, \sigma_{\partial i}) P(\sigma_i) - r_i(-\sigma_i, \sigma_{\partial i}) P(-\sigma_i) \right\}, \quad (5.10)$$

where  $r_i(\sigma_i, \sigma_{\partial i})$  is given by Eq. (5.9).

To obtain Eq. (5.10), it has been assumed that  $P(\sigma_i, \sigma_{\partial i}) \approx \left[ \prod_{k \in \partial i} P(\sigma_k) \right] P(\sigma_i)$ . This factorization leads to a fully closed master equation in site probabilities  $P(\sigma_i)$ . In studies on epidemic spreading, this is called the Individual-Based Mean-Field (IBMF) approximation [27]; in other cases, it is simply called the mean-field approximation.

It should be noted that there are practically no techniques in the literature to study the dynamics of local search algorithms in K-SAT. Here, IBMF is useful as a reference to decide whether the effort to deduce and numerically integrate the CME translates into a non-trivial description of FMS. Fortunately, Fig. (5.5) shows that the CME significantly improves the much simpler results obtained from IBMF.

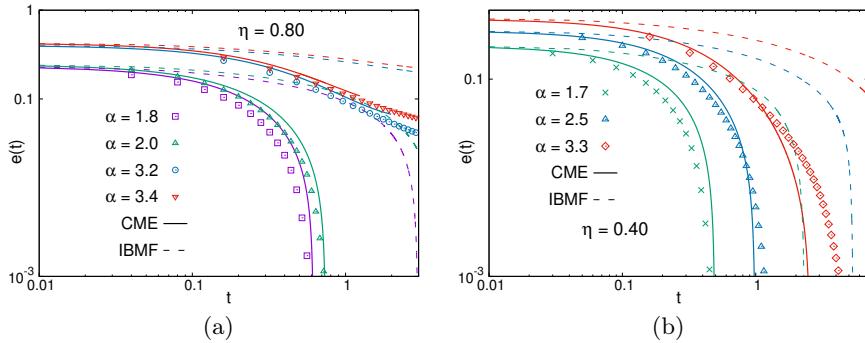


Figure 5.5: Comparison between CME (lines), FMS (points), and IBMF (dashed lines) in 3-SAT instances. Both panels show the average number of unsatisfied clauses per variable as a function of time. A logarithmic scale is used for both axes. Each point is the average of  $10^4$  FMS algorithm runs on instances of size  $N = 2000$ . **a)**  $\eta = 0.80$  was used. **b)**  $\eta = 0.40$ .

To estimate the location of the algorithmic transition with CME, it is sufficient to observe when the curvature of the  $e(t)$  functions changes in plots like (5.4a) and (5.4b), which are in logarithmic scale on both axes. In Appendix J, it is explained how this relates to solution times and how the value of  $\alpha_{\text{CME}}(\eta)$  is determined in practice.

Finally, Fig. (5.6) compares the values of  $\alpha_{\text{FMS}}(\eta)$  and  $\alpha_{\text{CME}}(\eta)$  for 3-SAT. This type of figure is often called a phase diagram, where the boundary between the SAT and UNSAT regions appears. For  $\eta > 0.3$ , the CME prediction of the algorithmic transition has the same qualitative shape as that obtained directly from the algorithm, although it is shifted to smaller values of  $\alpha$ . However, the CME cannot detect that near  $\eta = 0.3$ , the algorithm's performance is optimal. It continues to predict an increase in  $\alpha_{\text{CME}}(\eta)$ .

In the low  $\eta$  region, the numerical integration of the CME becomes more challenging as larger values of  $\alpha$  need to be explored. Indeed, when  $\alpha > 4$ , it is likely to encounter nodes with very high connectivities. Since the CME requires summing over

all neighbor configurations, evaluating the sums on the right-hand side of Eq. (5.3) is computationally expensive. For  $\eta = 0.1$ , there is low confidence in the reported value of  $\alpha_{\text{CME}}(\eta)$ , which is thus represented with a different point type. However, it is certain that the transition is to the right of this point, never to the left.

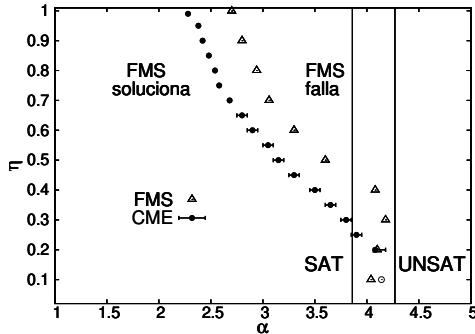


Figure 5.6: Comparison between phase diagrams for FMS and CME in 3-SAT. The boundary between SAT and UNSAT regions for the algorithm was obtained by running it on 100 problem instances for different values of  $\alpha$  for a time  $t = 10^5 N$ , with  $N = 50000$ . For each  $\eta$ , the value  $\alpha_{\text{FMS}}(\eta)$  where half of the runs solve the problem was determined (triangles). The CME phase diagram was obtained by numerically integrating it over 24 instances (filled circles) or at least 4 instances (unfilled circles) for each value of  $\alpha$ . The CME results were plotted in logarithmic scale graphs (see Fig. (5.4)) to determine the value of  $\alpha_{\text{CME}}(\eta)$  where half of them did not converge quickly to zero. For more details, see Appendix J. CME used  $N = 5000$  for  $\eta \geq 0.65$ ,  $N = 2000$  for  $0.4 \leq \eta < 0.65$ , and  $N = 500$  for  $\eta < 0.4$ .

The issues with the CME in this region are likely related to the presence of high correlations. Note that above  $\alpha_d = 3.86$ , it is expected that the dynamics slow down and point-to-set correlations diverge [51], as discussed earlier. In the following chapters, another example is presented where the CME fails precisely in the spin glass phase.

The next section examines the WalkSAT algorithm in a version for which a theoretical description with differential equations for local probabilities has already been written [21].

## 5.4 WalkSAT algorithm with greedy steps

WalkSAT is the most well-known stochastic local search algorithm used for K-SAT. Its first version was introduced in Ref. [101] by Papadimitriou and is known as *Random* WalkSAT. In each step of this algorithm, an unsatisfied clause is randomly selected, and then a variable within this clause is chosen to be flipped. It has since been demonstrated [21, 95] that *Random* WalkSAT can solve 3-SAT in linear time with

respect to the system size  $N$  for all values of  $\alpha$  less than  $\alpha_{RWSAT} = 2.7$ .

Subsequent work combined this random step with greedy steps where, after selecting an unsatisfied clause randomly, another criterion is used to choose the variable to flip. The most efficient version is due to Selman, Kautz, and Cohen [102] and is capable of solving 3-SAT in linear time up to  $\alpha_{SKC} = 4.2$  [52].

In this section, a less efficient formulation of WalkSAT is studied, which has been previously described by differential equations for local probabilities [21]. The corresponding pseudo-code, which depends on a parameter  $q \in [0, 1]$ , is presented below:

```

1: Let  $\vec{\sigma}$  be a random assignment of the variables
2: while  $\vec{\sigma}$  is not a solution do
3:   Select a clause  $C$  uniformly from the unsatisfied clauses
4:   Generate a random number  $x$  uniformly in  $[0, 1]$ 
5:   if  $x < q$  then
6:     Select a variable  $\sigma_i \in C$  uniformly from the  $K$  possible
7:   else
8:     Flip the variable that participates in the fewest satisfied clauses
9:   end if
10:  end while
```

It is known that this version of WalkSAT does not significantly improve the limit  $\alpha = 2.7$  of the *Random* WalkSAT. This allows predictions to be made with theoretical techniques away from the *hard* zone  $\alpha_d < \alpha < \alpha_s$  (with  $\alpha_d = 3.86$ ).

#### 5.4.1 Dynamic Independent Neighbor Approximation (DINA)

In subsection 4.2.2, the results of various levels of the hierarchical master equation system were compared with a technique from the literature: the Dynamic Independent Neighbor Approximation (DINA). Although subsection 4.2.2 studies the Ising ferromagnetic model on random regular graphs, the original approach of DINA was designed specifically to describe WalkSAT on 3-SAT instances [21].

Starting from approximations for the conditional probabilities equivalent to Eqs. (4.20) and (4.21), closed differential equations for the probabilities  $\hat{P}^t(u, s)$  of finding a variable in exactly  $s$  satisfied clauses and  $u$  unsatisfied clauses are derived:

$$\begin{aligned} \frac{d\hat{P}^t(u, s)}{dt} = & -r(u, s)\hat{P}^t(u, s) + \\ & + \left(\frac{1}{2^K - 1}\right)^u \sum_{k=0}^s \binom{u+k}{k} \left(1 - \frac{1}{2^K - 1}\right)^k r(s-k, u+k) \hat{P}^t(s-k, u+k) - \\ & - \frac{(K-1)}{2^K - 1} \frac{\langle sr(u, s) \rangle_t}{\langle s \rangle_t} \left[ s\hat{P}^t(u, s) - (s+1)\hat{P}^t(u-1, s+1) \right] - \\ & - (K-1) \frac{\langle ur(u, s) \rangle_t}{\langle u \rangle_t} \left[ u\hat{P}^t(u, s) - (u+1)\hat{P}^t(u+1, s-1) \right], \end{aligned} \quad (5.11)$$

where  $r(u, s)$  is the probability per unit time of flipping a variable that is in  $s$  satisfied and  $u$  unsatisfied clauses. Additionally, the notation  $\langle \cdot \rangle$  denotes the average of any quantity weighted by the probabilities  $\hat{P}^t(u, s)$ . This means that  $\langle \cdot \rangle \equiv \sum_u \sum_s [\cdot] \hat{P}^t(u, s)$ .

Note that  $\hat{P}^t(u, s)$  is the joint probability of a central variable  $\sigma$  and all its neighbors. There are as many values of  $\hat{P}^t(u, s)$  as there are combinations of the quantities  $(u, s)$  of unsatisfied and satisfied clauses in that set. The first term on the right-hand side of Eq. (5.11) represents the probability of flipping that central variable ( $\sigma \rightarrow -\sigma$ ) and contributes negatively to the derivative of  $\hat{P}^t(u, s)$ .

The second term accounts for all positive contributions to the derivative of  $\hat{P}^t(u, s)$  related to transitions  $-\sigma \rightarrow \sigma$ . When  $\sigma$  is flipped, all  $u$  unsatisfied clauses it belongs to become satisfied. Additionally,  $k \leq s$  of the remaining clauses will remain satisfied. Only  $s - k$  of them will become unsatisfied after the flip. Therefore, in the second line of Eq. (5.11), the probabilities  $\hat{P}^t(s - k, u + k)$  of having exactly  $s - k$  unsatisfied clauses and  $u + k$  satisfied clauses are included. These are multiplied by the binomial probability  $p(k) = \binom{u+k}{k} \left(\frac{1}{2^K-1}\right)^u \left(1 - \frac{1}{2^K-1}\right)^k$  and summed over all possible values of  $k$ . Note that  $2^K - 1$  is the number of possible configurations that satisfy a clause.

From the above description, it follows that the approximations (4.20) and (4.21) are now augmented by the assumption that all configurations satisfying a clause are equiprobable, which is generally not true.

The third and fourth lines of Eq. (5.11) contain contributions to the derivative due to flips of the neighbors of the central variable  $\sigma$ . They correspond to the variables in satisfied and unsatisfied clauses, respectively.

If the rules  $r(u, s)$  are known, Eq. (5.11) can be numerically integrated to obtain the temporal dependence of  $\hat{P}^t(u, s)$  and with it a prediction for the evolution of the number of unsatisfied clauses per variable  $e(t) = \langle u \rangle_t / K$ . Similarly to what was done in Section 5.3, in Ref. [21]  $r(u, s)$  for the WalkSAT algorithm with any parameter  $q$  is provided.

The probability of selecting an unsatisfied clause  $C$  is  $1/N e(t)$ . Recall that the algorithm can take either a random step ( $q$ ) or a greedy step ( $1 - q$ ). The probability that a specific variable is flipped in a random step is  $q u / N e(t) K$ . On the other hand, for it to be flipped in a greedy step, it must belong to fewer  $s$  satisfied clauses compared to the other variables in  $C$ . In general, we have

$$r(u, s) = \frac{q u}{e(t) K} + \frac{(1 - q) u}{e(t)} p_g^{\text{DINA}}(s), \quad (5.12)$$

where again the factor  $1/N$  has been omitted as it is absorbed by the time scale.

The quantity  $p_g^{\text{DINA}}(s)$  is the probability that a variable belonging to  $s$  satisfied clauses is selected in a specific unsatisfied clause, over the other  $K - 1$  variables in the clause. This is written as:

$$p_g^{\text{DINA}}(s) = \sum_{u_1=0}^{\infty} \dots \sum_{u_{K-1}=0}^{\infty} \sum_{s_1=s}^{\infty} \dots \sum_{s_{K-1}=s}^{\infty} \frac{\prod_{i=1}^{K-1} \hat{P}^t(u_i, s_i | U)}{1 + \sum_{i=1}^{K-1} \delta_{s_i, s}}, \quad (5.13)$$

where  $\delta_{s_i, s}$  is the Kronecker delta and  $\hat{P}^t(u_i, s_i | U)$  is the conditional probability that a variable belongs to  $u_i$  unsatisfied and  $s_i$  satisfied clauses given that at least one of the clauses is unsatisfied. Following the same spirit as the approximations (4.20) and (4.21), this probability is calculated as:

$$\hat{P}^t(u_i, s_i | U) = \frac{u_i \hat{P}^t(u_i, s_i)}{\langle u \rangle_t}. \quad (5.14)$$

This is sufficient to write the dynamic rule  $r(u, s)$  in Eq. (5.12). Results can then be obtained if initial conditions for  $\hat{P}^t(u, s)$  are defined. The initial configuration of all variables is chosen randomly, so:

$$\hat{P}^0(u, s) = e^{-K\alpha} \frac{(K\alpha)^{u+s}}{u!s!} \left(1 - \frac{1}{2^K}\right)^s \left(\frac{1}{2^K}\right)^u. \quad (5.15)$$

With this, it is possible to integrate the DINA equations for the average dynamics of the algorithm and obtain the temporal evolution of the average number of unsatisfied clauses per variable  $e(t)$ . The following section also adapts the CME to the description of the WalkSAT dynamics and compares its results with those of the DINA.

## 5.5 Cavity Master Equation and numerical results

To apply the CME to WalkSAT, it is sufficient to define the transition probabilities per unit time in a manner analogous to what was done with Eq. (5.9) for the FMS. In the definition of  $r_i(\sigma_i, \sigma_{\partial i})$  (see Eqs. (5.3) and (5.4)), information about the local structure of the interaction hypergraphs representing the K-SAT can be included. Indeed, the CME is written directly for a specific instance of the problem, which provides some advantages over DINA.

Thus, the rules  $r_i$  are:

$$r_i(\sigma_i, \sigma_{\partial i}) = \frac{q}{e(t)K} E_i(\sigma_i, \sigma_{\partial i}) + \frac{1-q}{e(t)} \sum_{a \subset \partial i} \left[ \prod_{j \in a} \frac{1 - J_j^a \sigma_j}{2} \right] p_g^{\text{CME}}(S_i(\sigma_i, \sigma_{\partial i})). \quad (5.16)$$

The quantity  $E_i(\sigma_i, \sigma_{\partial i})$  was already defined in Eq. (5.6). Additionally, the number  $S_i(\sigma_i, \sigma_{\partial i})$  of satisfied clauses to which variable  $\sigma_i$  belongs has been introduced:

$$S_i(\sigma_i, \sigma_{\partial i}) = \sum_{a \subset \partial i} \left\{ 1 - \prod_{j \in a} \frac{1 - J_j^a \sigma_j}{2} \right\}. \quad (5.17)$$

To simplify the notation, we will use  $S_i(\sigma_i, \sigma_{\partial i}) \equiv s_i$ . The probability  $p_g^{\text{CME}}(s_i)$  of flipping the variable  $\sigma_i$  within clause  $a$  is obtained by summing over all events where  $s_j \geq s_i$  for all variables  $j \in a \setminus i$ . It should be noted that when implementing the CME, the connectivities  $c_j$  of all neighbors of node  $i$  in the graph are available. To simplify computational calculations, a simple form for the probability that a neighbor  $j \in \partial i$  sharing an unsatisfied clause with  $i$  is in  $s_j$  satisfied clauses will be proposed. This will be simply the binomial  $P_{c_j}^t(s_j) = \binom{c_j-1}{s_j} (e(t)/\alpha)^{c_j-1-s_j} (1-e(t)/\alpha)^{s_j}$ . On one hand, the local information of the connectivities is preserved. On the other hand, the local probabilities  $P_{c_j}^t(s_j)$  are approximated using the average value of the number of unsatisfied clauses per variable  $e(t)$ .

The result is:

$$p_g^{\text{CME}}(s_i) = \sum_{\substack{\{s_j\}_{j \in a \setminus i} \\ s_j \geq s_i}}^{\infty} \left[ \prod_{j \in a \setminus i} P_{c_j}^t(s_j) \right] / (1 + \sum_{j \in a \setminus i} \delta_{s_j, s_i}). \quad (5.18)$$

With Eq. (5.18), the definition of the transition probabilities per unit time  $r_i(\sigma_i, \sigma_{\partial i})$  that appear in the CME is completed. It is important to highlight that this description of the WalkSAT algorithm considers local information about the structure of the boolean formula that is not accessible with DINA.

Fig. (5.7) presents the dynamics of WalkSAT and its description using DINA and CME for two different values of  $q$ . Although the WalkSAT results are obtained by averaging over different instances, there are no significant differences from instance to instance. This self-averaging property allowed saving computational effort and integrating the CME into a single instance of 3-SAT. It should be remembered that DINA, on the other hand, is already an analytical average over all instances.

For  $q = 1$  (Fig. (5.7a)), there are no greedy steps. In this limit, the *Random* WalkSAT algorithm is recovered, and Ref. [21] had already shown that DINA provides an accurate description of the steady state and the algorithmic transition ( $\alpha_{\text{RWSAT}} = 2.7$ ). Indeed, for the three values of  $\alpha$  shown, there is a match between *Random* WalkSAT and DINA for long times. At small  $\alpha$ , both converge rapidly to zero, while at larger  $\alpha$  they reach the same final value for the number of unsatisfied clauses per variable.

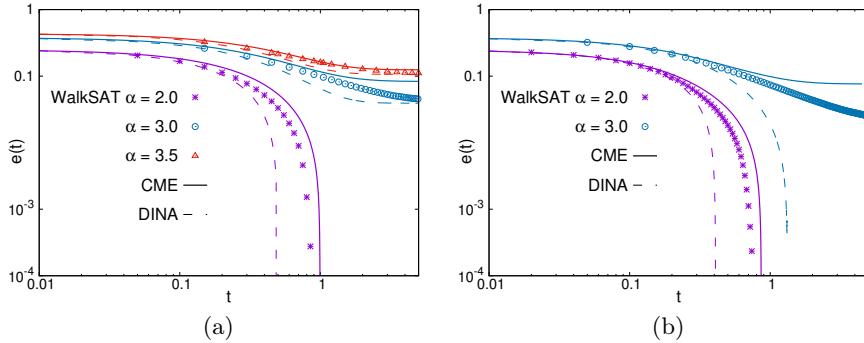


Figure 5.7: Dynamics of the WalkSAT algorithm in 3-SAT instances. The points represent the average of 50 instances of size  $N = 10000$  with 20 different initial conditions each. Predictions from DINA (dashed lines) and CME (solid lines) for the dynamics are shown. Both panels display the average number of unsatisfied clauses per variable as a function of time. A logarithmic scale is used for both axes. The numerical integration of the CME was performed for systems of sizes  $N = 200$  and  $N = 500$ , with one instance in each case. **a)** Random steps only ( $q = 1$ ). **b)** Presence of greedy steps ( $q = 0.5$ ).

The stationary values of  $e(t)$  predicted by the CME for  $\alpha = 3.0$  and  $\alpha = 3.5$  in Fig. (5.7a) do not match the behavior of WalkSAT. However, the CME description for the transient regime is not worse than that of DINA. For small  $\alpha$ , the continuous curve representing CME results is even appreciably closer to the algorithm.

When greedy steps appear, the situation changes. The WalkSAT dynamics for  $q = 0.5$  in Fig. (5.7b) deviate from the theoretical description provided by DINA. For  $q = 0.5$  and  $\alpha = 3.5$ , the corresponding stationary states are also different. Although WalkSAT is no longer able to solve the problem for  $\alpha = 3.5$ , DINA continues to predict that the algorithm solves the instances and converges rapidly to  $e = 0$ . For small  $\alpha$ , CME again has more accurate results. For  $\alpha = 3.5$ , although it does not quantitatively describe WalkSAT well, CME is able to predict that solutions should not be reached and converges to  $e \neq 0$ .

Similarly to how it was done for the FMS (see Appendix J), theoretical CME and DINA techniques can be used to predict the location of the WalkSAT algorithmic transition. Fig. (5.8) shows the results for 3-SAT. Note that the WalkSAT algorithm is much less efficient than FMS (compare with Fig. (5.6)). While the latter solves instances up to  $\alpha$  values close to  $\alpha = 4.23$ , WalkSAT is only effective below  $\alpha = 2.9$ . Its phase diagram is even distant from the hard region ( $\alpha_d = 3.86$ ), marked in the figure with a black vertical line.

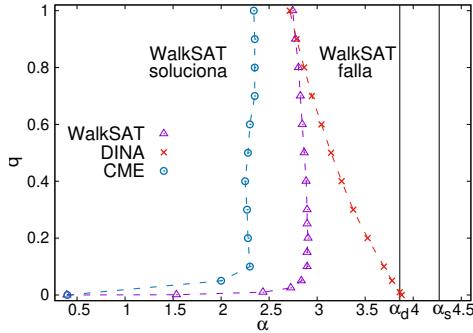


Figure 5.8: Comparison between phase diagrams for WalkSAT, CME, and DINA in 3-SAT. The boundary between the SAT and UNSAT zones of the algorithm was obtained by running it on 50 instances of the problem with 50 different initial conditions for a time  $t = 50N$ , with  $N = 10000$ . For each  $q$ , the value  $\alpha_{\text{WSAT}}(q)$  was determined at which half of the runs solve the problem (triangles). The CME phase diagram (circles) was obtained by numerically integrating it in 10 instances for each value of  $\alpha$ . CME results were plotted on logarithmic scale graphs (see Fig. (5.7)) to determine the value  $\alpha_{\text{CME}}(q)$  at which half of them did not converge rapidly to zero. For more details see Appendix J. For CME,  $N = 500$  was used. Since DINA is already an analytical average over all instances with the same  $\alpha$ , it was only necessary to generate the same type of logarithmic scale graphs and decide when a single curve  $e(t)$  obtained with DINA converges rapidly to zero.

The DINA's prediction of the transition has a qualitatively different shape compared to the algorithm's behavior. According to these equations, the efficiency of WalkSAT should always increase as  $q$  is decreased until it can solve instances close to the hard region. However, the algorithm is very ineffective for  $q$  close to zero.

On the other hand, the CME does describe this decrease in effectiveness for  $q$  close to zero. Its prediction for the algorithmic transition is qualitatively very similar to the actual behavior of WalkSAT. There is only a shift towards lower values of  $\alpha$ , which is, in fact, what happened for most values of the parameter  $\eta$  in FMS (see Fig. (5.6)).

## Partial conclusions

In this chapter, the CME is adapted to describe two local search algorithms for K-SAT: FMS and WalkSAT. For FMS applied to 3-SAT, the predictions for the algorithmic transition are qualitatively correct for high values of the algorithmic parameter  $\eta$ . The CME's issues arise in two different ways. First, for any  $\eta$ , the equations fail near the algorithmic transition  $\alpha_{\text{FMS}}(\eta)$ , although good results are obtained once sufficiently far from those problematic  $\alpha$  values. Second, it is important to note that for small  $\eta$ , CME does not qualitatively describe the behavior of FMS well. Here, changes in

the structure of solutions above  $\alpha_d$  and the corresponding slowdown of dynamics are expected to affect the theory's accuracy.

For WalkSAT, however, being far from the hard region ( $\alpha \ll \alpha_d$ ) for 3-SAT, CME achieves better results. The fact that CME performs better than DINA for describing WalkSAT indicates the importance of considering the local structure of instances. The dynamics with greedy steps are not well captured by DINA's average equations. However, for the Ising ferromagnetic model in subsection 4.2.2, accurate results were obtained with DINA. This suggests caution when applying average equations to models with local disorder in interactions.

The differences between the two algorithms help identify a source of difficulties for CME. The algorithmic transition of WalkSAT remains well away from the hard region  $\alpha_d < \alpha < \alpha_s$  for 3-SAT, and in this case, it qualitatively describes the entire phase diagram. However, FMS is a much more efficient algorithm and ventures into the region  $\alpha_d < \alpha < \alpha_s$  for 3-SAT. Here, the theoretical description made with CME also qualitatively fails. It can be assumed that the mechanisms controlling the efficiency of FMS for  $\alpha > \alpha_d$  are different from those governing the region  $\alpha < \alpha_d$ . This partially aligns with what is expected in the literature where  $\alpha_d$  was obtained as a limit for all local search algorithms but does not explain why FMS continues to work efficiently well beyond  $\alpha_d$ .

This chapter has applied CME to describe two local search algorithms for K-SAT. Its results are better than the few techniques available in the literature and qualitatively predict algorithmic transitions for broad regions of phase diagrams. Nevertheless, there are several limitations related to the approximations made to derive CME that prevent obtaining quantitatively correct results. The third part of this thesis introduces a new theoretical technique that performs better in several models: Approximate Conditional Dynamics.

# Part III

## Conditional Dynamic Approximation

## Chapter 6

# Spin glass dynamics

In the previous chapters, it has been shown that the cavity master equation is a powerful tool for describing the dynamics of discrete variables in continuous time. Its stationary state is equivalent to the cavity method in equilibrium when the transition probabilities per unit time satisfy detailed balance. Accurate average equations for different models can be written and results from the literature can be recovered. It has also been generalized to better account for the contribution of correlations through a hierarchical system of approximations. Moreover, it is capable of making qualitatively correct predictions for two local search algorithms used to solve K-SAT.

However, the uncontrolled approximations introduced in the derivation of CME impose limitations that are difficult to overcome. This technique fails when high temporal and spatial correlations arise. The ultimate testing ground for such situations is spin glass dynamics. Section 1.3 reviewed a portion of the literature that has studied this topic [55, 56, 60–64].

In models that exhibit a spin glass phase at low temperatures, such as the p-spin ferromagnetic model discussed in Section 1.1, the dynamics display very rich phenomenology. Below the transition temperature  $T_d$  to the spin glass, the relaxation time to equilibrium diverges and multiple time scales coexist. Long-range correlations cause qualitative changes in the system’s behavior.

This chapter precisely describes the spin glass dynamics for the p-spin ferromagnetic model. Appendix N shows that CME does not provide a good description of this dynamics. Here, there is no choice but to correct the CME or replace it with another theory capable of making more accurate predictions.

Appendix K contains a first attempt to modify the methodology, which does not improve the description of the dynamics under high correlation conditions. The results are also unfavorable for the Viana-Bray model of spin glass. Appendix N applies the same equations to the p-spin ferromagnetic model, with identical conclusions.

Next, Section 6.1 presents the derivation of a new technique that describes the dynamics of the p-spin ferromagnetic model spin glass much better: approximate conditional dynamics. It also includes numerical results corresponding to the model defined on regular random hypergraphs. Subsequently, Section 6.2 introduces some

theoretical corrections that enhance the accuracy of the new technique.

## 6.1 Conditional Dynamic Approximation

Recall that the Hamiltonian corresponding to the ferromagnetic  $p$ -spin model is given by  $H(\vec{\sigma}) = -\sum_{a=1}^M \prod_{i \in a} \sigma_i$  (see Section 1.1). The interaction is structured in groups of  $p$  variables called plaquettes. This type of model is usually represented using hypergraphs with variable nodes (denoted by  $i, j, \dots$ ) and factor nodes (denoted by  $a, b, \dots$ ).

As explained in Chapter 4, it is possible to write equations for the time derivatives of probabilities defined over any set of connected nodes of the hypergraph. For example:

$$\frac{d}{dt} P(\sigma_a) = - \sum_{i \in a} \sum_{\sigma_{\partial i \setminus a}} [r_i(\sigma_i, \sigma_{\partial i}) P(\sigma_{\partial i \setminus a}, \sigma_a) - r_i(-\sigma_i, \sigma_{\partial i}) P(\sigma_{\partial i \setminus a}, F_i[\sigma_a])] \quad (6.1)$$

$$\begin{aligned} \frac{d}{dt} P(\sigma_i, \sigma_{\partial i}) &= -r_i(\sigma_i, \sigma_{\partial i}) P(\sigma_i, \sigma_{\partial i}) + r_i(-\sigma_i, \sigma_{\partial i}) P(-\sigma_i, \sigma_{\partial i}) - \\ &- \sum_{b \subset \partial i} \sum_{j \in b \setminus i} \sum_{\sigma_{\partial j \setminus b}} [r_j(\sigma_j, \sigma_{\partial j}) P(\sigma_{\partial j \setminus b}, \sigma_{\partial i}, \sigma_i) - r_j(-\sigma_j, \sigma_{\partial j}) P(\sigma_{\partial j \setminus b}, F_j[\sigma_{\partial i}], \sigma_i)]. \end{aligned} \quad (6.2)$$

Equation (6.1) expresses the time derivative of the probability  $P(\sigma_a)$  of finding a specific configuration  $\sigma_a$  of the variables that make up plaquette  $a$ . On the other hand, the fundamental quantities in Equation (6.2) are the probabilities  $P(\sigma_i, \sigma_{\partial i})$  associated with the configuration of a spin and all its neighbors.

The symbol  $\partial i$  represents the set of variables that also belong to the plaquettes  $b$  in which node  $i$  appears. As in Chapter 5, in Equation (6.2) a plaquette  $b$  is considered to be a set of  $p$  variables. Therefore, it is correct to say that  $b$  is a *subset* of  $i \cup \partial i$ . To further simplify the notation, it is written simply as  $b \subset \partial i$ .

If a larger number of variables is chosen, a hierarchical system of differential equations is constructed that must be truncated at some point. The simplest approximation would be to neglect all connected correlations  $C_{ij} = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle$ , yielding closed equations for the time derivatives of the site probabilities  $P(\sigma_i)$ . The previous approximation, known as the Individual-Based Mean Field (IBMF) approximation, was already described in Section 5.3 (see Equation (5.10)). If the initial condition  $P^0(\sigma_i) = 1/2$  is taken regardless of the value of  $\sigma_i$ , the time derivatives of the IBMF vanish and the system remains with local probabilities  $P(\sigma_i) = 1/2$  at all times. The energy density is easily calculated and  $e = 0$  is obtained for all times  $t$ .

A less trivial approximation, similar to that used in the original CME, is:

$$\begin{aligned} P(\sigma_{a \setminus i}, \sigma_i, \sigma_{b \setminus i}) &= P(\sigma_{b \setminus i} | \sigma_i, \sigma_{a \setminus i}) P(\sigma_a) \approx P(\sigma_{b \setminus i} | \sigma_i) P(\sigma_a) \\ P(\sigma_{a \setminus i}, \sigma_i, \sigma_{b \setminus i}) &\approx \frac{P(\sigma_a) P(\sigma_b)}{P(\sigma_i)}. \end{aligned} \quad (6.3)$$

In Equation (6.3), the joint probability associated with two plaquettes sharing a spin  $\sigma_i$  appears. In the first line, the conditional probability  $P(\sigma_{b \setminus i} | \sigma_i, \sigma_{a \setminus i})$  was

approximated by  $P(\sigma_{b \setminus i} | \sigma_i)$ . This means that the configuration of  $\sigma_{a \setminus i}$  is irrelevant once the value of  $\sigma_i$  is known. Information about the variable at distance  $d = 1$  from  $\sigma_{b \setminus i}$  is sufficient to define the conditional probability.

This is reflected in the *conditional* connected correlations, obtained from the conditional probabilities:

$$\begin{aligned} C_{\sigma_i} &\equiv \langle \sigma_{a \setminus i} \sigma_{b \setminus i} \rangle_{\sigma_i} - \langle \sigma_{a \setminus i} \rangle_{\sigma_i} \langle \sigma_{b \setminus i} \rangle_{\sigma_i} \\ C_{\sigma_i} &= \sum_{\sigma_{a \setminus i}, \sigma_{b \setminus i}} \left[ \prod_{k \in a \setminus i} \sigma_k \right] \left[ \prod_{j \in b \setminus i} \sigma_j \right] P(\sigma_{a \setminus i}, \sigma_{b \setminus i} | \sigma_i) - \\ &- \left\{ \sum_{\sigma_{a \setminus i}} P(\sigma_{a \setminus i} | \sigma_i) \prod_{k \in a \setminus i} \sigma_k \right\} \left\{ \sum_{\sigma_{b \setminus i}} P(\sigma_{b \setminus i} | \sigma_i) \prod_{j \in b \setminus i} \sigma_j \right\}. \end{aligned} \quad (6.4)$$

In principle,  $C_{\sigma_i}$  may be different from zero. However, Equation (6.3) allows the factorization  $P(\sigma_{a \setminus i}, \sigma_{b \setminus i} | \sigma_i) = P(\sigma_{a \setminus i} | \sigma_i)P(\sigma_{b \setminus i} | \sigma_i)$ , which immediately forces  $C_{\sigma_i} = 0$  for both values of  $\sigma_i$ .

With the same initial conditions  $P(\sigma_i) = \frac{1}{2}$  for all sites  $i$  in a random regular hypergraph, one obtains a single equation for  $P(\sigma_a)$ . Something similar was discussed in subsection 4.2.2 for another model. Everything can be written in terms of the average probability  $\Phi \equiv \sum_{\sigma_a} P(\sigma_a) \delta(\prod_{k \in a} \sigma_k, -1)$ . The corresponding differential equation is:

$$\frac{1}{p} \frac{d\Phi}{dt} = - \sum_{u=0}^{c-1} \binom{c-1}{u} r_c(u+1) \Phi^{u+1} (1-\Phi)^{c-1-u} + \sum_{u=0}^{c-1} \binom{c-1}{u} r_c(u) \Phi^u (1-\Phi)^{c-u}, \quad (6.5)$$

where  $u$  is the number of unsatisfied interactions ( $\prod_{k \in a} \sigma_k = -1$ ) in which a variable participates. The positive number  $c$  is the connectivity of each variable in the regular hypergraph. It has been assumed that the transition probabilities per unit time  $r_c(u)$  depend only on  $u$  and  $c$ .

The equation following this Equation (6.5) will be referred to as the first-order approximate conditional dynamics (*CDA-1*) for reasons that will be clarified later. This equation can be integrated numerically over time to obtain a non-trivial relaxation of the energy density  $e(t)$  that already surpasses the results obtained with the simpler IBMF approximation. Appendix L shows that *CDA-1* is equivalent to the known Pair-Based Mean Field (PBMF) approximation [32] in the case of pair interactions.

The results for the Glauber dynamics [70] of the ferromagnetic  $p$ -spin model, with  $p = 3$ , are compared with Monte Carlo simulations in Fig. (6.1) in random regular hypergraphs with connectivity  $c = 3$ . In this case,  $r_i(\sigma_i, \sigma_{\partial_i}) \equiv r_c(u) = \frac{1}{2}(1 - \tanh[\beta(c - 2u)])$  is chosen. *CDA-1* is represented with dashed lines in the figure. For high temperatures (Fig. (6.1a)), the approximation works well for the transient regime and correctly provides the stationary values of the energy density. However, below the spin glass transition temperature  $T_d(c = p = 3) \approx 0.51$  (Fig. (6.1b)), *CDA-1* has very poor results and predicts a relaxation that is far from the behavior of the simulations.

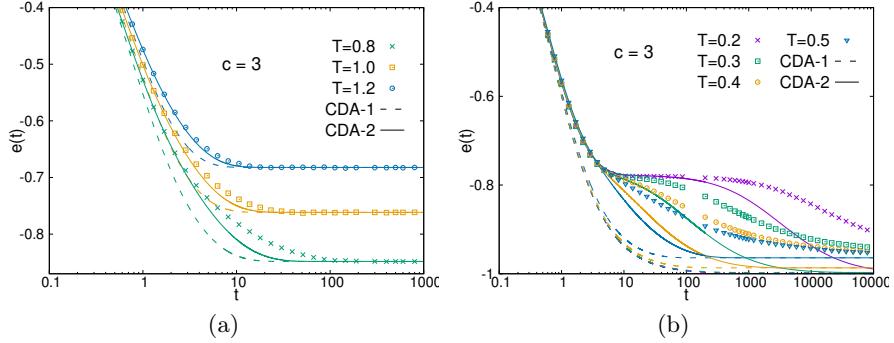


Figure 6.1: Time dependence of the energy density  $e(t)$  for the ferromagnetic  $p$ -spin model with  $p = 3$  defined on random regular hypergraphs with connectivity  $c = 3$ . Each point is an average of  $10^4$  Monte Carlo simulations with system size  $N = 10^4$ . The numerical integration results of *CDA-1* are represented with dashed lines, while the solid lines represent *CDA-2*. In each case, the system starts in a random configuration and is brought into contact with a thermal bath at fixed temperature  $T$ . **a)** Dynamics at high temperatures ( $T > T_d \approx 0.51$ ). **b)** Dynamics at low temperatures ( $T < T_d$ ).

At low temperatures, the simulations exhibit a two-step relaxation. The first step occurs over a short time and reaches an intermediate energy  $e_p \approx -0.78$ . The second step is slower and further decreases to  $e_d \approx -0.96$ . Since *CDA-1* is insufficient to explain this behavior, an additional step is necessary.

Analogous to approximation (6.3), to close Equation (6.2), one can write:

$$\begin{aligned} P(\sigma_{\partial j \setminus b}, \sigma_i, \sigma_{\partial i}) &= P(\sigma_{\partial j \setminus b} | \sigma_i, \sigma_{\partial i}) P(\sigma_i, \sigma_{\partial i}) \approx P(\sigma_{\partial j \setminus b} | \sigma_j, \sigma_{b \setminus j}) P(\sigma_i, \sigma_{\partial i}) \\ P(\sigma_{\partial j \setminus b}, \sigma_i, \sigma_{\partial i}) &\approx \frac{P(\sigma_j, \sigma_{\partial j})}{P(\sigma_i, \sigma_{b \setminus i})} P(\sigma_i, \sigma_{\partial i}) = \frac{P(\sigma_j, \sigma_{\partial j})}{\sum_{\sigma_{\partial j \setminus b}} P(\sigma_j, \sigma_{\partial j})} P(\sigma_i, \sigma_{\partial i}), \end{aligned} \quad (6.6)$$

where  $b$  is one of the plaquettes containing  $i$ , while  $j \neq i$  is another variable belonging to  $b$ .

Equation (6.6) leads to the second-order approximate conditional dynamics (*CDA-2*). It focuses on the conditional probabilities  $P(\sigma_{\partial j \setminus b} | \sigma_i, \sigma_{\partial i})$ , which are defined over the configurations of all nodes in the neighborhoods of  $i$  and  $j$ . By approximating this probability as  $P(\sigma_{\partial j \setminus b} | \sigma_j, \sigma_{b \setminus j})$ , the effect of nodes at a distance  $d = 3$  from  $\partial j \setminus b$  is neglected. Unlike *CDA-1*, the dependence on nodes at distance  $d = 2$  is retained. This procedure can be easily generalized for any  $d$  in the graph to obtain *CDA-3*, *CDA-4*, etc.

After combining Equations (6.2) and (6.6), a closed system of differential equations for  $P(\sigma_i, \sigma_{\partial i})$  is obtained.

$$\begin{aligned}
\frac{d}{dt} P(\sigma_i, \sigma_{\partial i}) = & -r_i(\sigma_i, \sigma_{\partial i}) P(\sigma_i, \sigma_{\partial i}) + r_i(-\sigma_i, \sigma_{\partial i}) P(-\sigma_i, \sigma_{\partial i}) - \\
& - \sum_{b \subset \partial i} \sum_{j \in b \setminus i} \sum_{\sigma_{\partial j \setminus b}} r_j(\sigma_j, \sigma_{\partial j}) \frac{P(\sigma_j, \sigma_{\partial j})}{\sum_{\sigma_{\partial j \setminus b}} P(\sigma_j, \sigma_{\partial j})} P(\sigma_i, \sigma_{\partial i}) + \\
& + \sum_{b \subset \partial i} \sum_{j \in b \setminus i} \sum_{\sigma_{\partial j \setminus b}} r_j(-\sigma_j, \sigma_{\partial j}) \frac{P(\sigma_j, \sigma_{\partial j})}{\sum_{\sigma_{\partial j \setminus b}} P(-\sigma_j, \sigma_{\partial j})} P(\sigma_i, F_j[\sigma_{\partial i}]). \quad (6.7)
\end{aligned}$$

Equation (6.7) makes no additional assumptions about the structure of the interactions. The rules  $r_i(\sigma_i, \sigma_{\partial i})$  for the dynamics can take any form depending on the spin at site  $i$  and its neighbors. Therefore, *CDA-2*, written as in Equation (6.7), is general-purpose.

When considering random regular hypergraphs with homogeneous initial conditions ( $P(\sigma_i) = \frac{1}{2}$ ), the treatment simplifies. It is possible to write a single equation for the probability  $P(u)$  of finding a node participating in exactly  $u$  unsatisfied interactions. In Appendix M, the differential equations for  $P(u)$  are obtained. It is also shown that the Dynamic Approximation of Independent Neighbors (DINA) [21, 22], which was previously discussed in subsection 4.2.2 and in Section 5.4, is a special case of *CDA-2*.

Figure (6.1a) illustrates the behavior of the energy density  $e(t)$  for  $p = c = 3$  at high temperatures. The results of *CDA-2* are represented with solid lines alongside Monte Carlo simulations (points) and *CDA-1* (dashed lines). In all cases, the description of the dynamics provided by *CDA-2* is more accurate. As the temperature decreases, its quality deteriorates in the transient regime but the steady state is predicted for all  $T > T_d$ .

It is more important to note in Figure (6.1b) that *CDA-2* qualitatively reproduces the two-step relaxation exhibited by Monte Carlo simulations. This means that considering local correlations more accurately is sufficient to obtain this non-trivial form of dynamics. These correlations begin to play an important role when the first plateau appears at an energy density  $e_p \approx -0.78$ . As the temperature decreases, the length of the plateau increases, and eventually, the system becomes stuck exactly at  $e_p$ .

Despite this significant result, *CDA-2* exits the plateau very quickly and reaches lower energies than those obtained in simulations. Similarly to the approach taken in Section 3.2, it can be shown that these final energies predicted by *CDA-2* correspond to the solution of the equilibrium message-passing equations (BP). Below  $T_d$ , BP is known to provide incorrect results [36].

These failures can be attributed to the following approximation inherent to any mean-field approximation such as *CDA-2*: at each moment  $t$ , such techniques assume that all relevant configurations can be explored in a short time and that the dynamics is a weighted average of these configurations. This generally holds at high energies, where correlations are weak. If the energy is low enough, correlations become important and the system takes more time to explore the configuration space. In the spin glass region, an extreme situation occurs. At some point, the timescale diverges and

relaxation simply takes so long that, from the observer's point of view, the system appears to freeze.

Figure (6.1b) also shows that Monte Carlo simulations reach an asymptotic energy  $e_d \approx -0.96$ , which is higher than the minimum possible ( $e = -1$  when all spins point up). In the following Section, this dynamic threshold energy is analytically obtained. For this, the equilibrium calculation of point-to-set correlation is extended (see Appendix A) to systems out of equilibrium.

## 6.2 New time scale

In Section 1.1, it was shown that in the ferromagnetic p-spin model, a transition to the spin glass occurs at a temperature  $T_d$  [43, 103]. To detect when this phase appears, the point-to-set correlation can be calculated. Ref. [48] proves that exactly at  $T_d$ , the associated correlation length diverges, and with it, the relaxation time to equilibrium.

In summary, this method reduces to studying the relationship between a *root* variable  $\sigma_i$  and the variables  $\vec{\sigma}_l$  that are at a distance  $l$  from  $\sigma_i$ . For a given configuration of  $\vec{\sigma}_l$ , the message-passing equations (BP) can be applied to calculate the expected value  $m^l$  of the *root* spin  $\sigma_i$ . Appendix A provides more details on how to obtain the distribution of expected values  $Q^l(m^l)$ , which is a function of the distance  $l$  and is defined over the set of values  $m^l$  obtained by varying the *boundary conditions*  $\vec{\sigma}_l$ . The important magnitude is then the point-to-set correlation, which for the p-spin model can be expressed as  $C^l = \int dm Q^l(m) m$ .

Here, this method is extended to consider systems that are completely out of equilibrium. With a technique for describing the dynamics such as *CDA-2* from the previous section, it is possible to obtain the local probabilities:

$$P(S_{a \setminus i}, S_{b \setminus i} | \sigma_i) = \sum_{\sigma_{a \setminus i}} \sum_{\sigma_{b \setminus i}} P(\sigma_{a \setminus i}, \sigma_{b \setminus i} | \sigma_i) \delta(S_{a \setminus i}, \prod_{j \in \partial a \setminus i} \sigma_j) \delta(S_{b \setminus i}, \prod_{k \in \partial b \setminus i} \sigma_k) \quad (6.8)$$

where  $\delta(x, y)$  is the Kronecker delta evaluated at  $x$  and  $y$ .

With the information contained in the local probabilities from Eq. (6.8), self-consistent equations can be written for the cavity distributions  $Q^l(\mu)$  and  $Q^l(\hat{\mu})$ . The distribution  $Q^l(\mu)$  is taken over the mean values  $\mu$  of the *root* variable in an auxiliary graph where an interaction involving the variable has been removed. That is, a factor node or hyperedge in the hypergraph containing  $\sigma_i$  is removed. On the other hand, to calculate  $Q^l(\hat{\mu})$ , all hyperedges containing  $\sigma_i$  except one are removed.

Thus, a set of iterative equations can be written where  $Q^l(\mu)$  depends on the contribution of  $c - 1$  hyperedges, each of which distributes according to  $Q^l(\hat{\mu})$ . Similarly, the equation for the distribution  $Q^l(\hat{\mu})$  of a single factor node depends on  $p - 1$  contributions that distribute according to  $Q^l(\mu)$ . The equations are:

$$\begin{aligned} Q^l(\mu) &= \sum_{S_1, \dots, S_{c-1}} \hat{\pi}(S_1, \dots, S_{c-1}) \int \left[ \prod_{i=1}^{c-1} d\hat{\mu}_i \hat{Q}^{l,S_i}(\hat{\mu}_i) \right] \delta(\mathcal{F}_1[\mu, \{\hat{\mu}_i\}_{i=1}^{c-1}]) \quad (6.9) \\ \hat{Q}^{l,S}(\hat{\mu}) &= 2^{-p+2} \sum_{\sigma_1, \dots, \sigma_{p-1}} \delta(S, \prod_{j=1}^{p-1} \sigma_j) \int \left[ \prod_{j=1}^{p-1} d\mu_j Q^{l-1}(\sigma_j \mu_j) \right] \delta(\mathcal{F}_2^S[\hat{\mu}, \{\mu_j\}_{j=1}^{p-1}]) \end{aligned}$$

In Eqs. (6.9) and (6.10), it is again assumed that  $P(\sigma_i) = \frac{1}{2}$  for all  $i = 1, \dots, N$ , and  $\hat{\pi}(S_1, \dots, S_{c-1}) \equiv P(S_1, \dots, S_{c-1} \mid \sigma = 1)$  is defined. These equations involve Dirac delta functions  $\delta(z)$  (not to be confused with the Kronecker delta  $\delta(x, y)$ ) which impose the relationships:

$$\mathcal{F}_1[\mu, \{\hat{\mu}_i\}_{i=1}^{c-1}] = \frac{1+\mu}{2} - \frac{\prod_{i=1}^{c-1} (1 + \hat{\mu}_i)}{\sum_{\sigma} [\prod_{i=1}^{c-1} (1 + \sigma \hat{\mu}_i)]} \quad (6.11)$$

$$\mathcal{F}_2^S[\hat{\mu}, \{\mu_j\}_{j=1}^{p-1}] = \hat{\mu} - \langle S \rangle \prod_{j=1}^{p-1} \mu_j, \quad (6.12)$$

where  $\langle S \rangle = \sum_S S \hat{\pi}(S)$ .

The solutions of Eqs. (6.9) and (6.10) can be explored to locate the divergence of the point-to-set correlation distance (see Appendix A). To simplify numerical calculations and handle any value of connectivity  $c$ , the approximation  $\hat{\pi}(S_1, \dots, S_{c-1}) \approx \prod_{i=1}^{c-2} \hat{\pi}(S_i, S_{i+1}) / \prod_{i=2}^{c-2} \hat{\pi}(S_i)$  is proposed for any local probability  $\hat{\pi}$ . The probabilities associated with a pair of factor nodes ( $\hat{\pi}(S_i, S_{i+1})$ ) can be written exactly in terms of two parameters: the energy density  $e$  and the conditional correlation between a pair of neighboring hyperedges  $C_{\sigma_i=1} \equiv C$  (see Eq. (6.4)). Figure (6.2a) shows the position of the divergence of the point-to-set correlation distance in the  $(e, C)$  plane. At this point, the physical time scale associated with the dynamics diverges.

Figure (6.2b) presents the same data from Section 6.1 but in a parametric plot of time in the  $(e(t), C(t))$  plane. In this case, the time scales become irrelevant, and a good match is observed between Monte Carlo simulations and the analytical results based on *CDA-2*. This indicates that an appropriate redefinition of the microscopic time scale in the dynamic equations could provide a solution much closer to the physical behavior of the system (see Figure (6.1b)).

The inset graph in Figure (6.2b) approximates the region reached for very long times. The black line with points marks the location where the point-to-set correlation distance diverges. As noted by Ref. [48], this is associated with the divergence of the physical time scale, and Monte Carlo simulations cannot cross to the left of this line. They stop before crossing into the non-ergodic region where it is not possible to visit the relevant configurations in a finite time (see Figure (6.2a)). In contrast, *CDA-2* produces trajectories that fully enter this non-ergodic region, clearly ignoring the dynamic transition to spin glass.

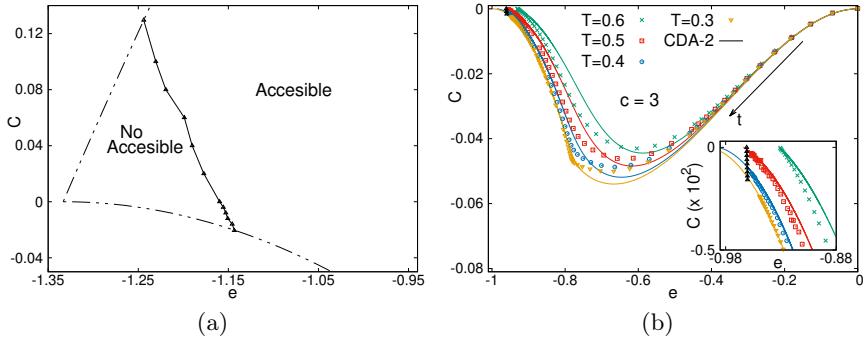


Figure 6.2: **a)** Phase diagram in the  $(e, C)$  plane for the ferromagnetic  $p$ -spin model with  $p = 3$  and  $c = 4$ . The solid line with points marks the critical line where the time scale associated with relaxation to equilibrium diverges. The dashed lines are mathematical bounds for  $e$  and  $C$  derived from the relations  $0 \leq \hat{\pi}(S_1, S_2) \leq 1$ . **b)** Parametric plots  $(e(t), C(t))$  for  $p = c = 3$  at low temperatures. The black arrow indicates the direction in which time  $t$  progresses. Each point is the average of  $10^4$  Monte Carlo simulations with a system size  $N = 10^4$ . Numerical integration results from *CDA-2* are shown as solid lines. In each case, the system is placed in a random initial configuration and evolves in contact with a thermal bath at a fixed temperature  $T$ .

The results from Eqs. (6.9) and (6.10) allow for the independent calculation of the energy density to which Monte Carlo simulations converge for long times. Figure (6.2b) shows that this value  $e_d$  depends very weakly on the correlation  $C$  and can be taken as  $e_d \approx -0.96$ , confirming what was observed in Figure (6.1b).

Next, a new timescale will be proposed for the range  $e_d < e < e_p$ , taking advantage of the fact that both  $e_p$  and  $e_d$  were calculated analytically. Recall that  $e_p$  is the energy density of the plateau where the system arrives after the first relaxation. Since this is the point where the Monte Carlo algorithm blocks at zero temperature ( $T = 0$ ), it is expected that this is precisely where the dynamics encounters the first finite-sized *energy barriers*.

Indeed, when  $T = 0$ , the system cannot afford to make local changes to the values of the variables that increase its energy. When Monte Carlo simulations reach the energy density  $e_p$ , they visit configurations of the variables from which the only way to escape is by slightly increasing the energy. These are the aforementioned *energy barriers* that block the dynamics. The slowdown observed at finite temperatures  $T < T_d$ , where a plateau near  $e_p$  still occurs, is probably due to the complexity of this energy landscape in which the system evolves. Ref. [104] supports the idea that this is where *entropic barriers* start to act, forcing the system to remain for a long time exploring similar configurations until it finds one that significantly reduces the energy.

Following the same description in terms of exploring minima of a dynamically relevant probability distribution, as the dynamics continues to decrease  $e$  at finite temperature  $T$ , the energy barriers increase in size. It becomes increasingly difficult to escape from a local minimum. Eventually, it becomes impossible to continue decreasing. At the energy density  $e_d$ , the barriers become extensive in the limit of large system size  $N$  and the relaxation halts.

The redefinition of the timescale for energy densities such that  $e_d < e < e_p$  can be argued as follows: in a mean-field approximation like *CDA-2*, it is assumed that at each time  $t$ , the system behaves as an average taken over several dynamically relevant configurations. While the exploration time is short, the assumption that this average represents everything instantaneously is reasonable. However, as barriers appear, it is no longer possible for the dynamics to quickly visit several different configurations.

One way to include this effect in the equations is to consider a slower timescale. The system has to take more time to explore the accessible configurations and evolve. Therefore, what is incorrect is not the probability distribution obtained with *CDA-2*, but rather the assumption that it represents the system over a short time. The same idea has extreme consequences when the *CDA-2* trajectories approach the critical line (see Figure (6.2b)). At that point, the relaxation timescale diverges, and if this is added to *CDA-2*, its dynamics also stops.

As already discussed, the microscopic timescale  $\hat{\tau}(e)$  only needs to be modified below  $e_p$ , when barriers appear. It will be assumed that the exploration of accessible configurations slows down by a factor of  $\exp(\Delta_S)$ . The entropic barrier  $\Delta_S$  is constant with respect to time and does not depend on temperature. Finally,  $\hat{\tau}(e)$  must diverge at  $e_d$ , and it will be simply assumed to have a power-law dependence  $(e - e_d)^{-\gamma}$ .

Thus, the *ansatz* for the effective timescale is:

$$\hat{\tau}(e) = \begin{cases} 1 & e_p < e \\ e^{\Delta_S} \left( (e_p - e_d) / (e - e_d) \right)^\gamma & e_d < e < e_p \\ \infty & e < e_d \end{cases} \quad (6.13)$$

This *ansatz* depends on two values of the energy density ( $e_p$  and  $e_d$ ) and two numbers  $\Delta_S$  and  $\gamma$ . As mentioned,  $e_p$  is calculated with the *CDA-2* dynamics at  $T = 0$  and  $e_d$  is obtained with Eqs. (6.9) and (6.10). Therefore,  $\Delta_S$  and  $\gamma$  are the only free parameters. These are adjusted to reproduce the results of the Monte Carlo simulations. It should be noted that neither depends on temperature and their values are not strongly affected by the connectivity  $c$  of the interaction hypergraph.

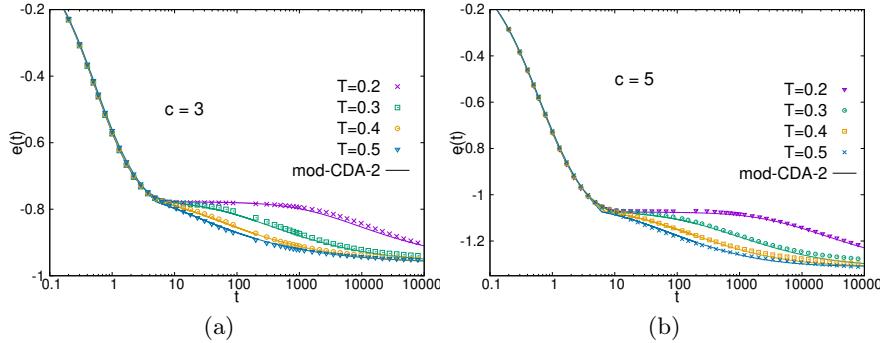


Figure 6.3: Temporal evolution of the energy density  $e(t)$  in the ferromagnetic  $p$ -spin model with  $p = 3$  at low temperatures. Each point represents the average of  $10^4$  Monte Carlo simulations with system size  $N = 10^4$ . The numerical integration of *CDA-2* included in this case the effective timescale defined in Eq. (6.13) and is shown as continuous lines. In each case, the system is placed in a random initial configuration and evolves in contact with a thermal bath at a fixed temperature  $T$ . **a)** The connectivity of the hypergraph is  $c = 3$  and the adjusted parameters are  $\Delta_S = 0.15$  and  $\gamma = 2$ . **b)** The connectivity of the hypergraph is  $c = 5$  and the adjusted parameters are  $\Delta_S = 0.17$  and  $\gamma = 2$ .

In Figure (6.3), the Monte Carlo simulations and the results of *CDA-2* with the appropriate timescales added are compared. There is a very good agreement between the actual physical dynamics and the theoretical predictions. Two different connectivities ( $c = 3$  and  $c = 5$ ) were considered to show the robustness of the dynamic description provided by *CDA-2*. Although  $\Delta_S$  varies slightly with  $c$ , the parameter  $\gamma$  is stable.

The value of  $\gamma$  is between the theoretically calculated limits in Refs. [48, 50]. In particular, it is tempting to compare it with the results for Monte Carlo simulations reported in Ref. [50] ( $\gamma_{MC} \approx 3.2$ ). However, it should be noted that all those predictions were made with an equilibrium probability distribution for the configurations, while here a completely out-of-equilibrium process is studied. The latter is prone to making more efficient decisions to relax to  $e_d$  and it is not surprising that the corresponding timescale divergence is less severe.

## Partial conclusions

In this chapter, a new technique was presented for studying the dynamics of discrete variables on random graphs, known as Approximate Conditional Dynamics. The CDA reduces to known results in the literature for some specific models and provides a more accurate description of the spin glass dynamics of the ferromagnetic  $p$ -spin model compared to the cavity master equation. With the redefinition of the timescale,

it is possible to fit the numerical results very well to Monte Carlo simulations.

In Section 6.1, important qualitative differences between the behavior of *CDA-1* and *CDA-2* at low temperatures are observed. It was necessary to reach the second level of approximation to reproduce the existence of an intermediate plateau in the temporal evolution of the energy density. *CDA-2* completely stops at this plateau at temperature  $T = 0$ . Including the effect of local correlations between different plaquettes is sufficient to accurately predict the energy value at which the spin glass dynamics is blocked at zero temperature.

The CDA has close relationships with other methods in the literature. Its first level of approximation, *CDA-1*, reduces to the Pair-Based Mean Field Approximation (PBMF) in epidemic spreading on graphs with pairwise interactions. The second level of approximation, *CDA-2*, can be reduced to the average equations known as the Independent Neighbors Dynamic Approximation (DINA) proposed for the ferromagnetic p-spin model.

The next chapter returns to local search algorithms for K-SAT and shows that the average equations obtained from the CDA are more accurate than those corresponding to CME. This reinforces the idea presented here of the superiority of CDA and allows for better profiling of the prospects for future work beyond this thesis.



## Chapter 7

# Average case equations in K-SAT

In this last chapter, we return to the description of local search algorithms used to solve K-SAT instances. To do this, we primarily use two theoretical methods introduced in the thesis: the Conditional Dynamic Approximation (CDA) and the Cavity Master Equation (CME). While Appendix N shows that *CDA-2* provides more accurate predictions about the spin glass dynamics of the p-spin ferromagnetic model than *CME-2*, it is unclear what happens with other models.

The main objective of this chapter is to compare the results of *CDA-2* with those of *CME-2* in describing the FMS algorithm (see Section 5.3). As explained in Chapter 5, numerical integration of these differential equations is computationally costly in the case of K-SAT. Therefore, in Section 7.1, we develop average equations that represent both techniques. Then, numerical results are presented to help decide which of them to continue applying in future studies on algorithm dynamics. Section 7.2 contains new predictions about the phase diagram of FMS that complement those presented in Section 5.3.

### 7.1 Derivation of the equations

To obtain average equations, some simplifications must be made while trying to retain the essential aspects of the local interaction structure. This operation is especially complicated in K-SAT. In the hypergraph that represents each Boolean formula (see Chapter 5), the variable nodes  $\sigma_i$  have different connectivities and their couplings  $J_i^a$  with factor nodes (clauses) are random variables with two values:  $J_i^a = \pm 1$ . Recall the expression for the problem Hamiltonian from Eq. (5.2).

This disorder can be difficult to capture with average equations. In Section 5.4, the average description provided by the Dynamic Independent Neighborhood Approximation (DINA) for the WalkSAT algorithm [21] was compared with the predictions of the CME applied to a given instance of K-SAT. Qualitatively, the CME outperformed

the DINA results in that case.

Therefore, care must be taken when deriving average equations from *CDA-2*:

$$\begin{aligned} \frac{d}{dt} P_{c_i}(\sigma_i, \sigma_{\partial i}) = & -r_{c_i}(\sigma_i, \sigma_{\partial i}) P_{c_i}(\sigma_i, \sigma_{\partial i}) + r_{c_i}(-\sigma_i, \sigma_{\partial i}) P_{c_i}(-\sigma_i, \sigma_{\partial i}) - \\ & - \sum_{b \subset \partial i} \sum_{j \in b \setminus i} \sum_{\sigma_{\partial j} \setminus b} r_{c_j}(\sigma_j, \sigma_{\partial j}) \frac{P_{c_j}(\sigma_j, \sigma_{\partial j})}{\sum_{\sigma_{\partial j} \setminus b} P_{c_j}(\sigma_j, \sigma_{\partial j})} P_{c_i}(\sigma_i, \sigma_{\partial i}) + \\ & + \sum_{b \subset \partial i} \sum_{j \in b \setminus i} \sum_{\sigma_{\partial j} \setminus b} r_{c_j}(-\sigma_j, \sigma_{\partial j}) \frac{P_{c_j}(\sigma_j, \sigma_{\partial j})}{\sum_{\sigma_{\partial j} \setminus b} P_{c_j}(-\sigma_j, \sigma_{\partial j})} P_{c_i}(\sigma_i, F_j[\sigma_{\partial i}]). \end{aligned} \quad (7.1)$$

In Eq. (7.1), it is highlighted that the joint probability  $P_{c_i}(\sigma_i, \sigma_{\partial i})$  is defined over a central node  $i$  with a certain connectivity  $c_i$ . The form of the transition probabilities per unit of time  $r_{c_i}$  generally also depends on  $c_i$ . Otherwise, the equation is identical to Eq. (6.7).

By averaging over all the connectivities of the nodes  $j \in \partial i$ , the equation no longer corresponds to a specific site in a hypergraph. The average probability  $P_{c_i} \equiv P_c$  will be parameterized taking into account the configurations of the  $c$  clauses to which it is associated.

Let:

- $u$ : The number of unsatisfied clauses.  $u = \sum_{a \in \partial i} \prod_{k \in a} (1 - J_k^a \sigma_k) / 2$
- $s_0$ : The number of clauses that become unsatisfied once the value of  $\sigma_i$  is inverted. That is,  $s_0 = \sum_{a \in \partial i} (1 + J_i^a \sigma_i) / 2 \prod_{k \in a \setminus i} (1 - J_k^a \sigma_k) / 2$

In most algorithms, the probability per unit of time  $r_{c_i}(\sigma_i, \sigma_{\partial i})$  of flipping the variable  $\sigma_i$  can be written as a function of the numbers  $u$  and  $s_0$ . For FMS, for example, we have:

$$r(u, s_0) = \frac{u}{K e(t)} \min\{\eta^{s_0-u}, 1\}, \quad (7.2)$$

where  $e(t)$  is the average number of unsatisfied clauses per variable.

Given the relevance of  $u$  and  $s_0$  in the dynamics, the following *ansatz* can be proposed:  $P_c \equiv P_c(u, s_0) / (2^K - 2)^{c-u-s_0}$ . This assumes that the  $2^K - 2$  configurations of the clauses that do not contribute to the quantities  $u$  and  $s_0$  are equiprobable and therefore each appears with a probability of  $1/(2^K - 2)$ . The unknown becomes the non-trivial probability distribution  $P_c(u, s_0)$  of finding  $u$  unsatisfied clauses and  $s_0$  clauses that will become unsatisfied if the central variable is flipped when it participates in exactly  $c$  clauses.

After inserting the previous *ansatz* into Eq. (7.1), an equation for the new probabilities  $\hat{P}_c(u, s_0) \equiv P_c(u, s_0) e^{-\alpha K} (\alpha K)^c / u! s_0! (c - u - s_0)!$  can be found, which resembles the Dynamic Independent Neighborhood Approximation (DINA) for K-SAT (see Eq. (5.11)):

$$\begin{aligned} \frac{d\hat{P}_c(u, s_0)}{dt} = & -r(u, s_0)\hat{P}_c(u, s_0) + r(s_0, u)\hat{P}_c(s_0, u) - \\ & -(K-1)\langle r \rangle_U \left( u\hat{P}_c(u, s_0) - (u+1)\hat{P}_c(u+1, s_0) \right) - \\ & -\frac{(K-1)}{2^K - 2} \langle r \rangle_{S_0} \left( (c-u-s_0)\hat{P}_c(u, s_0) - (c-u-s_0+1)\hat{P}_c(u-1, s_0) \right) - \\ & -(K-1)\langle r \rangle_S \left( s_0\hat{P}_c(u, s_0) - (s_0+1)\hat{P}_c(u, s_0+1) \right) - \\ & -\frac{(K-1)}{2^K - 2} \langle r \rangle_S \left( (c-u-s_0)\hat{P}_c(u, s_0) - (c-u-s_0+1)\hat{P}_c(u, s_0-1) \right). \end{aligned} \quad (7.3)$$

The first line of Eq. (7.3) contains the contributions to the derivative that are due to the flips of the variable  $\sigma_i$  in Eq. (7.1). That is, it is obtained by inserting the *ansatz* into the first line of Eq. (7.1). The other four lines correspond to the flips of the neighbors of  $\sigma_i$ .

In Eq. (7.3), averages  $\langle \cdot \rangle_\Omega$  calculated with the conditional probability distribution  $\hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | \Omega)$  have been introduced, where  $\Omega$  can take three values:  $U$ ,  $S_0$ , and  $S$ . The number  $\gamma$  is equal to the connectivity  $c$  minus one. When  $\Omega = U$ , the variable belongs to at least one unsatisfied clause. The quantity  $P_\gamma(\tilde{u}, \tilde{s}_0 | U)$  is interpreted as the conditional probability that it belongs to other  $\tilde{u}$  unsatisfied clauses and  $\tilde{s}_0$  clauses that become unsatisfied when the variable is flipped. Similarly, in  $\hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | S_0)$ , it is conditioned that at least one clause becomes unsatisfied when the variable is flipped. Finally, when  $\Omega = S$ , at least one clause that was already satisfied maintains its condition.

More explicitly, the averages are written as:

$$\langle r \rangle_\Omega = \sum_{\gamma=0}^{\infty} e^{-\alpha K} \frac{(\alpha K)^\gamma}{\gamma!} \sum_{\tilde{u}=0}^{\gamma} \sum_{\tilde{s}_0=0}^{\gamma-\tilde{u}} r(\tilde{u} + \delta_{\Omega, U}; \tilde{s}_0 + \delta_{\Omega, S_0}) \hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | \Omega) \quad (7.4)$$

The relations that actually close the system of average master equations are those that express the conditional probabilities in terms of  $\hat{P}_c(u, s_0)$ . In the case of this average *CDA-2*, we have:

$$\hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | U) = \frac{(\tilde{u}+1) \hat{P}_{\gamma+1}(\tilde{u}+1, \tilde{s}_0)}{\sum_{u'=0}^{\gamma} \sum_{s'_0=0}^{\gamma-u'} (u'+1) \hat{P}_{\gamma+1}(u'+1, s'_0)} \quad (7.5)$$

$$\hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | S_0) = \frac{(\tilde{s}_0+1) \hat{P}_{\gamma+1}(\tilde{u}, \tilde{s}_0+1)}{\sum_{u'=0}^{\gamma} \sum_{s'_0=0}^{\gamma-u'} (s'_0+1) \hat{P}_{\gamma+1}(u', s'_0+1)} \quad (7.6)$$

$$\hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | S) = \frac{(\gamma - \tilde{u} - \tilde{s}_0 + 1) \hat{P}_{\gamma+1}(\tilde{u}, \tilde{s}_0)}{\sum_{u'=0}^{\gamma} \sum_{s'_0=0}^{\gamma-u'} (\gamma - u' - s'_0 + 1) \hat{P}_{\gamma+1}(u', s'_0)}. \quad (7.7)$$

It is precisely Eqs. (7.5), (7.6), and (7.7) that define the average *CDA-2*. Other ways of expressing the conditional probabilities lead to other types of approximations. The formulation of the average *CME-2*, for example, is obtained by substituting  $\hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | \Omega)$  with the cavity conditional probability  $\hat{p}_\gamma(\tilde{u}, \tilde{s}_0 \| U)$  which is obtained by integrating an auxiliary differential equation. Appendix O explains how to obtain the average version of *CME-2* in K-SAT.

Eq. (7.3) can be further simplified by making a new *ansatz* for  $\hat{P}_c(u, s_0)$ . Considering that the only relevant parameter is the number of unsatisfied clauses  $u$  and that all other configurations are equivalent, we have the relation:  $\hat{P}_c(u, s_0) \equiv \hat{P}_c(u) [1/(2^K - 1)]^{s_0} [1 - 1/(2^K - 1)]^{c-u-s_0}$ . Appendix P shows that after inserting this last *ansatz* into Eq. (7.3) and redefining the conditional probabilities, the equations of DINA are recovered (see Eq. (5.11)).

The next section presents the numerical results of the three mentioned average equations: the *CDA-2*, the *CME-2*, and the DINA. Their respective predictions on the dynamics and phase diagram of the FMS algorithm serve to compare these three techniques and make recommendations for future work.

## 7.2 Average phase diagrams

Figure (7.1a) shows the results of the numerical integration of Eq. (7.3) for the dynamics of FMS (see Eq. (7.2)). A time dependence of the number of unsatisfied clauses per variable  $e(t)$  is obtained that is quantitatively far from the actual behavior of FMS runs. However, for the value of the FMS parameter  $\eta$  used in the figure, the average *CDA-2* does provide a qualitatively acceptable description. For small values of  $\alpha$  where the algorithm quickly finds solutions, the theoretical curves also converge rapidly to  $e = 0$ . When  $\alpha$  is larger, a finite number of unsatisfied clauses is reached for long times in both cases. This means that, at least for  $\eta = 0.7$ , the prediction that can be made with the average *CDA-2* about the algorithmic transition is close to the real transition of FMS (see Appendix J).

On the other hand, Figure (7.1b) shows the results of the average *CME-2* for the same value of  $\eta$ . The corresponding equations are very similar to Eq. (7.3) and are deduced in Appendix O. When compared with the results presented in Figures (5.4) and (5.5) of Chapter 5, it is clear that information is lost when the analytical average over instances is performed. It should be remembered that to produce the figures in Chapter 5, the first-level approximation equations were integrated on random hypergraphs. The *CME-1* (or simply CME) was written in that case for individual instances of 3-SAT and then numerically averaged. The predictions were quantitatively closer to the algorithm's dynamics.

The average *CME-2* does not describe the qualitative behavior of the algorithm. It predicts that for all the values of  $\alpha$  represented in Figure (7.1b), a solution should be found in a short time, which is evidently not true. This indicates that, for  $\eta = 0.7$ , the average *CME-2* overestimates the value of the algorithmic transition  $\alpha$ .

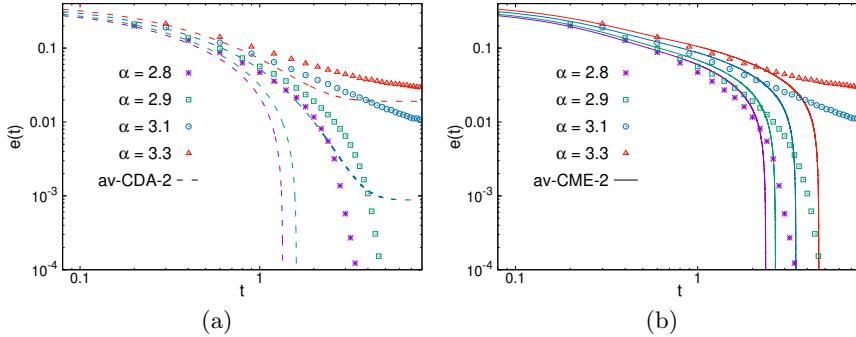


Figure 7.1: Comparison between the dynamics of the FMS algorithm applied to 3-SAT instances and the results of the numerical integration of the average equations for *CME-2* (solid lines) and *CDA-2* (dashed lines). The time evolution of the number of unsatisfied clauses per variable is represented. Both axes are on a logarithmic scale. Each point is the average of 100 runs of the algorithm on instances of size  $N = 10^4$  and different values of  $\alpha$ . In both panels,  $\eta = 0.7$  was taken. **a)** Comparison between the average *CDA-2* and the dynamics of FMS. **b)** Average *CME-2* vs. FMS.

Figure (7.2a) compares the predictions of the average *CDA-2* and *CME-2* on the algorithmic transition. As already indicated by the results in Figure (7.1), the average *CDA-2* is much closer to the results of the FMS algorithm. However, this does not mean that it provides an accurate description of the problem for two reasons: i) the dynamics of the average *CDA-2* are quantitatively very far from the algorithm, and ii) everything fails for low values of  $\eta$ .

Below  $\eta = 0.4$ , the average *CDA-2* and FMS enter the region  $\alpha_d < \alpha < \alpha_s$ , with  $\alpha_d = 3.86$ . In this region, temporal and spatial correlations are expected to play a fundamental role in slowing down the dynamics. As discussed in Chapter 6, mean-field techniques such as CDA and CME fail to predict the frustration of the dynamics, even in their version written for individual instances. Therefore, results for small  $\eta$  will not improve even if a more suitable method for deriving the average equations is developed.

However, this Fig. (7.2a) suggests that *CDA-2* is the most appropriate technique to continue studying algorithm dynamics. Moreover, Fig. (7.2b) provides information that allows evaluating the procedures used to obtain the average equations. As explained in detail in Appendix P, it is possible to further simplify Eq. (7.3). DINA, as introduced in Ref. [21], is obtained by assuming that all configurations satisfying a clause are equally probable.

Fig. (7.2b) presents predictions for the algorithmic transition of FMS in 3-SAT using this approach. Similar to Section 5.4 for the WalkSAT algorithm, the results here are very far from the real behavior of the system. This highlights the importance of considering at least a non-trivial probability distribution  $\hat{P}_c(u, s_0)$  for the number

of clauses that will become unsatisfied when flipping a variable.

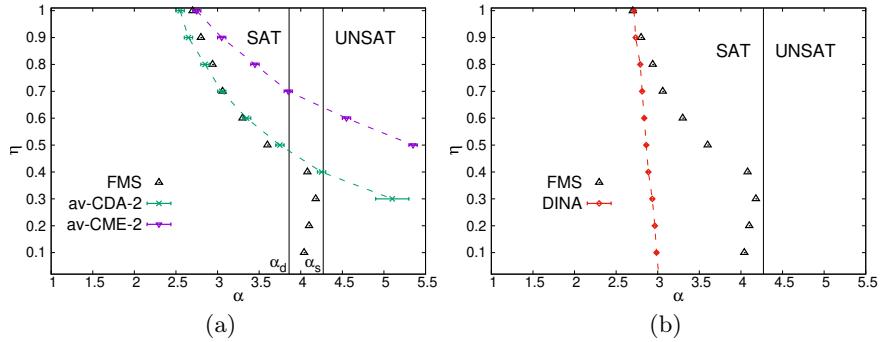


Figure 7.2: Comparison between the phase diagram of the FMS algorithm in 3-SAT and theoretical predictions with average equations. The boundary between SAT and UNSAT zones of the algorithm was obtained by running it on 100 instances of the problem for different values of  $\alpha$  for a time  $t = 10^5 N$ , with  $N = 50000$ . For each  $\eta$ , the value  $\alpha_{\text{FMS}}(\eta)$  was determined where half of the runs solve the problem (triangles). The algorithmic transition corresponding to the average equations was obtained by plotting theoretical curves on a logarithmic scale for different values of  $\alpha$  and  $\eta$ . This determined the region where the curves quickly converge to zero. For more details on the procedure, see Appendix J. a) Results of CDA-2 and CME-2 in their averaged version. b) Results of DINA applied to FMS.

Fig. (7.2) is useful for comparing different average equations with each other, not for evaluating the accuracy of each one separately. In fact, Fig. (7.1) indicates that none of them appropriately describes the dynamics of the FMS algorithm. Not even the averaged version of CDA-2, which shows the best performance in Fig. (7.2), correctly predicts the time scales of the relaxation of FMS to its stationary state.

The results can be improved by working in two independent directions. The first is to add more details about the local structure of the probabilities  $\hat{P}_c$ . For example, it is possible to obtain differential equations for the temporal evolution of  $\hat{P}_c(u, s_0, u_1)$ , where  $u_1$  is the number of clauses that are only unsatisfied by a variable other than the central variable. This considers a non-trivial dependence of probabilities with the new parameter  $u_1$ . It would also be possible to include all  $u_l$ , with  $l = 2, \dots, K - 1$ . Each new inclusion represents a higher computational cost in numerical integration and a lower simplicity of the equations.

The above procedure does not guarantee success. A description that uses only the mean values  $\hat{P}_c$  of the probabilities is perhaps insufficient in problems that exhibit disorder both in connectivities and in couplings  $J_i^a$ . To overcome the limitations of mean-field techniques in K-SAT, it would be necessary to include fluctuations around  $\hat{P}_c$ . This is fundamentally different from adding details to the structure of  $\hat{P}_c$  and requires deriving new types of equations.

## Partial conclusions

In this chapter, average equations representing *CDA-2* and *CME-2* have been derived. The numerical results reinforce the idea of the superiority of the former in describing the dynamics of discrete variables in random graphs, as previously discussed in Chapter 6. When compared with another average technique from the literature, DINA, significant advantages of the equations proposed here are observed in Section 7.1.

However, predictions for the algorithmic transition of FMS fail at low values of the parameter  $\eta$ . This is qualitatively similar to what happens at low temperatures in the p-spin ferromagnetic model (Chapter 6). The presence of long-range correlations is not captured by mean-field techniques such as CDA.

In the future, it is advisable to use CDA written for a given instance of the problem or, alternatively, to develop new average equations that consider more details about the local structure of the interactions.



# Conclusions

This thesis contributes to consolidating the Cavity Master Equation (CME), already present in the literature, as a relevant technique for studying the continuous-time dynamics of discrete variables on random graphs. Additionally, a new technique called Approximate Conditional Dynamics (CDA) is introduced as an alternative that surpasses the CME in various scenarios.

The CME is equivalent to several known results in the literature. For equilibrium, it recovers the well-known message passing algorithm (BP) equations. Moreover, the average equations obtained for the asymmetric ferromagnetic Ising model reproduce the exact solution known in the literature.

On the other hand, the hierarchical system of equations presented in this thesis corrects the CME by adding the effects of spatial correlations. The results improve monotonically as the level of approximation is increased.

The CME was adapted to two algorithms used in the K-Satisfiability (K-SAT) problem: Focused Metropolis Search (FMS) and WalkSAT. In both cases, it predicts the qualitative form of the algorithmic transition across broad regions of the phase diagrams for 3-SAT. The CME fails near the algorithmic transition and in the difficult region  $\alpha_d < \alpha < \alpha_s$ , where the effects of correlations are significant. The results indicate that the mechanisms affecting the efficiency of FMS for  $\alpha_d < \alpha < \alpha_s$  are different from those causing the WalkSAT algorithmic transition.

The CDA was employed to describe the spin glass dynamics of the ferromagnetic p-spin model. The inclusion of local correlation effects between plaquettes is sufficient to reproduce the dynamics at zero temperature ( $T = 0$ ). At  $T > 0$ , introducing a new timescale allows the modified equations to fit Monte Carlo simulations for random graphs very well.

The CME and CDA have been simplified to obtain average equations that have been applied to ferromagnetic models, epidemic spreading, and the description of algorithms used in the K-Satisfiability (K-SAT) problem. The approximations made lose relevant information about the actual behavior of the algorithms in K-SAT and need to be improved.

The theoretical techniques for dynamics explored in this thesis still face significant limitations in high-correlation regimes. Nevertheless, in several cases, they are more accurate than the methods available in the literature. Thus, the CME and CDA become attractive for future studies of out-of-equilibrium systems.



# Recommendations

Future work directly related to this thesis should be divided into three lines: i) apply Approximate Conditional Dynamics (CDA) to individual instances of local search algorithms for K-SAT, ii) improve the approximations made in deriving the average equations, and iii) derive new equations capable of capturing the presence of long-range correlations and the consequent slowdown of dynamics.

The results of the thesis recommend using CDA over the Cavity Master Equation (CME). The comparison with the equations written for individual instances also suggests that where computational cost is not too high, CDA results should be explored at the instance level.

The inclusion of more details about the local structure of probabilities improved the results of the average equations in K-SAT. Further progress should be made in this direction. Beyond this, it is possible that in these models with high disorder in interactions, average magnitudes may be simply insufficient to describe the temporal evolution of the system. It may be useful to include additional parameters that account for fluctuations around those average values.

On the other hand, none of the methods presented here can account for the presence of long-range correlations in the system. This issue could be addressed by writing differential equations for probabilities defined at two or more times:  $P(\sigma(t), \sigma(t'), \dots)$ . This would include the effects of previous times  $t' < t$  on the dynamics at time  $t$ . It is also possible to design a procedure analogous to replica symmetry breaking to apply to message passing dynamic equations. However, at first glance, it seems complex to proceed to obtain a closed system of differential equations.

Finally, this thesis may also pave the way for new applications. There are other models, apart from the ferromagnetic p-spin, that exhibit spin glass dynamics under certain conditions. CDA is presented here as a relevant technique when accurately determining the stationary state of the system at zero temperature. The CME and CDA may also be useful in inference problems, especially if equations at two times can be written.



## Part IV

# Appendices



## Appendix A

# Point-to-set correlation in the $p$ -spin model

This appendix shows how to calculate the point-to-set correlations  $C_{PS}(i, l)$  mentioned in Section 1.1 and how to use them to determine the temperature  $T_d$  of transition to the spin glass. These correlations are defined as:

$$C_{PS}(i, l) = \sum_{\vec{\sigma}_l} P(\vec{\sigma}_l) \left( \sum_{\sigma_i} P(\sigma_i | \vec{\sigma}_l) \right)^2 - \sum_{\sigma_i} \left( P(\sigma_i) \sigma_i \right)^2 \quad (\text{A.1})$$

In the case of the simulations represented by blue points in Fig. (1.2b), the system is kept all the time with magnetization  $m \approx 0$  and  $P(\sigma_i) = 1/2$  for all  $i$ . If no additional information is provided, the spin  $\sigma_i$  can point up or down with equal probability. Calculating  $C_{PS}(i, l)$  is equivalent to determining whether it is likely to reconstruct the value of  $\sigma_i$  given the configuration  $\vec{\sigma}_l$ .

To calculate  $P(\sigma_i | \vec{\sigma}_l)$ , it must be noted that, once  $\vec{\sigma}_l$  is fixed, the subgraph of nodes within distance no greater than  $l$  from  $\sigma_i$  is a boundary tree  $\vec{\sigma}_l$  with center at  $\sigma_i$ . The equilibrium cavity method is valid and can be applied if the values of messages on the boundary are appropriately defined. Here, we will use the belief propagation (BP) equations presented in Section 3.1. Then, the average over all possible boundary conditions and all tree-like graphs connecting  $\sigma_i$  with spins  $\vec{\sigma}_l$  must be taken.

Consider a tree-like graph where the distance between the root node  $i$  and the leaves is exactly  $l$ . For all  $\sigma_k \in \vec{\sigma}_l$ , it will hold that  $\nu_{k \rightarrow a}(s_k) = \delta_{s_k, \sigma_k}$ . It is convenient to parameterize the messages in equations (3.1) and (3.2) as:

$$\nu_{i \rightarrow a}(\sigma_i) = \frac{1 + \sigma_i \mu_{i \rightarrow a}}{2} \quad \eta_{a \rightarrow i}(\sigma_i) = \frac{1 + \sigma_i \hat{\mu}_{a \rightarrow i}}{2} \quad (\text{A.2})$$

If different realizations of the interaction hypergraph are averaged, the indices lose their meaning. We can simply write  $\mu_{i \rightarrow a} \equiv \mu$ ,  $\hat{\mu}_{a \rightarrow i} \equiv \hat{\mu}$ , and  $\sigma_i \equiv \sigma$ . Applying the BP equations to different boundary configurations compatible with the central value  $\sigma$  yields the probability distributions  $Q_\sigma^l(\mu)$  and  $\hat{Q}_\sigma^l(\hat{\mu})$ . These satisfy the following iterative equations:

$$Q_\sigma^l(\mu) = \int \left[ \prod_{i=1}^{c-1} d\hat{\mu}_i \hat{Q}_\sigma^l(\hat{\mu}_i) \right] \delta(\mathcal{F}_1[\mu, \{\hat{\mu}_i\}_{i=1}^{c-1}]) \quad (\text{A.3})$$

$$\hat{Q}_\sigma^l(\hat{\mu}) = \sum_{\sigma_1, \dots, \sigma_{p-1}} P(\sigma_1, \dots, \sigma_{p-1} | \sigma) \int \left[ \prod_{j=1}^{p-1} d\mu_j Q_{\sigma_j}^{l-1}(\mu_j) \right] \delta(\mathcal{F}_2[\hat{\mu}, \{\mu_j\}_{j=1}^{p-1}]) \quad (\text{A.4})$$

where  $\mathcal{F}_1$  and  $\mathcal{F}_2$  are the same equations (3.1) and (3.2) written in terms of  $\mu$  and  $\hat{\mu}$ :

$$\mathcal{F}_1[\mu, \{\hat{\mu}_i\}_{i=1}^{c-1}] = \frac{1+\mu}{2} - \frac{\prod_{i=1}^{c-1} (1+\hat{\mu}_i)}{\sum_\sigma \left[ \prod_{i=1}^{c-1} (1+\sigma \hat{\mu}_i) \right]} \quad (\text{A.5})$$

$$\mathcal{F}_2[\hat{\mu}, \{\mu_j\}_{j=1}^{p-1}] = \hat{\mu} - \tanh(\beta J) \prod_{j=1}^{p-1} \mu_j \quad (\text{A.6})$$

The conditional probability  $P(\sigma_1, \dots, \sigma_{p-1} | \sigma)$  is taken according to the Boltzmann distribution:

$$P(\sigma_1, \dots, \sigma_{p-1} | \sigma) = \frac{\exp\left(-\beta J \sigma \prod_{k=1}^{p-1} \sigma_k\right)}{\sum_{\sigma_1, \dots, \sigma_{p-1}} \exp\left(-\beta J \sigma \prod_{k=1}^{p-1} \sigma_k\right)} \quad (\text{A.7})$$

Since the Hamiltonian of the  $p$ -spin model is symmetric under simultaneous spin inversions ( $\sigma \rightarrow -\sigma$ ), it follows that  $Q_+^l(\mu) = Q_-^l(-\mu)$  (and a similar expression for  $\hat{Q}$ ). The equations can be simplified:

$$Q^l(\mu) = \int \left[ \prod_{i=1}^{c-1} d\hat{\mu}_i \hat{Q}^l(\hat{\mu}_i) \right] \delta(\mathcal{F}_1[\mu, \{\hat{\mu}_i\}_{i=1}^{c-1}]) \quad (\text{A.8})$$

$$\hat{Q}_\sigma^l(\hat{\mu}) = \sum_{\sigma_1, \dots, \sigma_{p-1}} \pi(\sigma_1, \dots, \sigma_{p-1}) \int \left[ \prod_{j=1}^{p-1} d\mu_j Q^{l-1}(\sigma_j \mu_j) \right] \delta(\mathcal{F}_2[\hat{\mu}, \{\mu_j\}_{j=1}^{p-1}]) \quad (\text{A.9})$$

where  $\pi(\sigma_1, \dots, \sigma_{p-1}) = P(\sigma_1, \dots, \sigma_{p-1} | +)$ .

The equations can then be solved with a population dynamics algorithm [10, 48] to calculate  $C_{PS}(l) = \int d\mu Q(\mu) \mu$ .

To determine the temperature at which  $C_{PS}(l) \not\rightarrow 0$  as  $l \rightarrow \infty$ , we define the correlation length:

$$l^*(\epsilon) = \min\{l : C_{PS}(l) < \epsilon\} \quad (\text{A.10})$$

The point-to-set correlation length  $l^*(\epsilon)$  diverges at the temperature  $T_d$  of the dynamic transition to spin glass [48–50]. Fig.(A.1a) shows that near  $T_d$ , the value of

$l^*$  increases, and Fig.(A.1b) shows a nonlinear fit of these values according to the law  $l^* = A/(T - T_d)^{1/2}$ . The result is  $T_d \approx 0.753$ .

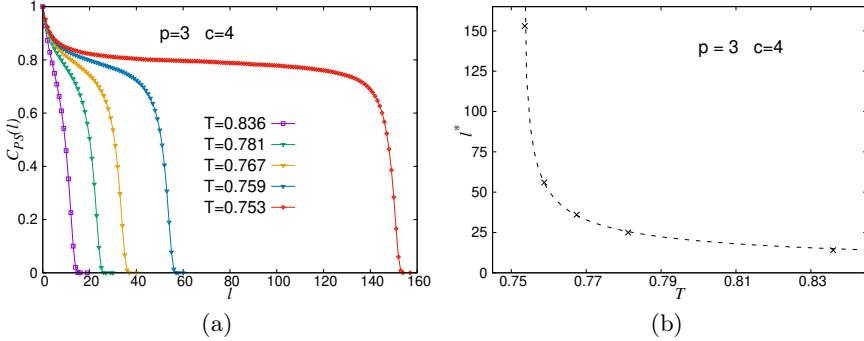


Figure A.1: Point-to-set correlation for the 3-spin ferromagnet defined on random regular hypergraphs with  $c = 4$ . **a)** Dependence of the point-to-set correlation with distance for various temperatures. **b)** Point-to-set correlation length calculated with the small parameter  $\epsilon = 0.05$ . A fit with the law  $l^* = A/(T - T_d)^{1/2}$  (dashed lines) gives the value  $T_d \approx 0.753$  for the temperature of the dynamic transition to spin glass.



## Appendix B

# Derivation of the master equation for $N$ spins

Now we will consider the dynamics of a system of  $N$  interacting spins:  $\vec{\sigma} = \{\sigma_1, \dots, \sigma_N\}$ . The Master Equation is given by Eq. (2.1). Analogous to Eq. (2.5), the probability  $P^t(\vec{\sigma})$  can be written as:

$$P^t(\vec{\sigma}) = \sum_{\vec{X}|\vec{\sigma}}^t Q^t(\vec{X}) \quad (\text{B.1})$$

where  $\vec{X} = \{X_1, \dots, X_N\}$  is the vector of the histories of all the spins.

Now we must differentiate (B.1) analogously to (2.5):

$$\frac{dP^t(\vec{\sigma})}{dt} = \lim_{\Delta t \rightarrow 0} \frac{P^{t+\Delta t}(\vec{\sigma}) - P^t(\vec{\sigma})}{\Delta t} \quad (\text{B.2})$$

The probability density of the set of individual histories of the  $N$  spins,  $Q^t(\vec{X})$ , can be expressed as a product of the probability densities  $\Phi_i^t(X_i | \vec{X}_{\setminus i})$  of the occurrence of the history  $X_i$  when the trajectories of the rest of the variables  $\vec{X}_{\setminus i}$  are fixed [87]:

$$Q^t(X_1, X_2, \dots, X_N) = \prod_{i=1}^N \Phi_i^t(X_i | \vec{X}_{\setminus i}) \quad (\text{B.3})$$

Each  $\Phi_i^t(X_i | \vec{X}_{\setminus i})$  can be parameterized similarly to how it was done in Eq. (2.10) [54, 86]:

$$\Phi_i^t(X_i | \vec{X}_{\setminus i}) = \left[ \prod_{l_i=1}^{s_i} r_i(\vec{\sigma}(t_{l_i})) \right] \exp \left\{ - \int_{t_0}^t r_i(\vec{\sigma}(\tau)) d\tau \right\} \quad (\text{B.4})$$

As in the previous subsection,  $P^{t+\Delta t}(\vec{\sigma})$  will be written as a marginal of  $Q^{t+\Delta t}(\vec{X})$  and then expanded to the first order in  $\Delta t$ . Explicitly writing this operation as in (2.9), we get:

$$P^{t+\Delta t}(\vec{\sigma}) = \sum_{s_1=0}^{\infty} \sum_{s_2=0}^{\infty} \dots \sum_{s_N=0}^{\infty} \left[ \prod_{i=1}^N \int_{t_0}^{t+\Delta t} dt_1^i \int_{t_1^i}^{t+\Delta t} dt_2^i \dots \int_{t_{s_i-1}^i}^{t+\Delta t} dt_{s_i}^i \right] Q^{t+\Delta t}(\vec{X}) \quad (\text{B.5})$$

To expand  $Q^{t+\Delta t}(\vec{X})$ , we first need to work with  $\Phi_i^{t+\Delta t}(X_i | \vec{X}_{\setminus i})$ , similarly to how it was done with (2.11):

$$\Phi_i^{t+\Delta t}(X_i | \vec{X}_{\setminus i}) = \Phi_i^t(X_i | \vec{X}_{\setminus i}) [1 - r_i(\vec{\sigma}(t)) \Delta t] + o(\Delta t) \quad (\text{B.6})$$

Thus:

$$\begin{aligned} Q^{t+\Delta t}(\vec{X}) &= \prod_{i=1}^N \Phi_i^{t+\Delta t}(X_i | \vec{X}_{\setminus i}) \\ Q^{t+\Delta t}(\vec{X}) &= \prod_{i=1}^N \Phi_i^t(X_i | \vec{X}_{\setminus i}) \left[ 1 - \Delta t \sum_{k=1}^N r_k(\vec{\sigma}(t)) \right] + o(\Delta t) \\ Q^{t+\Delta t}(\vec{X}) &= Q^t(\vec{X}) \left[ 1 - \Delta t \sum_{k=1}^N r_k(\vec{\sigma}(t)) \right] + o(\Delta t) \end{aligned} \quad (\text{B.7})$$

To expand  $Q^{t+\Delta t}(\vec{X})$ , we first need to work with  $\Phi_i^{t+\Delta t}(X_i | \vec{X}_{\setminus i})$ , similarly to how it was done with (2.11).

We can now use (B.7) to obtain two contributions of order  $O(1)$  and  $O(\Delta t)$ , respectively:

$$I_0 = \sum_{s_1=0}^{\infty} \dots \sum_{s_N=0}^{\infty} \left[ \prod_{i=1}^N \int_{t_0}^t dt_1^i \dots \int_{t_{s_i-1}^i}^t dt_{s_i}^i \right] Q^t(\vec{X}) = P^t(\vec{\sigma}) \quad (\text{B.8})$$

$$I_1 = \left( \sum_{s_1=0}^{\infty} \dots \sum_{s_N=0}^{\infty} \left[ \prod_{i=1}^N \int_{t_0}^t dt_1^i \dots \int_{t_{s_i-1}^i}^t dt_{s_i}^i \right] Q^t(\vec{X}) \right) \left( -\Delta t \sum_{k=1}^N r_k(\vec{\sigma}(t)) \right)$$

$$I_1 = -P^t(\vec{\sigma}) \Delta t \sum_{k=1}^N r_k(\vec{\sigma}(t)) \quad (\text{B.9})$$

As in Eq. (2.12), we can split the sum of iterated integrals on the right-hand side of Eq. (B.5) to explicitly write the  $\Delta t$  order contribution. Again, we obtain an operator that acts in the space of trajectories where a single spin has a single jump in  $[t, t + \Delta t]$ . The rest is essentially the same as what was done for (2.15):

$$\begin{aligned} I_2 &= \sum_{k=1}^N \sum_{s_1=0}^{\infty} \dots \sum_{s_N=0}^{\infty} \left[ \prod_{i \neq k} \int_{t_0}^t dt_1^i \dots \int_{t_{s_i-1}^i}^t dt_{s_i}^i \right] \int_{t_0}^t dt_1^k \dots \int_{t_{s_k-2}}^t dt_{s_k-1}^k \left[ \prod_{i \neq k} \Phi_i^t(X_i | \vec{X}_{\setminus i}) \right] \times \\ &\quad \times \left[ \prod_{l_k=1}^{s_k-1} r_k(\vec{\sigma}(t_{l_k})) \right] e^{-\int_{t_0}^t r_k(\vec{\sigma}(\tau)) d\tau} r_k(F_k[\vec{\sigma}(t)]) \Delta t + o(\Delta t) \\ I_2 &= \sum_{k=1}^N P^t(F_k[\vec{\sigma}]) r_k(F_k[\vec{\sigma}(t)]) \Delta t \end{aligned} \quad (\text{B.10})$$

Finally, the expansion of  $P^{t+\Delta t}(\vec{\sigma})$  is:

$$\begin{aligned} P^{t+\Delta t}(\vec{\sigma}) &= I_0 + I_1 + I_2 + o(\Delta t) \\ P^{t+\Delta t}(\vec{\sigma}) &= P^t(\vec{\sigma}) - \Delta t \sum_{k=1}^N P^t(\vec{\sigma}) r_k(\vec{\sigma}) + \Delta t \sum_{k=1}^N P^t(F_k[\vec{\sigma}]) r_k(F_k[\vec{\sigma}]) + o(\Delta t) \end{aligned} \quad (\text{B.11})$$

and using Eq. (B.2) we obtain the usual Master Equation for a set of  $N$  interacting spins (2.1):

$$\frac{dP^t(\vec{\sigma})}{dt} = - \sum_{i=1}^N [r_i(\vec{\sigma}) P^t(\vec{\sigma}) - r_i(F_i[\vec{\sigma}]) P^t(F_i[\vec{\sigma}])] \quad (\text{B.12})$$

Unfortunately, the high dimensionality of the probabilities  $P^t(\vec{\sigma})$  makes solving Eq. (B.12) generally difficult, even numerically.



## Appendix C

# The stationary CME

In this appendix, we show how to write fixed-point equations for the stationary solution of the CME in its original version (see Eq. (2.40)).

The stationary condition for the CME,  $\frac{dp(\sigma_i \parallel \sigma_j)}{dt} = 0$ , generates a set of coupled equations for the cavity conditional probabilities that is very similar in structure to Eq. (3.6). In the stationary state, the probabilities of leaving and entering any state are balanced:

$$\sum_{\sigma_{\partial i \setminus j}} r_i^+ p_i(\sigma_i \parallel \sigma_j) \prod_{k \in \partial i \setminus j} p(\sigma_k \parallel \sigma_i) = \sum_{\sigma_{\partial i \setminus j}} r_i^- p(-\sigma_i \parallel \sigma_j) \prod_{k \in \partial i \setminus j} p(\sigma_k \parallel -\sigma_i) \quad (\text{C.1})$$

where  $r_i^+ \equiv r_i(\sigma_i, \sigma_{\partial i})$  and  $r_i^- \equiv r_i(-\sigma_i, \sigma_{\partial i})$ .

Using  $p(-\sigma_i \parallel \sigma_j) = 1 - p(\sigma_i \parallel \sigma_j)$ , we obtain:

$$p(\sigma_i \parallel \sigma_j) = \frac{\sum_{\sigma_{\partial i \setminus j}} r_i^- \prod_{k \in \partial i \setminus j} p(\sigma_k \parallel -\sigma_i)}{\sum_{\sigma_i} \sum_{\sigma_{\partial i \setminus j}} r_i^+ \prod_{k \in \partial i \setminus j} p(\sigma_k \parallel \sigma_i)} \quad (\text{C.2})$$

The denominator of Eq. (C.2) is the same for  $p(\sigma_i \parallel \sigma_j)$  and for  $p(-\sigma_i \parallel \sigma_j)$ . It can be identified as a normalization factor similar to those in BP equations. This allows us to write:

$$p(\sigma_i \parallel \sigma_j) \propto \sum_{\sigma_{\partial i \setminus j}} r_i^- \prod_{k \in \partial i \setminus j} p(\sigma_k \parallel -\sigma_i) \quad (\text{C.3})$$

To solve the system of equations (C.3), the form of the transition probabilities per unit time  $r_i$  must be specified. In the literature, two main types of dynamics that respect detailed balance are used (see Eq. (3.5)): the Metropolis dynamics [46] and Glauber dynamics [70]. Here, Glauber transition probabilities will be used explicitly:

$$r_i(\sigma_i, \sigma_{\partial i}) = \frac{1}{2} \left[ 1 - \sigma_i \tanh \left( \beta \sum_{k \in \partial i} J_{ki} \sigma_k \right) \right] = \frac{\exp[-\beta \sigma_i \sum_{k \in \partial i} J_{ki} \sigma_k]}{2 \cosh[\beta \sum_{k \in \partial i} J_{ki} \sigma_k]} \quad (\text{C.4})$$

Then:

$$p(\sigma_i \parallel \sigma_j) \propto \sum_{\sigma_{\partial i \setminus j}} \frac{\exp[-\beta \sigma_i \sum_{k \in \partial i} J_{ki} \sigma_k]}{\cosh[\beta \sum_{k \in \partial i} J_{ki} \sigma_k]} \prod_{k \in \partial i \setminus j} p(\sigma_k \parallel -\sigma_i) \quad (\text{C.5})$$

In fact, the BP equation (3.6) and Equation (C.3) have a similar form. Applying the stationary condition to Equation (2.42) also yields an equation for  $P(\sigma_i)$ :

$$P(\sigma_i) \propto \sum_{\sigma_{\partial i}} \frac{\exp[-\beta \sigma_i \sum_{k \in \partial i} J_{ki} \sigma_k]}{\cosh[\beta \sum_{k \in \partial i} J_{ki} \sigma_k]} \prod_{k \in \partial i} p(\sigma_k \parallel -\sigma_i) \quad (\text{C.6})$$

It is possible to design an algorithm analogous to BP that finds the fixed point of Equations (C.3) and (C.6). With the resulting probability distributions, observables such as local magnetization can be calculated. This algorithm will be referred to as CME-BP from now on.

Figure (C.1) compares the results of CME-BP with the long-time limit obtained from the numerical integration of the original CME (Equations (2.40) and (2.42)). This is done for the ferromagnetic Ising model ( $J_{ij} = J$ ) defined on a single Erdős-Rényi (ER) graph [11] with average connectivity  $\kappa = 3$ . The local error is defined as:

$$\gamma(t) = \sqrt{\frac{1}{N} \sum_i (m_i^{CME}(t) - m_i^{CME-BP})^2} \quad (\text{C.7})$$

where  $m_i^{CME}(t)$  and  $m_i^{CME-BP}$  are the local magnetizations obtained from the original CME and from CME-BP, respectively.

It is important to note that the error  $\gamma(t)$  depends on time only because the original CME yields a time-dependent magnetization:  $m_i^{CME}(t)$ . It can be observed that  $\gamma(t) \rightarrow 0$  as  $t \rightarrow \infty$  for all temperatures, which means that the solution of Equations (C.3) indeed corresponds to the stationary state of (2.40).

Of course, an algorithm that finds the fixed-point solutions of (C.3) needs some convergence criterion. Similarly to how it is done for the usual BP algorithm, the parameter must be defined as:

$$\epsilon = \max_{\{i,j,\sigma_i,\sigma_j\}} \left[ \frac{\Delta p(\sigma_i \parallel \sigma_j)}{p(\sigma_i \parallel \sigma_j)} \right] \quad (\text{C.8})$$

where  $\Delta p(\sigma_i \parallel \sigma_j)$  is the change in  $p(\sigma_i \parallel \sigma_j)$  in one iteration of the algorithm. CME-BP stops when  $\epsilon < \epsilon_s$ , where  $\epsilon_s$  is a small positive real number.

With these tools, it is sufficient to compare CME-BP with the original BP algorithm. The inset plot in Fig. (C.2a) computes the local error between the magnetizations corresponding to the fixed points of BP and CME-BP. The new error is:

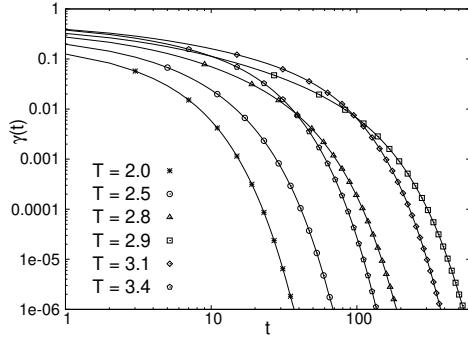


Figure C.1: Comparison between CME and CME-BP on a single Erdős-Rényi graph with average connectivity  $\kappa = 3$  and size  $N = 1000$ . The figure shows the time-dependent local error defined in (C.7) for various temperatures. All CME integrations start with a fully magnetized system in contact with a thermal bath at fixed temperature  $T$ . All CME-BP runs started with  $p(\sigma_i = 1 \parallel \sigma_j) = 1$  for all pairs  $(ij)$ . The algorithm's convergence parameter (see Eq. (C.8)) was set to  $\epsilon_s = 10^{-11}$ .

$$\mu(\infty) = \sqrt{\frac{1}{N} \sum_i (m_i^{CME-BP} - m_i^{BP})^2} \quad (C.9)$$

The main panel of Fig. (C.2a) shows the average of this error calculated over several ER graphs with the same average connectivity. This error is small for all temperatures and has a maximum near the critical temperature  $T_c^{ER}(\kappa = 3) \approx 2.89$ . This marks the order-disorder transition of the mentioned ferromagnetic Ising model on ER graphs. Below  $T_c$ , the stationary state is ferromagnetic ( $m \neq 0$ ) and above it is paramagnetic ( $m \approx 0$ ).

Fig. (C.2b) shows that, even under the influence of criticality, the error  $\mu(\infty)$  goes to zero as the quality of convergence of CME-BP and BP improves ( $\epsilon_s \rightarrow 0$ ).

In summary, Fig. (C.2) shows a good match between the magnetizations predicted by BP and the stationary solution of the CME. This corroborates the analytical results of Section 3.2. The CME is equivalent to the equilibrium cavity method (i.e., BP) when the transition probabilities satisfy detailed balance.

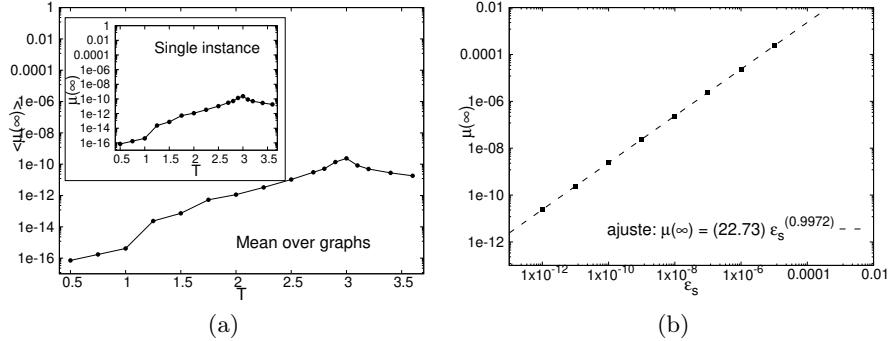


Figure C.2: Comparison between CME-BP and BP on Erdős-Renyi graphs with average connectivity  $\kappa = 3$  and size  $N = 1000$ . All CME-BP runs started with  $p(\sigma_i = 1 \mid \sigma_j) = 1$  for all pairs  $(ij)$ . Similarly, all BP runs initially had  $m_{i \rightarrow j}(\sigma_i = 1) = 1$ . **a)** Local error between the magnetizations corresponding to the fixed points of the algorithms. The inset plot shows the error (C.9) for various temperatures and a single Erdős-Rényi graph. The main panel shows the average of the local error (C.9) calculated over several Erdős-Rényi graphs with the same average connectivity. The number of graphs ranged from  $s = 20$  to  $s = 100$ . The convergence parameter of both algorithms (see Eq. (C.8)) was  $\epsilon_s = 10^{-11}$ . **b)** Dependence of the local error on the convergence parameter  $\epsilon_s$  near the critical temperature  $T_c^{FER}(\kappa = 3) \approx 2.89$ . Each point was obtained using the same parameter  $\epsilon_s$  for both algorithms in runs at the same temperature  $T = 3.0$ . The data were fitted to the curve  $\mu(\infty) = a \epsilon_s^b$ , with  $a = 22.73 \pm 0.02$  and  $b = 0.9972 \pm 0.0001$ .

## Appendix D

### Computation of the sums for the average case

Here we will derive the expression (3.24) that appears in subsection 3.3.1. In Eq. (3.22), we have a sum of the form:

$$S(\hat{m}, \beta J, q, \sigma) = \sum_{k=0}^{\infty} \frac{q^k}{k!} \sum_{n=0}^k \binom{k}{n} \left(\frac{1+\hat{m}}{2}\right)^n \left(\frac{1-\hat{m}}{2}\right)^{k-n} \times \tanh(\beta J(2n-k+\sigma)) \quad (\text{D.1})$$

First, we need to interchange the sums and rewrite everything as:

$$\begin{aligned} S(\hat{m}, \beta J, q, \sigma) &= \sum_{n=0}^{\infty} \sum_{k=n}^{\infty} \frac{q^k}{k!} \binom{k}{n} \left(\frac{1+\hat{m}}{2}\right)^n \left(\frac{1-\hat{m}}{2}\right)^{k-n} \tanh(\beta J(2n-k+\sigma)) \\ S(\hat{m}, \beta J, q, \sigma) &= \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \frac{q^{n+l}}{n! l!} \left(\frac{1+\hat{m}}{2}\right)^n \left(\frac{1-\hat{m}}{2}\right)^l \tanh(\beta J(n-l+\sigma)) \end{aligned} \quad (\text{D.2})$$

Between the first and second lines of (D.2), we applied the change of variables  $l = k - n$ . With another change ( $n' = n - l$ ), we obtain:

$$\begin{aligned} S(\hat{m}, \beta J, q, \sigma) &= \sum_{n'=-\infty}^{\infty} \tanh(\beta J(n'+\sigma)) \sum_{l=0}^{\infty} \frac{q^{2l+n'}}{(n'+l)! l!} \left(\frac{1+\hat{m}}{2}\right)^{n'+l} \left(\frac{1-\hat{m}}{2}\right)^l \\ S(\hat{m}, \beta J, q, \sigma) &= \sum_{n'=-\infty}^{\infty} \tanh(\beta J(n'+\sigma)) \left(\frac{1+\hat{m}}{1-\hat{m}}\right)^{n'/2} \times \\ &\quad \times \sum_{l=0}^{\infty} \frac{(q/2)^{2l+n'}}{(n'+l)! l!} \left(\sqrt{1-\hat{m}^2}\right)^{2l+n'} \end{aligned} \quad (\text{D.3})$$

The expression (3.24) is the result of simply recognizing in (D.3) the modified Bessel function of the first kind (Infeld)  $I_{\nu}(x)$ :

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(k + \nu + 1)} \left(\frac{x}{2}\right)^{2k+\nu} \quad (D.4)$$

From which, finally

$$S(\hat{m}, \beta J, q, \sigma) = \sum_{n'=-\infty}^{\infty} \tanh(\beta J(n' + \sigma)) \left(\frac{1 + \hat{m}}{1 - \hat{m}}\right)^{n'/2} I_{n'}(q \sqrt{1 - \hat{m}^2}) \quad (D.5)$$

## Appendix E

# Populations dynamics for the asymmetric ferromagnet

To complement the deduction of subsection 3.3.2, the approximations made to obtain the mean equations for the symmetric ferromagnetic model on random graphs will be numerically evaluated. To begin, it can be demonstrated that writing the CME in terms of probabilities that depend only on connectivity leads to numerically correct results. For this, the following population dynamics algorithm is defined to obtain the cavity conditional probabilities:

```

1: Choose two positive integers  $M$ ,  $n$ , and a real number  $t_{\max}$ 
2: Initialize the vector  $\{p_\gamma(\sigma' \parallel \sigma)\}$ , with  $\gamma = 1, 2, \dots, M$ 
3: while  $t < t_{\max}$  do
4:   for  $\gamma = (1, \dots, M)$  do
5:     for  $i = (1, \dots, n)$  do
6:       Choose a set  $\{\gamma_k^i\}_{k \in \{1, \dots, \gamma\}}$  using  $q(\{\gamma_k^i\})$ 
7:       Construct the corresponding vector  $\{p_{\gamma_k}^i(\sigma' \parallel \sigma)\}$ 
8:       Calculate  $\dot{p}_\gamma^i(\sigma' \parallel \sigma)$  using Eq. (2.40) and  $\{p_{\gamma_k}^i(\sigma' \parallel \sigma)\}$ 
9:     end for
10:    Define  $\dot{p}_\gamma(\sigma' \parallel \sigma)$  as the average of all  $\dot{p}_\gamma^i(\sigma' \parallel \sigma)$ 
11:  end for
12:  Update  $\{p_\gamma(\sigma' \parallel \sigma)\}$ 
13:  Calculate the average  $p_\lambda(\sigma' \parallel \sigma)$  of all  $\{p_\gamma(\sigma' \parallel \sigma)\}$  using  $q(\gamma)$ 
14: end while
```

An analogous procedure can be carried out to calculate the local probabilities  $P_c$ . This scheme will henceforth be called CME-sampled. Figure (E.1) compares its results with those of the original CME. It can be observed that the CME-sampled appropriately describes the original dynamics. It is important to note that in the previous algorithm nothing prevents a very large value of  $\gamma_k^i$  from being selected. For numerical implementation, it has been assumed that  $p_\gamma(\sigma' \parallel \sigma) = 1$  for all  $\gamma > \gamma_{\max}$ ,

where  $\gamma_{\max}$  is a sufficiently large natural number that must be chosen at the beginning.

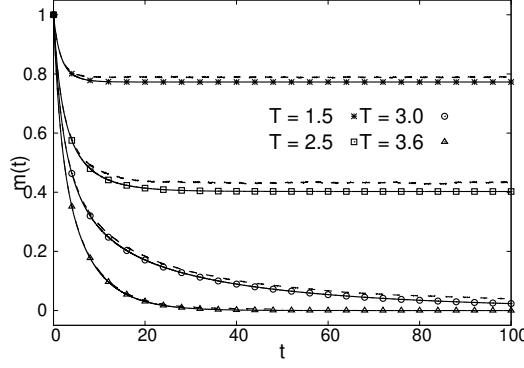


Figure E.1: Comparison between the magnetization predicted by the CME-sampled (dashed lines) and the average of the CME (solid lines) and Monte Carlo simulations (points) on various Erdős-Rényi graphs with average connectivity  $\kappa = 3$ .  $s = 350$  different graphs of size  $N = 4000$  were taken. In all cases, a completely magnetized system is placed in contact with a thermal bath at a fixed temperature  $T$ . It was assumed that  $p_{\gamma>50}(\sigma' \parallel \sigma) = 1$ . A total of 10000 Monte Carlo histories were run for each graph.

At another point in the deduction, equations for the average probabilities  $\hat{p}(\sigma_i \parallel \sigma_j)$  and  $\hat{P}(\sigma_i)$  were proposed. If the same equations are written omitting the sums over the connectivities, one obtains:

$$\begin{aligned} \dot{p}_{\gamma_i}(\sigma_i \parallel \sigma_j) &= -\frac{1}{2} \{ p_{\gamma_i}(\sigma_i \parallel \sigma_j) - p_{\gamma_i}(-\sigma_i \parallel \sigma_j) \} + \times \\ &+ \frac{\sigma_i}{2} \sum_{\{\sigma_k\}} \tanh(\beta J \sum_{k=1}^{\gamma_i} \sigma_k + \beta J \sigma_j) \sum_{\sigma'_i} \left[ \prod_{k=1}^{\gamma_i} \hat{p}(\sigma_k \parallel \sigma'_i) \right] p_{\gamma_i}(\sigma'_i \parallel \sigma_j) \end{aligned} \quad (\text{E.1})$$

$$\begin{aligned} \dot{P}_{c_i}(\sigma_i) &= -\frac{1}{2} \{ P_{c_i}(\sigma_i) - P_{c_i}(-\sigma_i) \} + \frac{\sigma_i}{2} \sigma_i \sum_{\{\sigma_k\}} \tanh(\beta J \sum_{k=1}^{c_i} \sigma_k) \\ &\times \sum_{\sigma'_i} \left[ \prod_{k=1}^{c_i} \hat{p}(\sigma_k \parallel \sigma'_i) P_{c_i}(\sigma'_i) \right] \end{aligned} \quad (\text{E.2})$$

The Eqs. (E.1) and (E.2) can be integrated simultaneously to obtain numerical results. It is not necessary to use the Eqs. (3.28), (3.29) to calculate the probabilities  $\hat{P}$  and  $\hat{p}$  since they are simply the averages of  $P_{c_i}$  and  $p_{\gamma_i}$  over all connectivity values. The following algorithm, which is also a population dynamics, allows calculating the values of  $p_{\gamma_i}$  for different times:

- 1: Choose a positive integer  $M$  and a real number  $t_{\max}$
- 2: Initialize the vector  $\{p_{\gamma}(\sigma' \parallel \sigma)\}$ , with  $\gamma = 1, 2, \dots, M$

---

```

3: while  $t < t_{max}$  do
4:   for  $\gamma = (1, \dots, M)$  do
5:     Calculate  $\dot{p}_\gamma(\sigma'|\sigma)$  with Eq. (E.1)
6:   end for
7:   Update  $\{p_\gamma(\sigma'|\sigma)\}$ 
8:   Calculate the average  $p_\lambda(\sigma'|\sigma)$  of all  $\{p_\gamma(\sigma'|\sigma)\}$  using  $q(\gamma)$ 
9: end while

```

Then the same idea should be followed to calculate the local probabilities  $P_c$ . Figure (E.2) compares the results with the average of the CME for several ER graphs. It can be seen that this second algorithm still accurately describes the numerical average of the original CME.

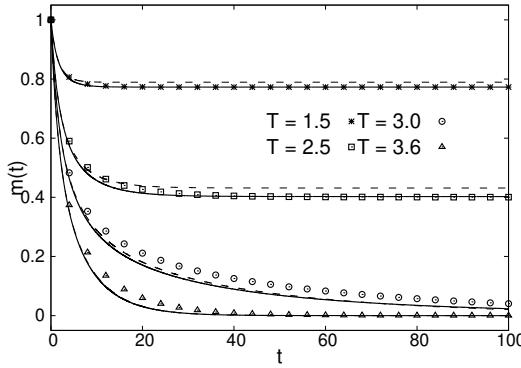


Figure E.2: Comparison between the results of Eqs. (E.1) and (E.2), represented by dashed lines, and the average of the CME (solid lines) and Monte Carlo simulations (points) on various Erdős-Rényi graphs with average connectivity  $\kappa = 3$  ( $T_c \approx 2.89$ ) and size  $N = 4000$ . In all cases, a completely magnetized system is placed in contact with a thermal bath at a fixed temperature  $T$ . Values of  $\gamma = 1, 2, \dots, 50$  were taken. Averages were made over  $s = 350$  different graphs and in each one  $n = 10^4$  Monte Carlo histories were run.

The results shown in this appendix provide a better understanding of the deduction presented in subsection 3.3.2. Figure (E.2) deserves special attention. Comparing it with Figure (3.3), the effects of the approximations made from Eq. (3.31) to Eqs. (3.33) and (3.34) can be directly appreciated. Apparently, it was precisely these last considerations that most affected the accuracy of the obtained average equations.



## Appendix F

# Other average equations

In this appendix, average equations for the CME in the symmetric ferromagnetic model defined on random graphs are derived. The transition probabilities per unit time will again be the Glauber rules (see Eq. (3.15)). Although the new formulation is more complex than the one presented in subsection 3.3.2, the numerical results are much more accurate.

It is convenient to start with the average of the CME (see Eq. (2.40)). Similarly to how it was done in subsection 3.3.2, equations for cavity magnetizations can be obtained:

$$\begin{aligned} \frac{d\tilde{m}(\sigma)}{dt} = & -\tilde{m}(\sigma) + \sum_{\sigma'} \sum_{\gamma=0}^{\infty} q(\gamma) p_{\gamma}(\sigma' \parallel \sigma) \times \\ & \times \sum_{n=0}^{\gamma} \binom{\gamma}{n} \tanh[\beta J(2n - \gamma + \sigma)] [\hat{p}(1 \parallel \sigma')]^n [\hat{p}(-1 \parallel \sigma')]^{\gamma-n} \end{aligned} \quad (\text{F.1})$$

where  $n$  is the number of neighbors of the variable  $\sigma$  pointing in the direction  $\sigma_k = 1$  and  $\binom{\gamma}{n}$  is the Newton binomial coefficient.

To go beyond Eq. (F.1), an *ansatz* for the dependence of  $p_{\gamma}(\sigma' \parallel \sigma)$  on  $\gamma$  will be made. It is reasonable to assume that  $p_{\gamma}(\sigma' \parallel \sigma)$  is a monotonic function of  $\gamma$  for each pair  $(\sigma', \sigma)$ . Given the form of  $q(\gamma)$  in Erdős-Rényi (ER) graphs, it is convenient to take  $p_{\gamma}(\sigma' \parallel \sigma)$  as a power law:

$$p_{\gamma}(\sigma' \parallel \sigma) = p_{\infty}(\sigma' \parallel \sigma) - [p_{\infty}(\sigma' \parallel \sigma) - p_0(\sigma' \parallel \sigma)] b^{\gamma} \quad (\text{F.2})$$

Here,  $p_0(\sigma' \parallel \sigma)$  is the value of  $p_{\gamma}(\sigma' \parallel \sigma)$  when  $\gamma = 0$ . This probability follows the differential equation:

$$\frac{dp_0(\sigma' \parallel \sigma)}{dt} = -\frac{1}{2} \sum_{\sigma''} \sigma' \sigma'' p_0(\sigma'' \parallel \sigma) [1 - \sigma'' \tanh(\beta J \sigma)] \quad (\text{F.3})$$

whose solutions are exponential functions of time. If  $p_0(\sigma' = 1 \parallel \sigma)(t = 0) = p_0^{\sigma}$  is used as notation for the initial conditions, then:

$$\begin{aligned} p_0(\sigma' = 1 \parallel \sigma) &= \frac{e^{-t}}{2} [2p_0^\sigma - \tanh(\beta J \sigma) - 1] + \frac{1 + \tanh(\beta J \sigma)}{2} \\ p_0(\sigma' = -1 \parallel \sigma) &= 1 - p_0(\sigma' = 1 \parallel \sigma) \end{aligned} \quad (\text{F.4})$$

On the other hand, the quantity  $p_\infty(\sigma' \parallel \sigma)$  is the limit of  $p_\gamma(\sigma' \parallel \sigma)$  as  $\gamma \rightarrow \infty$ . A differential equation can also be written for it:

$$\begin{aligned} \frac{dp_\infty(\sigma' \parallel \sigma)}{dt} &= -\frac{1}{2} \sum_{\sigma''} \sigma' \sigma'' p_\infty(\sigma'' \parallel \sigma) + \frac{\sigma'}{2} \lim_{\gamma \rightarrow \infty} \sum_{n=0}^{\gamma} \binom{\gamma}{n} \times \\ &\quad \times \tanh [\beta J(2n - \gamma + \sigma)] \sum_{\sigma''} [\hat{p}(1 \parallel \sigma'')]^n [\hat{p}(-1 \parallel \sigma'')]^{\gamma-n} \end{aligned} \quad (\text{F.5})$$

When  $\gamma$  is sufficiently large, the binomial weights in Eq. (F.5) can be replaced by Gaussian weights and the sum can be replaced by an integral. We have:

$$\begin{aligned} \frac{dp_\infty(\sigma' \parallel \sigma)}{dt} &= -\frac{1}{2} \sum_{\sigma''} \sigma' \sigma'' p_\infty(\sigma'' \parallel \sigma) + \\ &+ \sum_{\sigma''} \frac{\sigma'}{2} \lim_{\gamma \rightarrow \infty} \int_{f_-(\gamma, \tilde{m}(\sigma''))}^{f_+(\gamma, \tilde{m}(\sigma''))} Dy \tanh [\beta J(\gamma \tilde{m}(\sigma'') + y \sqrt{\gamma} \sqrt{1 - \tilde{m}^2(\sigma'')} + \sigma)] \end{aligned} \quad (\text{F.6})$$

where  $Dy \equiv dy \exp(-y^2/2)$  and

$$f_\pm(\gamma, \tilde{m}(\sigma'')) = \pm \sqrt{\gamma} \sqrt{\frac{1 \mp \tilde{m}(\sigma'')}{1 \pm \tilde{m}(\sigma'')}}$$

The presence of the Gaussian weight  $Dy$  ensures that the integral is well-behaved. The limit and the integral can be interchanged to obtain:

$$\frac{dp_\infty(\sigma' \parallel \sigma)}{dt} = -\frac{1}{2} \sum_{\sigma''} \sigma' \sigma'' p_\infty(\sigma'' \parallel \sigma) [1 - \sigma'' \operatorname{sign}(\tilde{m}(\sigma''))] \quad (\text{F.7})$$

provided that  $\tilde{m}(\sigma'') \neq 0$ , which is generally of interest when studying the dynamics of the system.

This indicates that when  $p_\infty(\sigma' = 1 \parallel \sigma) = 1$  at  $t = 0$ , the time derivative of  $p_\infty(\sigma' \parallel \sigma)$  is zero for all times, and this quantity is constant. In any other case, the equations are equally easy to integrate numerically.

Now, in Eq. (F.2), we only need to calculate the value of the free parameter  $b$ . It is known that  $\hat{p}(\sigma' \parallel \sigma) = \sum_\gamma q(\gamma) p_\gamma(\sigma' \parallel \sigma)$ , so:

$$\begin{aligned}\hat{p}(\sigma' \parallel \sigma) &= p_\infty(\sigma' \parallel \sigma) - [p_\infty(\sigma' \parallel \sigma) - p_0(\sigma' \parallel \sigma)] \sum_{\gamma=0}^{\infty} q(\gamma) b^\gamma \\ \hat{p}(\sigma' \parallel \sigma) &= p_\infty(\sigma' \parallel \sigma) - [p_\infty(\sigma' \parallel \sigma) - p_0(\sigma' \parallel \sigma)] e^{-\kappa(1-b)}\end{aligned}\quad (\text{F.8})$$

Solving for  $b$ , we obtain:

$$b = 1 + \frac{1}{\kappa} \ln \left[ \frac{p_\infty(\sigma' \parallel \sigma) - \hat{p}(\sigma' \parallel \sigma)}{p_\infty(\sigma' \parallel \sigma) - p_0(\sigma' \parallel \sigma)} \right] \quad (\text{F.9})$$

Everything is now ready to substitute the *ansatz* into Eq. (F.1):

$$\begin{aligned}\frac{d\tilde{m}(\sigma)}{dt} &= -\tilde{m}(\sigma) + \\ &+ e^{-\kappa} \sum_{\sigma'} p_\infty(\sigma' \parallel \sigma) \sum_{\gamma=0}^{\infty} \frac{\kappa^\gamma}{\gamma!} \sum_{n=0}^{\gamma} \binom{\gamma}{n} \tanh[\beta J(2n - \gamma + \sigma)] [\hat{p}(1 \parallel \sigma')]^n [\hat{p}(-1 \parallel \sigma')]^{\gamma-n} \\ &- e^{-\kappa} \sum_{\sigma'} [p_\infty(\sigma' \parallel \sigma) - p_0(\sigma' \parallel \sigma)] \sum_{\gamma=0}^{\infty} \frac{(\kappa b)^\gamma}{\gamma!} \sum_{n=0}^{\gamma} \binom{\gamma}{n} \tanh[\beta J(2n - \gamma + \sigma)] \times \\ &\times [\hat{p}(1 \parallel \sigma')]^n [\hat{p}(-1 \parallel \sigma')]^{\gamma-n}\end{aligned}\quad (\text{F.10})$$

The Eqs. (F.4), (F.7), (F.9), and (F.10) form a closed system of average equations. After some algebra (see Appendix D), the following differential equation for  $\tilde{m}(\sigma)$  can be written:

$$\begin{aligned}\frac{d\tilde{m}(\sigma)}{dt} &= -\tilde{m}(\sigma) + e^{-\kappa} \sum_{\sigma'} p_\infty(\sigma' \parallel \sigma) S(\tilde{m}(\sigma''), \beta J, \kappa, \sigma) \\ &- e^{-\kappa} \sum_{\sigma'} [p_\infty(\sigma' \parallel \sigma) - p_0(\sigma' \parallel \sigma)] S(\tilde{m}(\sigma''), \beta J, \kappa b, \sigma)\end{aligned}\quad (\text{F.11})$$

As written in subsection 3.3.2, the function  $S(\tilde{m}(\sigma''), \beta J, q, \sigma)$  has the form:

$$S(\tilde{m}(\sigma''), \beta J, q, \sigma) = \sum_{n'=-\infty}^{\infty} \tanh(\beta J(n' + \sigma)) \left( \frac{1 + \tilde{m}(\sigma'')}{1 - \tilde{m}(\sigma'')} \right)^{\frac{n'}{2}} I_{n'}(q \sqrt{1 - \tilde{m}^2(\sigma'')}) \quad (\text{F.12})$$

where  $I_\nu(x)$  is the modified Bessel function of the first kind or the Infeild function.

To calculate observables, instead of Eq. (2.42), a slightly different Master Equation will be used:

$$\begin{aligned}\frac{dP^t(\sigma_i, \sigma_j)}{dt} &= - \sum_{\sigma'_i} \sigma_i \sigma'_i P^t(\sigma'_i, \sigma_j) \sum_{\sigma_{\partial i \setminus j}} r_i(\sigma'_i, \sigma_j) \prod_{k \in \partial i \setminus j} p^t(\sigma_k \parallel \sigma'_i) - \\ &- \sum_{\sigma'_j} \sigma_j \sigma'_j P^t(\sigma_i, \sigma'_j) \sum_{\sigma_{\partial j \setminus i}} r_j(\sigma'_j, \sigma_i) \prod_{k \in \partial j \setminus i} p^t(\sigma_k \parallel \sigma'_j)\end{aligned}\quad (\text{F.13})$$

Eq. (F.13) is written for the probabilities corresponding to a pair of interacting variables:  $\sigma_i$ ,  $\sigma_j$ . Similar to Eq. (2.42), to obtain this new closed Master Equation, a factorization of the joint probability of a variable and its neighbors is used, substituting the conditional probabilities with cavity conditional probabilities:  $P(\sigma_i, \sigma_{\partial i}) = P(\sigma_i, \sigma_j) \prod_{k \in \partial i \setminus j} p(\sigma_k \mid \sigma_i)$ .

With Eq. (F.13), a procedure analogous to the one outlined for the CME can be followed. Fig. (F.1) compares the numerical integration of the new equations with the average of the CME and Monte Carlo simulations on various ER graphs. For both magnetization and energy, both theoretical descriptions give very similar results.

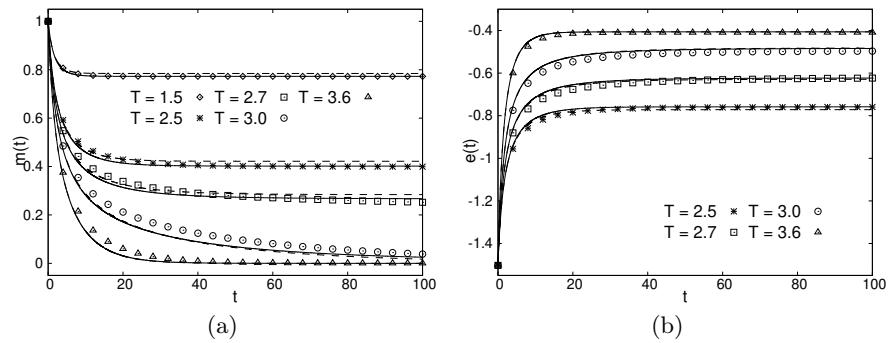


Figure F.1: New average equations (dashed lines) for a symmetric ferromagnet defined on Erdős-Rényi graphs with average connectivity  $\kappa = 3$  ( $T_c \approx 2.89$ ) and size  $N = 2000$ . In all cases, the system is initially fully magnetized and brought into contact with a thermal bath at a fixed temperature  $T$ . The solid lines are the average over different graphs of the integration of Eqs. (2.40) and (F.13). The points are the average of Monte Carlo simulations. In both cases,  $s = 300$  Erdős-Rényi graphs were used. For each,  $10^5$  Monte Carlo histories were run. **a)** Temporal dependence of the magnetization. **b)** Temporal dependence of the energy.

The *ansatz* used also allows obtaining the dependence on connectivity of the probabilities  $P_c(\sigma)$  through the equation:

$$P_c^{\text{AV}}(\sigma) = P_\infty(\sigma) - [P_\infty(\sigma) - P_0(\sigma)] \left\{ 1 + \frac{1}{\kappa} \ln \left[ \frac{P_\infty(\sigma) - \hat{P}(\sigma)}{P_\infty(\sigma) - P_0(\sigma)} \right] \right\}^c \quad (\text{F.14})$$

This equation is very similar to Eq. (F.2), and as before, expressions for  $P_\infty(\sigma)$  and  $P_0(\sigma)$  can be easily obtained.

At the same time, it should be noted that  $P_c(\sigma)$  is an average over local probabilities of nodes with connectivity  $c$ , taken over the family of Erdős-Rényi graphs with the same average connectivity  $\kappa$ . The original CME and Monte Carlo simulations allow the *experimental* estimation of these quantities through the numerical calculation of the average:

$$P_c^{\text{SI}}(\sigma) = \frac{\sum_{w=1} \sum_{i_w | c_{i_w} = c} P_{c_{i_w}}(\sigma_{i_w} = \sigma)}{\sum_{w=1} \sum_{i_w | c_{i_w} = c} 1} \quad (\text{F.15})$$

where  $w$  is an index running over all graphs and the index  $i_w$  runs over all nodes in graph  $w$ . The notation  $i_w | c_{i_w} = c$  indicates that the sum only considers nodes with connectivity  $c_{i_w} = c$ .

Figure (F.2) shows a comparison between the probabilities  $P_c^{\text{AV}}(\sigma)$ , obtained with the new average equations, and the probabilities  $P_c^{\text{SI}}(\sigma)$  obtained from the original CME and Monte Carlo simulations. It can be observed that the average equations not only describe the dynamics of magnetization and energy but also contain precise information about the structure of the local probabilities  $P_c^{\text{SI}}(\sigma)$ .

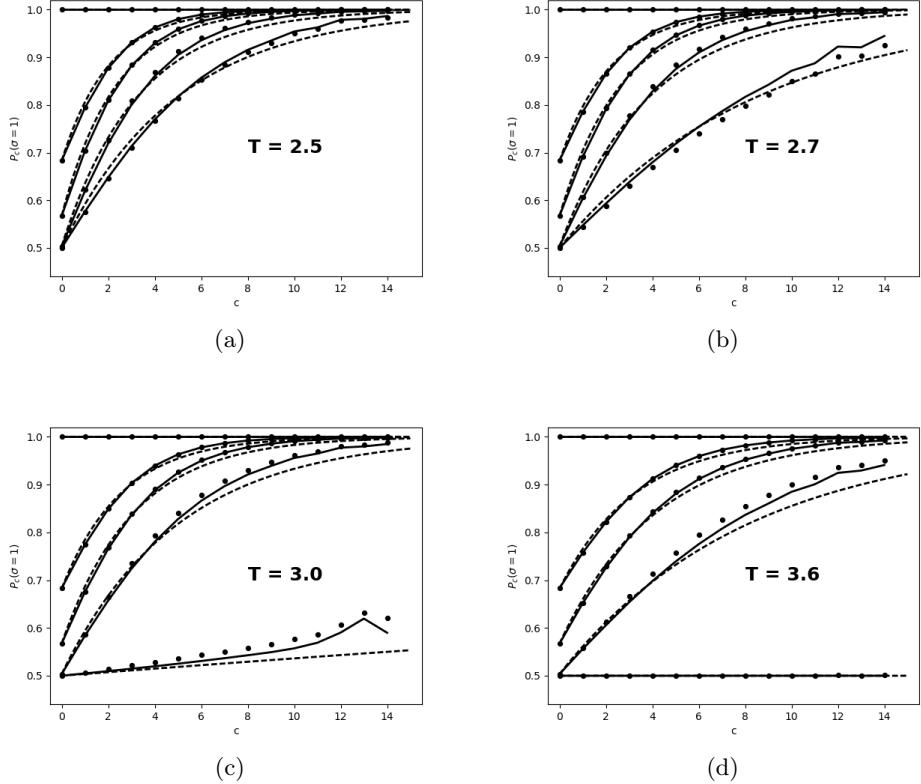


Figure F.2: Comparison between  $P_c^{\text{AV}}(\sigma = 1)$  and  $P_c^{\text{SI}}(\sigma = 1)$  (see Eqs. (F.14) and (F.15), respectively) for a symmetric ferromagnet defined on Erdős-Rényi graphs with average connectivity  $\kappa = 3$  ( $T_c \approx 2.89$ ) and size  $N = 2000$ . In all cases, the system is initially fully magnetized and brought into contact with a thermal bath at a fixed temperature  $T$ . The dashed lines represent the results of the new average equations. The solid lines are the numerical average over different graphs of the integration of Eqs. (2.40) and (F.13). The points are the average of Monte Carlo simulations. In both cases,  $s = 300$  Erdős-Rényi graphs were used. For each,  $10^5$  Monte Carlo histories were run. Each panel shows results for five different times. From top to bottom:  $t = \{0, 1, 2, 5, 90\}$ .

## Appendix G

# General closure for the hierarchy of master equations

Although in subsection 4.1.2 an initial idea was given on how to use the dynamic cavity method to close the Master Equation, here a more general presentation of the method will be provided, which allows constructing a hierarchical system of cavity master equations. This requires some notions about the structure of interactions since the different levels of approximation will be defined using the concept of distance in the graph.

The distance between two nodes  $d(k, l)$  will be the length of the shortest path between  $k$  and  $l$ . If  $k \neq l$ , then  $d$  is a positive integer. The connected node sets  $S = \{\mathcal{O} \setminus o_j, \mathcal{I}, o_j\}$  can be classified according to the maximum distance of the nodes from node  $o_j$ . The symbol  $\mathcal{O}$  represents again the set of external nodes and  $\mathcal{I}$  the set of internal nodes of  $S$  (the definitions of internal and external nodes are given in Section 4.1). The *origin node*  $o_j$  can be selected, without loss of generality, such that there is only one internal node  $i_j \in \mathcal{I}$  such that  $i_j$  is a neighbor of  $o_j$ . The maximum distance will be reached for some node  $o_k \in \mathcal{O}$  and will be denoted by the letter  $Z$ . Sets that have the same  $Z$  will belong to the *CS-Z* class. If all external nodes  $o_k$  are at the same distance from  $o_j$ , then the connected node set is *regular* and the *rCS-Z* notation will be used. Of course, elements of *rCS-Z* also belong to *CS-Z*. Figure (G.1) illustrates the meaning of this classification.

In this context, the notation can be used:

$$p^t(\vec{\sigma}_{\mathcal{O} \setminus o_j}, \vec{\sigma}_{\mathcal{I}} \parallel \sigma_{o_j}) \equiv p^t(\vec{\sigma}_Z, \vec{\sigma}_{Z-1}, \dots, \vec{\sigma}_2, \sigma_{i_j} \parallel \sigma_{o_j}) \quad (\text{G.1})$$

where the vector  $\vec{\sigma}_z$  ( $1 \leq z \leq Z$ ) contains the spins  $\sigma_k$  such that  $d(k, o_j) = z$ , with the condition that  $k \in \mathcal{O}$  or  $k \in \mathcal{I}$ . The probability distribution of Eq. (G.1) is then defined over a connected node set that belongs to *CS-Z*, for which the symbol  $p^t(\text{CS-Z})$  will be used. It will be simply said that  $p^t(\text{CS-Z})$  is defined *in CS-Z*.

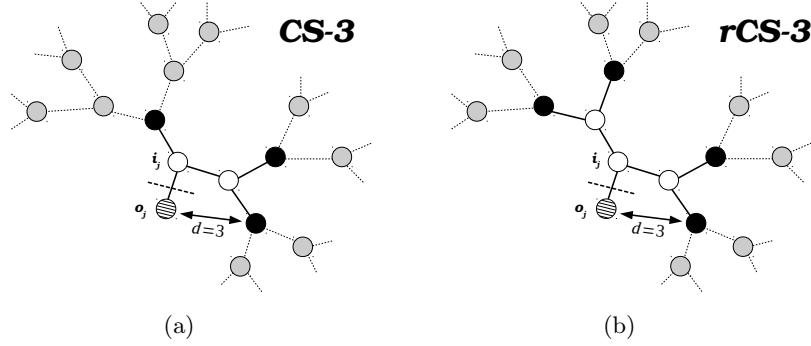


Figure G.1: The cavity conditional probability  $p^t(\sigma_{\mathcal{O} \setminus o_j}, \mathcal{I} \parallel o_j)$  is defined over a connected node set. This figure illustrates the classification of such sets according to the maximum distance from the origin node  $o_j$ . **a)** The maximum distance is  $Z = 3$  and the set belongs to the *CS-3* class. In this example, the distances of the external nodes from the origin  $o_j$  are not all the same. **b)** Each external node is at distance  $Z = 3$  from the origin node  $o_j$ , so the set belongs to *rCS-3* and therefore to *CS-3*.

With the new notation, in the cavity master equation corresponding to  $p^t(CS-Z)$ , probabilities defined in *CS-(Z+1)* appear (see Eq. (4.8)). It should be noted that each  $p^t(CS-Z)$  can be written as a marginal of a cavity conditional probability defined over the corresponding regular set *rCS-Z*. In view of this, what is needed is a system of equations of the form:

$$\frac{dp^t(rCS-Z)}{dt} = F[p^t(rCS-Z)] \quad (\text{G.2})$$

which in this case is equivalent to finding a relation  $p^t(CS-(Z+1)) = f[p^t(CS-Z)]$ .

With this classification of connected node sets, which essentially refers to the size of the set, it is possible to find the function  $f[\cdot]$ . It is given by:

$$p^t(\vec{\sigma}_{Z+1} \mid \vec{\sigma}_Z, \dots, \sigma_{i_j} \parallel \sigma_{o_j}) = \frac{p^t(\vec{\sigma}_{Z+1}, \vec{\sigma}_Z, \dots, \sigma_{i_j} \parallel \sigma_{o_j})}{p^t(\vec{\sigma}_Z, \dots, \sigma_{i_j} \parallel \sigma_{o_j})} \quad (\text{G.3})$$

If the distance  $Z + 1$  is large compared to the typical correlation length of the system, it can be assumed that the value of  $\sigma_{o_j}$  at time  $t$  is irrelevant for the configuration of  $\vec{\sigma}_{Z+1}$  once conditioned on the rest of the spins (see the left-hand side of Eq. (G.3)). Thus:

$$p^t(\vec{\sigma}_{Z+1} \mid \vec{\sigma}_Z, \dots, \sigma_{i_j} \parallel \sigma_{o_j}) \approx p^t(\vec{\sigma}_{Z+1} \mid \vec{\sigma}_Z, \dots, \vec{\sigma}_2, \sigma_{i_j}) \quad (\text{G.4})$$

Following the same reasoning, the dependence on all spins in  $\{\vec{\sigma}_Z, \dots, \sigma_{i_j}\}$  that are at a distance greater than  $Z$  from the nodes in  $\vec{\sigma}_{Z+1}$  can be removed. Approximations

like these have already been made to obtain *CME-1* and *CME-2* in 2.3 and 4.1.2, respectively. The remaining spins  $\{\vec{\sigma}_{Z+1}, \vec{\sigma}'_Z, \dots, \vec{\sigma}'_2, \sigma_{i_j}\}$  are defined over a connected set of the *CS-Z* class if  $i_j$  is taken as the new origin.

Moreover, when the correlations between  $\vec{\sigma}_{Z+1}$  and  $\sigma_{i_j}$  are low, the approximation [35] can be used:

$$p^t(\vec{\sigma}_{Z+1} | \vec{\sigma}'_Z, \dots, \vec{\sigma}'_2, \sigma_{i_j}) \approx p^t(\vec{\sigma}_{Z+1} | \vec{\sigma}'_Z, \dots, \vec{\sigma}'_2 \| \sigma_{i_j}) = \frac{p^t(\vec{\sigma}_{Z+1}, \vec{\sigma}'_Z, \dots, \vec{\sigma}'_2 \| \sigma_{i_j})}{p^t(\vec{\sigma}'_Z, \dots, \vec{\sigma}'_2 \| \sigma_{i_j})} \quad (\text{G.5})$$

It should be noted that the distance between  $i_j$  and all nodes in  $\vec{\sigma}_{Z+1}$  is  $Z$ , so it is assumed that correlations decay significantly at this distance. From Eqs. (G.4) and (G.5), it follows that:

$$p^t(\vec{\sigma}_{Z+1} | \vec{\sigma}_Z, \dots, \sigma_{i_j} \| \sigma_{o_j}) \approx \frac{p^t(\vec{\sigma}_{Z+1}, \vec{\sigma}'_Z, \dots, \vec{\sigma}'_2 \| \sigma_{i_j})}{\sum_{\vec{\sigma}_{Z+1}} p^t(\vec{\sigma}_{Z+1}, \vec{\sigma}'_Z, \dots, \vec{\sigma}'_2 \| \sigma_{i_j})} \quad (\text{G.6})$$

so that:

$$p^t(\vec{\sigma}_{Z+1}, \vec{\sigma}_Z, \dots, \sigma_{i_j} \| \sigma_{o_j}) \approx \frac{p^t(\vec{\sigma}_{Z+1}, \vec{\sigma}'_Z, \dots, \vec{\sigma}'_2 \| \sigma_{i_j})}{\sum_{\vec{\sigma}_{Z+1}} p^t(\vec{\sigma}_{Z+1}, \vec{\sigma}'_Z, \dots, \vec{\sigma}'_2 \| \sigma_{i_j})} p^t(\vec{\sigma}_Z, \dots, \sigma_{i_j} \| \sigma_{o_j})$$

While on the left-hand side there is a probability  $p^t(\text{CS-}(Z+1))$ , all probabilities on the right-hand side are of the form  $p^t(\text{CS-Z})$ . This means that the function  $f[p^t(\text{CS-Z})]$  has been found, which allows Eq. (G.2) to be written. It is important to note that after numerically solving Eq. (G.2), it is possible to calculate all probabilities defined in *CS-1*, *CS-2*, ..., *CS-(Z-1)* as they are simply marginals of  $p^t(\text{CS-Z})$ .

The same approach can be used to write a closed Master Equation similar to Eqs. (2.42) and (4.15). The local probabilities  $P^t$  are then used to calculate any observable defined locally.

It is important to state that these closure relations can be used in various ways, but the key point will be how to choose the distance  $Z$  at which correlations are neglected. This distance will be the parameter defining the *level of approximation* within the proposed hierarchical system.

In summary, a level of approximation was defined based on the form of the differential equations for the cavity conditional probabilities. If they are written for probabilities  $p^t(\text{CS-Z})$ , it is said that the equations are at level  $Z$  of approximation. In this sense, Eqs. (2.40) and (2.42) are the first level of approximation, while Eqs. (4.14), (4.13), and (4.15) are at the second level.



## Appendix H

# Average equations for random regular graphs

When initial conditions are site-independent and there is homogeneous connectivity in the graph, it is possible to replace the cavity master equations for different local probabilities with a single average equation. It is important to emphasize that these equations do not involve new approximations, contrary to what happened in Section 3.3. An exact representation of the behavior of the various approximation levels of the hierarchical system from Chapter 4 will be obtained. It just so happens that, in the particular case of interest here, these equations can be significantly simplified.

We will start with Eq. (4.11):

$$\begin{aligned} \frac{dp^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)}{dt} = & -r_i(\sigma_i, \sigma_{\partial i}) p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) + r_i(-\sigma_i, \sigma_{\partial i}) p^t(\sigma_{\partial i \setminus j}, -\sigma_i \| \sigma_j) \\ & - \sum_{l \in \partial i \setminus j} \sum_{\sigma_{\partial l \setminus i}} \left\{ r_l(\sigma_l, \sigma_{\partial l}) p^t(\sigma_{\partial l \setminus i} | \sigma_l \| \sigma_i) p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j) - \right. \\ & \left. - r_l(-\sigma_l, \sigma_{\partial l}) p^t(\sigma_{\partial l \setminus i} | -\sigma_l \| \sigma_i) p^t(F_l[\sigma_{\partial i \setminus j}], \sigma_i \| \sigma_j) \right\} \end{aligned} \quad (\text{H.1})$$

In a regular random graph with connectivity  $K$ , spatial symmetry is obtained by choosing an initial condition for  $p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)$  that does not depend on  $i$  or  $j$ . In this case, the equation governing the temporal evolution of  $p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)$  will not depend on the values of  $i, j$  for any time. The equation can be rewritten for all sites:

$$\begin{aligned} \frac{dp^t(\hat{u}, \sigma \| \sigma')}{dt} = & -r(K, \hat{u} + \delta_{\sigma, -\sigma'}) p^t(\hat{u}, \sigma \| \sigma') + \\ & + r(K, K - \hat{u} - \delta_{\sigma, -\sigma'}) p^t(K - \hat{u}, -\sigma \| \sigma') - \end{aligned}$$

$$\begin{aligned}
& - (K - 1 - \hat{u}) \sum_{u'=0}^{K-1} \binom{K-1}{u'} r(K, u') \left\{ p^t(u' | \sigma \| \sigma) p^t(\hat{u}, \sigma \| \sigma') - \right. \\
& \quad \left. - p^t(u' | -\sigma \| \sigma) p^t(\hat{u} + 1, \sigma \| \sigma') \right\} - \\
& - \hat{u} \sum_{u'=0}^{K-1} \binom{K-1}{u'} r(K, u') \left\{ p^t(u' | -\sigma \| \sigma) p^t(\hat{u}, \sigma \| \sigma') - \right. \\
& \quad \left. - p^t(u' | \sigma \| \sigma) p^t(\hat{u} - 1, \sigma \| \sigma') \right\} \tag{H.2}
\end{aligned}$$

In Eq. (H.2), regardless of the site where they are defined, the probabilities depend only on two neighboring spins ( $\sigma, \sigma'$ ) and the number  $\hat{u} = 0, \dots, K - 1$  of unsatisfied interactions that  $\sigma$  has with its neighbors excluding  $\sigma'$ . No new magnitudes have been defined so far; we have only denoted the probability  $p^t(\sigma_{\partial i \setminus j}, \sigma_i \| \sigma_j)$  as  $p^t(\hat{u}, \sigma \| \sigma')$  since the values of  $i$  and  $j$  become irrelevant. It is no longer necessary to distinguish between different configurations of  $\sigma_{\partial i \setminus j}$ . It is enough to know the number  $\hat{u}$  of unsatisfied interactions between  $\sigma_{\partial i \setminus j}$  and  $\sigma_i$ .

We will now introduce the probabilities  $\hat{p}^t(\hat{u}, \sigma \| \sigma') = \binom{K-1}{\hat{u}} p^t(\hat{u}, \sigma \| \sigma')$ , whose differential equation is:

$$\begin{aligned}
\frac{d\hat{p}^t(\hat{u}, \sigma \| \sigma')}{dt} &= -r(K, \hat{u} + \delta_{\sigma, -\sigma'}) \hat{p}^t(\hat{u}, \sigma \| \sigma') + \\
&\quad + r(K, K - \hat{u} - \delta_{\sigma, -\sigma'}) \hat{p}^t(K - \hat{u}, -\sigma \| \sigma') - \\
&\quad - (K - 1 - \hat{u}) \langle r(K, u') \rangle_{\sigma, \sigma} \hat{p}^t(\hat{u}, \sigma \| \sigma') + \\
&\quad + (\hat{u} + 1) \langle r(K, u') \rangle_{-\sigma, \sigma} \hat{p}^t(\hat{u} + 1, \sigma \| \sigma') - \\
&\quad - \hat{u} \langle r(K, u') \rangle_{-\sigma, \sigma} \hat{p}^t(\hat{u}, \sigma \| \sigma') + (K - \hat{u}) \langle r(K, u') \rangle_{\sigma, \sigma} \hat{p}^t(\hat{u} - 1, \sigma \| \sigma') \tag{H.3}
\end{aligned}$$

where  $\langle r(K, u') \rangle_{\sigma, \sigma'} = \sum_{u'=0}^{K-1} r(K, u') \hat{p}^t(u' | \sigma \| \sigma')$ . The cavity conditional probabilities in Eq. (H.3) are:

$$\hat{p}^t(\hat{u} | \sigma \| \sigma') = \frac{\hat{p}^t(\hat{u}, \sigma \| \sigma')}{\sum_{u'} \hat{p}^t(u', \sigma \| \sigma')} \tag{H.4}$$

Eqs. (H.3) and (H.4) form a closed system that can be numerically integrated. The results are used to integrate the equations for local probabilities:

$$\begin{aligned}
\frac{d\hat{P}^t(\sigma, u)}{dt} &= -r(K, u) \hat{P}^t(\sigma, u) + r(K, K - u) \hat{P}^t(-\sigma, K - u) - \\
&\quad - (K - u) \langle r(K, u') \rangle_{\sigma, \sigma} \hat{P}^t(\sigma, u) + (u + 1) \langle r(K, u') \rangle_{-\sigma, \sigma} \hat{P}^t(\sigma, u + 1) - \\
&\quad - u \langle r(K, u') \rangle_{-\sigma, \sigma} \hat{P}^t(\sigma, u) + (K - u + 1) \langle r(K, u') \rangle_{\sigma, \sigma} \hat{P}^t(\sigma, \hat{u} - 1) \tag{H.5}
\end{aligned}$$

With this, the average equations for the *CME-2* level are obtained. To complete all the levels present in Fig. (4.4), it is still necessary to write the equations for the first and third levels: *CME-1* and *CME-3*. In the first case, we start with Eqs. (2.40) and (2.42):

$$\frac{dp^t(\sigma_i \parallel \sigma_j)}{dt} = - \sum_{\sigma_{\partial i \setminus j}} \left\{ r_i(\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i \setminus j} p^t(\sigma_k \parallel \sigma_i) \right] p^t(\sigma_i \parallel \sigma_j) - \right. \\ \left. - r_i(-\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i \setminus j} p^t(\sigma_k \parallel -\sigma_i) \right] p^t(-\sigma_i \parallel \sigma_j) \right\} \quad (\text{H.6})$$

$$\frac{dP^t(\sigma_i)}{dt} = - \sum_{\sigma_{\partial i}} \left\{ r_i(\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i} p^t(\sigma_k \parallel \sigma_i) \right] P^t(\sigma_i) - \right. \\ \left. - r_i(-\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i} p^t(\sigma_k \parallel -\sigma_i) \right] P^t(-\sigma_i) \right\} \quad (\text{H.7})$$

It is easy to see that the average versions of these equations in regular random graphs with homogeneous initial conditions are:

$$\frac{d\hat{p}^t(\sigma \parallel \sigma')}{dt} = - \sum_{u'=0}^{K-1} \left\{ r(K, u' + \delta_{\sigma, -\sigma'}) [\hat{p}(-\sigma \mid \sigma)]^{u'} [\hat{p}(\sigma \mid \sigma)]^{K-1-u'} \hat{p}^t(\sigma \parallel \sigma') - \right. \\ \left. - r(K, u' + \delta_{\sigma, \sigma'}) [\hat{p}(\sigma \mid -\sigma)]^{u'} [\hat{p}(-\sigma \mid -\sigma)]^{K-1-u'} \hat{p}^t(-\sigma \parallel \sigma') \right\} \quad (\text{H.8})$$

$$\frac{d\hat{P}^t(\sigma)}{dt} = - \sum_{u'=0}^K \left\{ r(K, u') [\hat{p}(-\sigma \mid \sigma)]^{u'} [\hat{p}(\sigma \mid \sigma)]^{K-u'} \hat{P}^t(\sigma) - \right. \\ \left. - r(K, u') [\hat{p}(\sigma \mid -\sigma)]^{u'} [\hat{p}(-\sigma \mid -\sigma)]^{K-u'} \hat{P}^t(-\sigma) \right\} \quad (\text{H.9})$$

Eqs. (H.8) and (H.9) also form a closed system, and the results of their numerical integration are shown in Fig. (4.4).

The equations for the third-level approximation, *CME-3*, require a bit of extra work. The probabilities are now  $p^t(\hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \parallel \sigma_0)$ , which are defined over the connected sets illustrated in Fig. (H.1). There, the spins  $\sigma_0$ ,  $\sigma_1$ , and  $\sigma_2$  are represented. The integer  $\hat{u}_2$  represents the number of unsatisfied interactions between  $\sigma_2$  and its neighbors, excluding  $\sigma_1$ . The other integer  $\tilde{u}_1$  is the number of unsatisfied interactions between  $\sigma_1$  and its neighbors, excluding  $\sigma_0$  and  $\sigma_2$ .

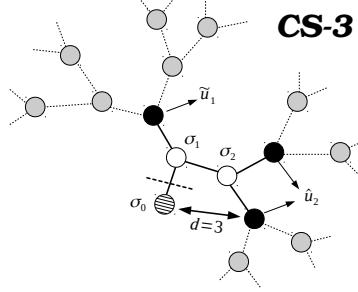


Figure H.1: The cavity conditional probability  $p^t(\hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0)$  is defined over a connected set belonging to the *CS-3* class (see Appendix G). The figure illustrates the meaning of each variable  $\hat{u}_2$ ,  $\sigma_2$ ,  $\tilde{u}_1$ ,  $\sigma_1$ , and  $\sigma_0$ .

As indicated by Eq. (4.8), the equations for these probabilities are:

$$\begin{aligned}
\frac{dp^t(\hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0)}{dt} = & -r(K, \tilde{u}_1 + \delta_{\sigma_0, -\sigma_1} + \delta_{\sigma_2, -\sigma_1}) p^t(\hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) + \quad (\text{H.10}) \\
& + r(K, K - \tilde{u}_1 - \delta_{\sigma_0, -\sigma_1} - \delta_{\sigma_2, -\sigma_1}) p^t(\hat{u}_2, \sigma_2, K - 2 - \tilde{u}_1, -\sigma_1 \| \sigma_0) - \\
& - r(K, \hat{u}_2 + \delta_{\sigma_1, -\sigma_2}) p^t(\hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) + \\
& + r(K, K - \hat{u}_2 - \delta_{\sigma_1, -\sigma_2}) p^t(K - 1 - \hat{u}_2, -\sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) - \\
& - (K - 2 - \tilde{u}_1) \sum_{u'_1=0}^{K-1} \binom{K-1}{u'_1} r(K, u'_1) \left\{ p^t(u'_1 | \hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) p^t(\hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) - \right. \\
& \quad \left. - p^t(u'_1 | \hat{u}_2, \sigma_2, \tilde{u}_1 + 1, \sigma_1 \| \sigma_0) p^t(\hat{u}_2, \sigma_2, \tilde{u}_1 + 1, \sigma_1 \| \sigma_0) \right\} - \\
& - \tilde{u}_1 \sum_{u'_1=0}^{K-1} \binom{K-1}{u'_1} r(K, u'_1) \left\{ p^t(u'_1 | \hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) p^t(\hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) - \right. \\
& \quad \left. - p^t(u'_1 | \hat{u}_2, \sigma_2, \tilde{u}_1 - 1, \sigma_1 \| \sigma_0) p^t(\hat{u}_2, \sigma_2, \tilde{u}_1 - 1, \sigma_1 \| \sigma_0) \right\} - \\
& - (K - 1 - \hat{u}_2) \sum_{u'_2=0}^{K-1} \binom{K-1}{u'_2} r(K, u'_2) \left\{ p^t(u'_2 | \hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) p^t(\hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) - \right. \\
& \quad \left. - p^t(u'_2 | \hat{u}_2 + 1, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) p^t(\hat{u}_2 + 1, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) \right\} - \\
& - \hat{u}_2 \sum_{u'_2=0}^{K-1} \binom{K-1}{u'_2} r(K, u'_2) \left\{ p^t(u'_2 | \hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) p^t(\hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) - \right. \\
& \quad \left. - p^t(u'_2 | \hat{u}_2 - 1, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) p^t(\hat{u}_2 - 1, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) \right\}
\end{aligned}$$

Unlike the previous ones, the variables in Eq. (H.10) are not defined over regular connected sets (see Fig. (H.1)). A somewhat different closure needs to be introduced, or more specifically, two new closures. This is necessary because the derivative of the

probabilities appearing in Eq. (H.10) depends on probabilities defined over connected sets of the *CS-4* class.

The first approximation focuses on the cavity conditional probabilities appearing between the fifth and eighth lines of Eq. (H.10). These quantities are reduced to:

$$p_{\text{SAT}}^t(u'_1 | \hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) \approx \frac{p^t(u'_1, \sigma_1, \bar{u}_1, \sigma_1 \| \sigma_0)}{\sum_{u'} p^t(u', \sigma_1, \bar{u}_1, \sigma_1 \| \sigma_0)} \quad (\text{H.11})$$

$$p_{\text{UNSAT}}^t(u'_1 | \hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) \approx \frac{p^t(u'_1, -\sigma_1, \bar{u}_1, \sigma_1 \| \sigma_0)}{\sum_{u'} p^t(u', -\sigma_1, \bar{u}_1, \sigma_1 \| \sigma_0)} \quad (\text{H.12})$$

Fig. (H.2a) illustrates the meaning of these equations. Although the fixed node at time  $t$  remains  $\sigma_0$ , the dependence of all nodes on the right side of the black line is eliminated, which in the left-hand side of Eqs. (H.11) and (H.12) corresponds to the nodes in  $\hat{u}_2$ . It can be seen that they are at a distance  $d = 4 > 3$  from the external nodes marked with the letter  $u'_1$  in the figure.

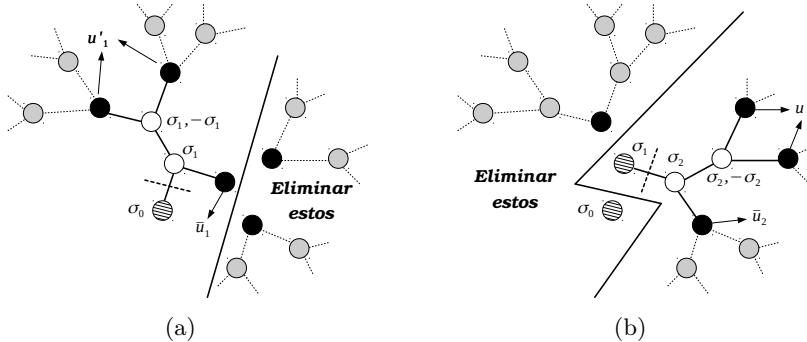


Figure H.2: Illustration of the closure relations (H.11), (H.12), (H.13), and (H.14) applied to cavity conditional probabilities. In each panel, the nodes not colored in gray belong to the connected set where the probability to be approximated is defined (left-hand sides of Eqs. (H.11), (H.12), (H.13), and (H.14)). It is clearly indicated which nodes' dependence is removed to close Eq. (H.10).

The second approximation will be written for the probabilities appearing between the ninth and twelfth lines of Eq. (H.10):

$$p_{\text{SAT}}^t(u'_2 | \hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) \approx \frac{p^t(u'_2, \sigma_2, \bar{u}_2, \sigma_2 \| \sigma_1)}{\sum_{u'} p^t(u', \sigma_2, \bar{u}_2, \sigma_2 \| \sigma_1)} \quad (\text{H.13})$$

$$p_{\text{UNSAT}}^t(u'_2 | \hat{u}_2, \sigma_2, \tilde{u}_1, \sigma_1 \| \sigma_0) \approx \frac{p^t(u'_2, -\sigma_2, \bar{u}_2, \sigma_2 \| \sigma_1)}{\sum_{u'} p^t(u', -\sigma_2, \bar{u}_2, \sigma_2 \| \sigma_1)} \quad (\text{H.14})$$

Fig. (H.2b) illustrates the meaning of these equations. In this case, the node on which the cavity condition is imposed is now  $\sigma_1$ , which is fixed at time  $t$  for all realizations of the dynamics. The dependence of those nodes that are at a distance  $d = 4$  from the nodes marked with the letter  $u'_2$  is eliminated. In the left-hand side of Eqs. (H.13) and (H.14), the removed nodes correspond to  $\sigma_0$  and  $\tilde{u}_1$ .

Eqs. (H.11), (H.12), (H.13), and (H.14) combine with the differential equation (H.10) to form a closed system that can be integrated numerically. The results appear in Fig. (4.4).

## Appendix I

# Simplification of the CME in the K-SAT problem

In this appendix, we will show how to simplify the CME equations under fairly general assumptions about the form of the time-unit transition probabilities  $r_i(\sigma_i, \sigma_{\partial i})$ .

A clause is said to be partially satisfied by the variables in  $a \setminus i$  if:

$$E_{a \setminus i} = \frac{1}{2^{K-1}} \prod_{j \in a \setminus i} (1 - l_j^a \sigma_j) = 0 \quad (\text{I.1})$$

Let  $\{\sigma_{a \setminus i SAT}\}$  be the set of configurations of the variables  $\sigma_{a \setminus i}$  that satisfy the clause  $a$ . Similarly,  $\sigma_{a \setminus i UNSAT}$  will be the unique configuration that does not satisfy the clause and for which  $E_{a \setminus i} = 1$ .

We can then define two variables:

$$p_u(\sigma_{a \setminus i} \parallel \sigma_i) = p(\sigma_{a \setminus i UNSAT} \parallel \sigma_i) \quad (\text{I.2})$$

$$p_s(\sigma_{a \setminus i} \parallel \sigma_i) = \sum_{\sigma_{a \setminus i} \in \{\sigma_{a \setminus i SAT}\}} p(\sigma_{a \setminus i SAT} \parallel \sigma_i) \quad (\text{I.3})$$

We now assume that the value of  $r_i(\sigma_i, \sigma_{\partial i})$  is the same for all combinations  $\{\sigma_{b \setminus i SAT}\}$ , with  $b \subset \partial i$ . That is, in the rules for the dynamics, only whether the clauses to which  $\sigma_i$  belongs are partially satisfied matters. Using Eqs. (I.2) and (I.3), the CME can be rewritten by summing over all configurations that partially satisfy the clauses. Introducing the binary variable  $S_{b \setminus j}$  such that  $S_{b \setminus j} = 1$  when  $b$  is partially satisfied and  $S_{b \setminus j} = -1$  when it is not, we obtain:

$$\begin{aligned} \frac{p(\sigma_{a \setminus i} \parallel \sigma_i)}{dt} &= \sum_{j \in a \setminus i} \sum_{S_{\partial j \setminus a}} \left\{ r_j(\sigma_j, S_{\partial j \setminus a}, \sigma_{a \setminus j}) \left[ \prod_{b \subset \partial j \setminus a} p(S_{b \setminus j} \parallel \sigma_j) \right] p(\sigma_{a \setminus i} \parallel \sigma_i) - \right. \\ &\quad \left. - r_j(-\sigma_j, S_{\partial j \setminus a}, \sigma_{a \setminus j}) \left[ \prod_{b \in \partial j \setminus a} p(S_{b \setminus j} \parallel -\sigma_j) \right] p(F_j[\sigma_{a \setminus i}] \parallel \sigma_i) \right\} \end{aligned} \quad (\text{I.4})$$

In Eq. (I.4), the probability  $p(S_{b \setminus j} \parallel \sigma_j)$  has two possible values corresponding to  $p_u$  and  $p_s$  (see Eqs. (I.2) and (I.3)).

With this procedure, the number of operations involved in the right-hand side of the CME has been considerably reduced. Previously, one had to sum over all  $2^{(K-1)c}$  configurations of the variables, where  $c$  is the connectivity of the node. In Eq. (I.4), one only needs to sum over the configurations of the new variables  $S_{b \setminus i}$ , which are  $2^c$ . This can be computationally implemented with a relatively low cost if the values of  $p_u$  and  $p_s$  are calculated at each step.

## Appendix J

# Determination of the algorithmic transition

In Fig. (5.4), the results of the CME numerically integrated with the appropriate  $r_i$  rules to describe the FMS algorithm are shown. Two qualitatively different behaviors appear: the CME curves either converge to zero quickly or stabilize at some finite value of the amount  $e$  of unsatisfied clauses per variable. However, to estimate the limit value  $\alpha_{\text{CME}}(\eta)$  that separates the two regions, a more rigorous criterion must be defined.

The first thing to note is that, unlike the algorithm, the CME does not directly work with the assignments of the K-SAT boolean variables. To decide whether it converges to zero, only the continuous function  $e(t)$  is available. Below, it will be explained why logarithmic scale graphs are helpful in this task and how they were used to estimate  $\alpha_{\text{CME}}(\eta)$  for any value of  $\eta$ . It will be seen that in this log-log scale, it is possible to distinguish when  $e(t)$  converges polynomially to zero simply by looking at its convexity. Polynomial time convergence occurs if the solution time  $t^*$  is bounded by some polynomial of the number of input parameters, or equivalently, the system size  $N$ .

If a logarithmic scale is chosen for both axes, a power function with a negative exponent  $f(x) = Ax^{-b}$  becomes a straight line  $F(z) = -bz + a$ , where  $z = \log(x)$ ,  $F(z) = \log(f(x))$ , and  $a = \log(A)$ . As illustrated in Fig. (J.1), if a curve  $G(z)$  is below the line  $-bz + a$  in logarithmic scale, then the corresponding  $g(x)$  converges to zero faster than a power law  $Ax^{-b}$ .

What has been explained has important implications for the dynamics of the algorithms. In this context, it is necessary to specify what exactly it means for the curve  $e(t)$  obtained with the CME to converge to zero. Recalling that  $e(t)$  is an intensive quantity, in a system of  $N$  variables its fluctuations should be proportional to  $1/\sqrt{N}$ . If a threshold  $e_{\text{th}}(N) = D/\sqrt{N}$  is introduced, the CME can be used to predict the time  $t^*$  it takes for the algorithm's dynamics to reach  $e(t^*) = e_{\text{th}}$ . From that moment, the typical fluctuations of the quantity  $e$  can lead the system to a solution.

What happens if the curve  $e(t)$  is below some line in the logarithmic scale of figures such as (5.4) and (J.2)? In that case, the relation holds (see Fig. (J.1)):

$$e(t^*) < A(t^*)^{-b} \Rightarrow \frac{D}{\sqrt{N}} < A(t^*)^{-b}$$

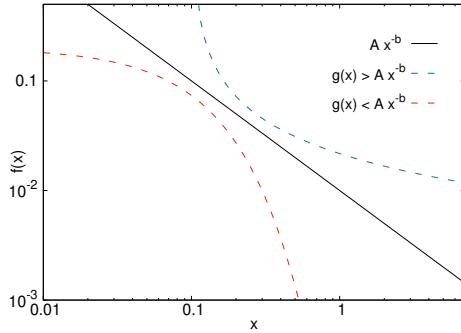


Figure J.1: Schematic representation of the relationships between functions in logarithmic scale. In the graph, the functions below a line decrease faster than a power law (red dashed line). The functions above a line decrease slower than a power law (blue dashed line).

Recall that the time  $t$  is defined as the number of iterations  $k$  of the algorithm divided by the system size  $N$  (see Section 5.3). Therefore

$$\frac{D}{\sqrt{N}} < A \left( \frac{k}{N} \right)^{-b} \Rightarrow k < A' N^{1/2b+1} \quad (\text{J.1})$$

where  $A' = (A/D)^{1/b}$ .

The last inequality of (J.1) means that the actual runtime of the algorithm is bounded by a polynomial of the system size  $N$ . In summary: **if a curve  $e(t)$  is obtained that is below some line for long times, the CME predicts that the algorithm solves the problem in polynomial time.** Similarly, it can be seen that if  $e(t)$  is above a line for long times, the solution cannot be polynomial.

With this, to determine the prediction  $\alpha_{\text{CME}}(\eta)$ , instances of K-SAT with different values of  $\alpha$  are taken. The CME is numerically integrated on each one, and the convexity is observed. Fig. (J.2) shows the results for  $\eta = 0.75$  and five different values of  $\alpha$ . For each  $\alpha$ , there are 20 different instances of size  $N = 2000$ . It is observed that for  $\alpha = (2.3, 2.4, 2.5)$  all the  $e(t)$  curves are below a line for long times. They are convex upward and converge rapidly to zero. When studying  $\alpha = 2.7$ , curves are obtained that are above a line for long times. They are convex downward. The case of  $\alpha = 2.6$  is more diverse, although most instances continue to converge rapidly to zero. It is then said that  $\alpha_{\text{CME}}(\eta)$  should be between  $\alpha = 2.6$  and  $\alpha = 2.7$ . Ideally, it is the point where exactly half of the curves are convex upward and the other

half downward. As a first approximation, simply choose  $\alpha_{\text{CME}}(\eta) \approx 2.65$ . The same procedure can be applied for any  $\eta$  to obtain Fig. (5.6).

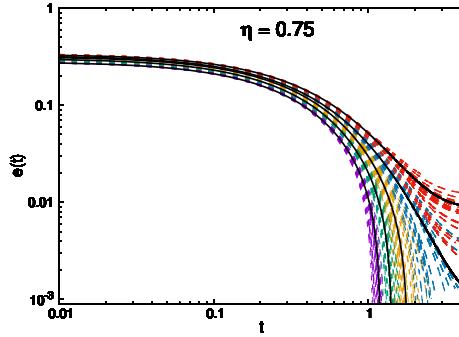


Figure J.2: Finite size effects on the algorithmic transition prediction with the CME. The curves are the result of integrating the CME with  $N = 5000$  (black lines) and with  $N = 2000$  (colored dashed lines).  $\eta = 0.75$  was used. From left to right,  $\alpha = (2.3, 2.4, 2.5, 2.6, 2.7)$ . The transition occurs approximately at  $\alpha \approx 2.65$  regardless of  $N$ . For  $N = 2000$ , the CME was integrated over 20 different instances for each value of  $\alpha$ .

Fig. (J.2) also compares the results of  $N = 2000$  with those of integrating the CME for a larger size. The black curves were obtained with  $N = 5000$ . It is important to ensure that the prediction for the algorithmic transition  $\alpha_{\text{CME}}(\eta)$  is not heavily dependent on  $N$ . It can be verified that for the values of  $\alpha$  where all the colored lines ( $N = 2000$ ) have the same qualitative behavior, this coincides with what is obtained for the black curve ( $N = 5000$ ).



## Appendix K

# The Instantaneous Master Equation

This appendix presents an original result from the thesis published in Ref. [41]. A new closure for the Master Equation derived from the Theory of Point Processes (TRPP) will be deduced. In Chapter 2, it was explained how to obtain the Cavity Master Equation (CME) from TRPP. Similarly, in Section K.1, a new Instantaneous Cavity Master Equation (iCME) is derived from TRPP. Later, in K.2, the numerical results of iCME are compared with Monte Carlo simulations and the numerical integration of CME in the Ising ferromagnetic model and in the Viana-Bray model for spin glasses.

### K.1 Closure of the Master Equation

In Chapter 2, dynamic cavity messages  $\mu_{i \rightarrow (ij)}(X_i | X_j)$  were introduced, which satisfy the message propagation equation:

$$\mu_{i \rightarrow (ij)}^t(X_i | X_j) = \sum_{X_{\partial i \setminus j}} \Phi_i^t(X_i | X_{\partial i}) \prod_{k \in \partial i \setminus j} \mu_{k \rightarrow (ki)}^t(X_k | X_i) \quad (\text{K.1})$$

where  $\Phi_i^t$  is a local weight expressing the probability of history  $X_i$  given the histories of its neighboring spins  $X_{\partial i}$ . The key to CME is to obtain differential equations for the marginals

$$p^t(\sigma_i \| X_j) = \sum_{X_i | \sigma_i}^{[t_0, t]} \mu_{i \rightarrow (ij)}^t(X_i | X_j) \quad (\text{K.2})$$

These cavity conditional probabilities still depend on the full trajectory  $X_j$  associated with site  $j$ . To close the equations, in Section 2.3,  $X_j$  was replaced by its final value  $\sigma_j$  at time  $t$ .

The objective of this section is to avoid making this last uncontrolled approximation. The result will be an equation for quantities that, by their definition, do

not depend on the full trajectory  $X_j$  but only on its final value. Hence, the name Instantaneous Cavity Master Equation (iCME) for the new closure.

Following the definitions from Ref. [83], an auxiliary graph is constructed in which all neighbors of node  $j$  are removed except for one. In this case,  $j$  becomes an *end node*, with a single neighbor in the graph denoted by  $i$ . The joint probability of the trajectories  $X_j$  and  $X_i$  in the auxiliary graph is then defined as:

$$m_{i \rightarrow (ij)}^t(X_i, X_j) = \mu_{i \rightarrow (ij)}^t(X_i | X_j) \Phi^t(X_j | X_i) \quad (\text{K.3})$$

The notation  $m_{i \rightarrow (ij)}$  indicates that this is a cavity probability. While in Section 2.2 conditional cavity probabilities  $\mu_{i \rightarrow (ij)}(X_i | X_j)$  were defined in an auxiliary stochastic process where the trajectory  $X_j$  is fixed as an external parameter, now a different modification is introduced where  $j$  is an *end node*.

Analogous to Section 2.3, a differential equation must be found for the marginal probabilities:

$$p_{i \rightarrow (ij)}^t(\sigma_i, \sigma_j) = \sum_{X_i | \sigma_i}^t \sum_{X_j | \sigma_j}^t m_{i \rightarrow (ij)}^t(X_i, X_j) \quad (\text{K.4})$$

It is necessary to expand the sum on the right side of Eq. (K.4) up to order  $\Delta t$ . More explicitly, the sums are expanded as follows:

$$\begin{aligned} \sum_{X_i | \sigma_i}^{t+\Delta t} \sum_{X_j | \sigma_j}^{t+\Delta t} m_{i \rightarrow (ij)}^{t+\Delta t}(X_i, X_j) &= \sum_{s_i} \int_{t_0}^{t+\Delta t} dt_1^i \dots \int_{t_{s_i-1}^i}^{t+\Delta t} dt_{s_i}^i \sum_{s_j} \int_{t_0}^{t+\Delta t} dt_1^j \dots \\ &\quad \dots \int_{t_{s_j-1}^j}^{t+\Delta t} dt_{s_j}^j m_{i \rightarrow (ij)}^{t+\Delta t}(X_i, X_j) \end{aligned} \quad (\text{K.5})$$

where  $s_i$  is the number of jumps in trajectory  $X_i$ , chosen such that the final state is always  $\sigma_i$ . The jumps occur at times  $\{t_1^i, \dots, t_{s_i}^i\}$ . For trajectory  $X_j$ , similarly, quantities  $s_j$  and  $\{t_1^j, \dots, t_{s_j}^j\}$  are defined.

To keep only the terms of order  $O(\Delta t)$ , only two possibilities can be allowed:

- a) All integrals are taken up to time  $t$ , which means no jumps occur between  $t$  and  $t + \Delta t$ .
- b) Only one integral is taken between  $t$  and  $t + \Delta t$ , which means only one jump occurs in that interval. This can correspond to either  $\sigma_i$  or  $\sigma_j$ .

Recalling Eqs. (2.28) and (2.29)

$$\mu_{i \rightarrow (ij)}^{t+\Delta t}(X_i | X_j) = [1 - \lambda^t(X_i, X_j) \Delta t] \mu_{i \rightarrow (ij)}^t(X_i | X_j) + o(\Delta t) \quad (\text{K.6})$$

$$\Phi^{t+\Delta t}(X_j | X_i) = [1 - r(\sigma_j, \sigma_i) \Delta t] \Phi^t(X_j | X_i) + o(\Delta t) \quad (\text{K.7})$$

the first of these contributions can be written as:

$$I_0 = p_{i \rightarrow (ij)}^t(\sigma_i, \sigma_j) - \Delta t r(\sigma_j, \sigma_i) p_{i \rightarrow (ij)}^t(\sigma_i, \sigma_j) - \\ - \Delta t \sum_{X_i | \sigma_i}^t \sum_{X_j | \sigma_j}^t \lambda^t(X_i, X_j) m_{i \rightarrow (ij)}^t(X_i, X_j) \quad (\text{K.8})$$

What happens when the last jump occurs between  $t$  and  $t + \Delta t$  is slightly different. In this case, the local weight  $\Phi^{t+\Delta t}(X_j | X_i)$  can be expressed as follows:

$$\Phi^{t+\Delta t}(X_j | X_i) = \Phi^t(X_j^- | X_i) r_j(-\sigma_j, \sigma_i) e^{-(t_{sj} - t) r(-\sigma_j, \sigma_i)} \times \\ \times e^{-(t + \Delta t - t_{sj}) r(\sigma_j, \sigma_i)} \quad (\text{K.9})$$

where  $X_j^-$  is a trajectory that ends with the value  $X_j^-(t) = -\sigma_j$ .

Since Eq. (K.9) will be within an integral that is already of order  $\Delta t$ , only the zero-order terms can be retained:

$$\Phi^{t+\Delta t}(X_j | X_i) = \Phi^t(X_j^- | X_i) r_j(-\sigma_j, \sigma_i) + O(\Delta t) \quad (\text{K.10})$$

The second contribution is then:

$$I_1 = \Delta t r(-\sigma_j, \sigma_i) p_{i \rightarrow (ij)}^t(\sigma_i, -\sigma_j) \quad (\text{K.11})$$

In the case where only  $\sigma_i$  jumps in the interval  $[t, t + \Delta t]$ , we have:

$$I_2 = \Delta t \sum_{X_i^- | -\sigma_i}^t \sum_{X_j | \sigma_j}^t \lambda^t(X_i^-, X_j) m_{i \rightarrow (ij)}^t(X_i^-, X_j) \quad (\text{K.12})$$

From Eqs. (K.8) and (K.12), it is still necessary to eliminate the time unit probabilities  $\lambda^t$  whose form is unknown. However, similarly to the deduction of the CME (see Eq. (2.34)), the relation can be obtained:

$$\lambda^t(X_i, X_j) m_{i \rightarrow (ij)}(X_i, X_j) = \sum_{\sigma_{\partial i \setminus j}} r_i(\sigma_i, \sigma_{\partial i \setminus j}, \sigma_j) p_{i \rightarrow (ij)}^t(\sigma_{\partial i \setminus j}, X_i, X_j) \quad (\text{K.13})$$

With this, it is easy to rewrite the contributions (K.8) and (K.12). Combining the results with Eq. (K.11) gives an exact differential equation on trees for the cavity probabilities of a pair of variables:

$$\frac{dp_{i \rightarrow (ij)}(\sigma_i, \sigma_j)}{dt} = - \sum_{\sigma_{\partial i \setminus j}} r_i(\sigma_i, \sigma_{\partial i}) p_{i \rightarrow (ij)}(\sigma_{\partial i \setminus j}, \sigma_i, \sigma_j) + \\ + \sum_{\sigma_{\partial i \setminus j}} r_i(-\sigma_i, \sigma_{\partial i}) p_{i \rightarrow (ij)}(\sigma_{\partial i \setminus j}, -\sigma_i, \sigma_j) - \\ - r_j(\sigma_j, \sigma_i) p_{i \rightarrow (ij)}^t(\sigma_i, \sigma_j) + r_j(-\sigma_j, \sigma_i) p_{i \rightarrow (ij)}^t(\sigma_i, -\sigma_j) \quad (\text{K.14})$$

Equation (K.14) can be closed using an approximate expression for the joint probabilities  $p_{i \rightarrow (ij)}(\sigma_{\partial i \setminus j}, \sigma_i, \sigma_j)$  in terms of the pair probabilities  $p_{i \rightarrow (ij)}(\sigma_i, \sigma_j)$ . The result is:

$$\begin{aligned} \frac{dp_{i \rightarrow (ij)}(\sigma_i, \sigma_j)}{dt} = & - \sum_{\sigma_{\partial i \setminus j}} \left\{ r_i(\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i \setminus j} p_{k \rightarrow (ki)}(\sigma_k | \sigma_i) \right] p_{i \rightarrow (ij)}(\sigma_i, \sigma_j) - \right. \\ & - r_i(-\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i \setminus j} p_{k \rightarrow (ki)}(\sigma_k | -\sigma_i) \right] p_{i \rightarrow (ij)}(-\sigma_i, \sigma_j) \Big\} - \\ & - r_j(\sigma_j, \sigma_i) p_{i \rightarrow (ij)}(\sigma_i, \sigma_j) + r_j(-\sigma_j, \sigma_i) p_{i \rightarrow (ij)}(\sigma_i, -\sigma_j) \quad (\text{K.15}) \end{aligned}$$

where the factorization was applied

$$p_{i \rightarrow (ij)}(\sigma_{\partial i \setminus j}, \sigma_i, \sigma_j) = \left[ \prod_{k \in \partial i \setminus j} p_{k \rightarrow (ki)}(\sigma_k | \sigma_i) \right] p_{i \rightarrow (ij)}(\sigma_i, \sigma_j)$$

and the conditional probabilities have been defined as  $p_{i \rightarrow (ij)}(\sigma_i | \sigma_j) = \frac{p_{i \rightarrow (ij)}(\sigma_i, \sigma_j)}{\sum_s p_{i \rightarrow (ij)}(s, \sigma_j)}$

The rules  $r_i(\sigma_i, \sigma_{\partial i})$  and  $r_j(\sigma_j, \sigma_i)$  are the transition probabilities per unit time of the spins  $\sigma_j$  and  $\sigma_i$  when  $j$  is an *extreme node* (it only has  $i$  as its neighbor).

Equation (K.15) is used to complement the differential equation for the single-site probabilities  $P_i(\sigma_i) = \sum_{X_i^t: \sigma_i(t)=\sigma_i} Q_i^t(X_i^t)$ . Finally:

$$\begin{aligned} \frac{d}{dt} P_i(\sigma_i) = & \sum_{\sigma_{\partial i}} \left\{ r_i(\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i} p_{k \rightarrow (ki)}(\sigma_k | \sigma_i) \right] P_i(\sigma_i) - \right. \\ & \left. - r_i(-\sigma_i, \sigma_{\partial i}) \left[ \prod_{k \in \partial i} p_{k \rightarrow (ki)}(\sigma_k | -\sigma_i) \right] P_i(-\sigma_i) \right\}. \quad (\text{K.16}) \end{aligned}$$

In summary, Eqs. (K.15) and (K.16) constitute a closure for the Master Equation that allows the calculation of local probabilities and is directly written for probabilities that only depend on the states of the variables at time  $t$ . This technique will be called the Instantaneous Cavity Master Equation (iCME) and its numerical results will be shown in the following section.

## K.2 Numerical results

In Fig. (K.1), the descriptions of CME and iCME for the Ising ferromagnetic model are compared in two different types of graphs. Fig. (K.1a) applies the techniques to the one-dimensional ferromagnet where the spins align forming a chain. Its exact solution was deduced by Glauber in 1963 [70]. In this simple case, the original CME is surprisingly far from the solution, while the new closure (iCME) allows for considerably more accurate results.

On the other hand, Fig. (K.1b) presents the dynamics of the Ising ferromagnetic model defined on Erdös-Rényi graphs. Here, each spin interacts with  $k$  neighbors

and  $k$  follows a Poisson distribution with mean  $\kappa$ . The CME and iCME exhibit very similar behavior to the dynamics of Monte Carlo simulations. However, it is possible to appreciate that iCME provides a small improvement. The local errors in the inset graph point towards this conclusion for all temperatures studied.

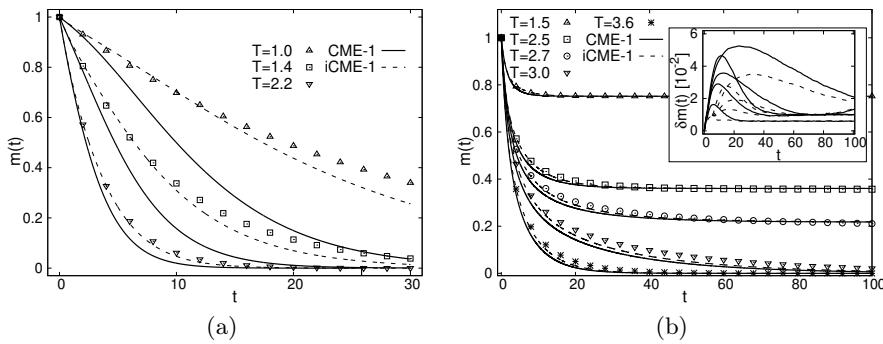


Figure K.1: Continuous-time dynamics of the Ising ferromagnet. Both panels compare the results of CME (thick solid lines) and iCME (dashed lines). In all cases, a completely magnetized system is placed in contact with a thermal bath at fixed temperature  $T$  and allowed to relax to equilibrium. **a)** Magnetization of the one-dimensional Ising ferromagnetic model. The points represent Glauber's exact solution for the average magnetization [70]. **b)** Magnetization of the Ising ferromagnetic model defined on a single instance of an Erdős-Rényi graph with size  $N = 5000$  and average connectivity  $\kappa = 3$ . The points represent the average of  $s = 10000$  Monte Carlo histories. The inset graph shows the local error  $\delta m(t) = (N^{-1} \sum_{i=1}^N (m_i^{DCAV}(t) - m_i^{MC}(t))^2)^{1/2}$  on a scale of order  $10^{-2}$ .

So far, integrating Eqs. (K.15) and (K.16) improves the results of the CME, at least in these ferromagnetic models. This is not necessarily true in other models. To continue the comparison, a model with richer phenomenology that includes a spin glass phase at low temperatures will be used. The Viana-Bray model defined on Erdős-Rényi graphs will be used (see subsection 4.2.1). The main difference from the ferromagnetic models is that now each pair of spins may have either a ferromagnetic or an antiferromagnetic interaction. This is chosen randomly, thus introducing another source of disorder into the system.

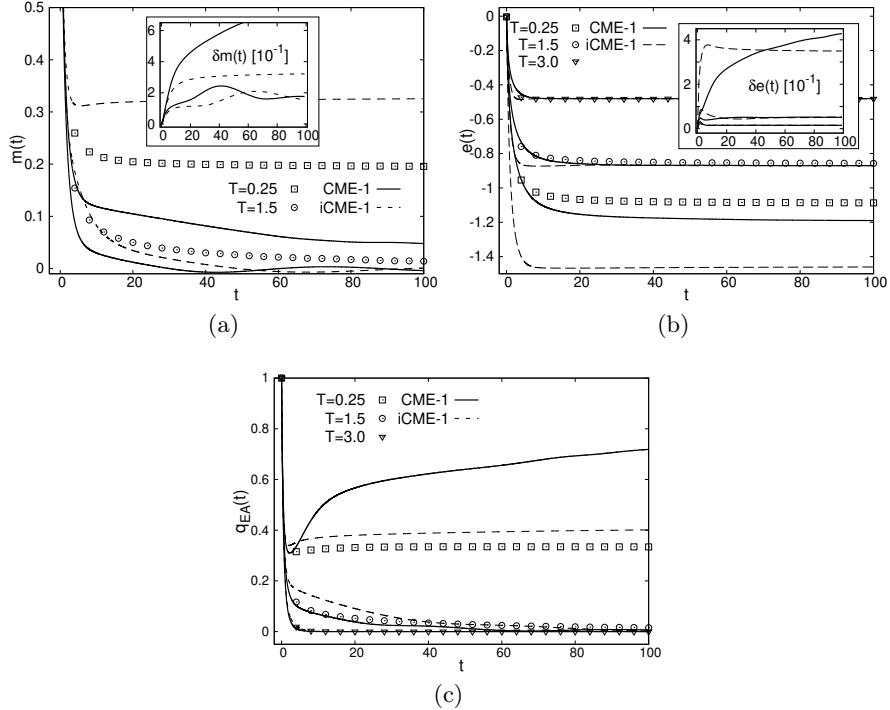


Figure K.2: Continuous-time dynamics of the Viana-Bray model for spin glass defined on an Erdős-Rényi graph of size  $N = 1000$  and average connectivity  $\kappa = 3$ . In all panels, the results of CME (thick solid lines), iCME (dashed lines), and Monte Carlo simulations (points) are compared. In each case, a completely magnetized system is placed in contact with a thermal bath at fixed temperature  $T$  and allowed to evolve over time. The points are the average of  $s = 10000$  Monte Carlo histories. **a)** Temporal evolution of magnetization. The inset graph shows the local error  $\delta m(t) = (N^{-1} \sum_{i=1}^N (m_i^{DCAV}(t) - m_i^{MC}(t))^2)^{1/2}$ . **b)** Temporal evolution of energy density. The inset graph shows the local error  $\delta e(t) = (N^{-1} c^{-1} \sum_{i \neq j} (e_{ij}^{DCAV}(t) - e_{ij}^{MC}(t))^2)^{1/2}$ . **c)** Temporal evolution of the Edwards-Anderson parameter  $q_{EA} = N^{-1} \sum_i (\sum_{\sigma_i} \sigma_i P(\sigma_i))^2$ .

Fig. (K.2) does not allow us to decide which of the two techniques is better in this case. The main panel of Fig. (K.2a) contains the temporal evolution of the average magnetization. At very low temperature ( $T = 0.25$ ), both CME and iCME are far from the Monte Carlo results. The inset graph, on the other hand, shows that the local errors  $\delta m(t) = (N^{-1} \sum_{i=1}^N (m_i^{DCAV}(t) - m_i^{MC}(t))^2)^{1/2}$  of iCME are somewhat smaller.

The temporal dependence of energy density is less straightforward to describe using average values. Apparently, at low temperatures, iCME provides a much worse

description of the simulations than CME (see Fig. (K.2b)). However, note that the corresponding local errors in energy density  $\delta e(t) = (N^{-1}c^{-1}\sum_{i\neq j}(e_{ij}^{DCAV}(t) - e_{ij}^{MC}(t))^2)^{1/2}$  are of the same order. Finally, Fig. (K.2c) shows that the temporal evolution of the Edwards-Anderson parameter  $q_{EA} = N^{-1}\sum_i(\sum_{\sigma_i}\sigma_i P(\sigma_i))^2$  is predicted more accurately by the new Eqs. (K.15) and (K.16).

In summary, this appendix presents a new technique to close the Master Equation for the dynamics of discrete variables in continuous time. Although the derivation is very similar to that of CME (see Section 2.3), the numerical results of the new iCME are better in some models. However, both theoretical methods fail to describe the dynamics in the spin glass phase of the Viana-Bray model.



## Appendix L

### Equivalence between the *CDA-1* and PBMF

The equations for the Pair-Based Mean Field (PBMF) approximation were shown in subsection 4.2.3 applied to the Susceptible-Infectious-Susceptible (SIS) model for epidemic propagation. Here, it will be shown that the main PBMF approximation is analogous to the one applied in the First Order Approximate Conditional Dynamics (*CDA-1*), introduced in Chapter 6. Numerical results of PBMF in the SIS will also be compared with those obtained using second-order (*CDA-2*). To complete the description, a figure is added that also contains results from the second level of approximation of the system presented in Chapter 4 (i.e., *CME-2*).

In the context of the SIS model, the time derivative of the pair probabilities  $P(\sigma_i, \sigma_j)$  is written in terms of the probabilities of groups of three variables connected in the graph  $P(\sigma_i, \sigma_j, \sigma_k)$ . To fix ideas, it will be assumed that  $i$  and  $k$  are neighbors to  $j$ . In order to close the system of equations,  $P(\sigma_i, \sigma_j, \sigma_k)$  needs to be expressed in terms of pair probabilities. Reference [78] proposed the approximation  $P(\sigma_i, \sigma_j, \sigma_k) \approx P(\sigma_i)P(\sigma_j, \sigma_k)$ . However, it was later shown [32] that results improved by applying the symmetric approximation:

$$P(\sigma_i, \sigma_j, \sigma_k) \approx \frac{P(\sigma_i, \sigma_j)P(\sigma_j, \sigma_k)}{P(\sigma_j)} \quad (\text{L.1})$$

It can be immediately seen that the approximation taken to reach *CDA-1*:

$$P(\sigma_{a \setminus j}, \sigma_j, \sigma_{b \setminus j}) \approx \frac{P(\sigma_a)P(\sigma_b)}{P(\sigma_j)} \quad (\text{L.2})$$

reduces to Eq. L.1 in the case of pair interactions. Indeed, factor nodes such as  $a$  end up having only two variable nodes each, so  $P(\sigma_a) \equiv P(\sigma_i, \sigma_j)$  and  $P(\sigma_{a \setminus j}) \equiv P(\sigma_i)$ . A similar situation occurs for factor node  $b$ .

This means that *CDA-1* is a more general formulation of PBMF, applicable to any model of discrete variables on random graphs with local dynamic rules  $r_j$ . It is

expected that the next level of approximation, *CDA-2*, will improve upon the well-known PBMF results.

Fig. (L.1a) compares the descriptions of both techniques for the SIS model and an epidemic spreading from a single node in a random regular graph. *CDA-2* is noticeably more accurate in its predictions of the average number of infected individuals, especially for low infection rates  $\lambda$ .

Building on the latest results from subsection 4.2.3, Fig. (L.1b) adds *CME-2* to the group of techniques considered. No significant differences are observed between the quality of the descriptions obtained with *CME-2* and those already shown for *CDA-2*.

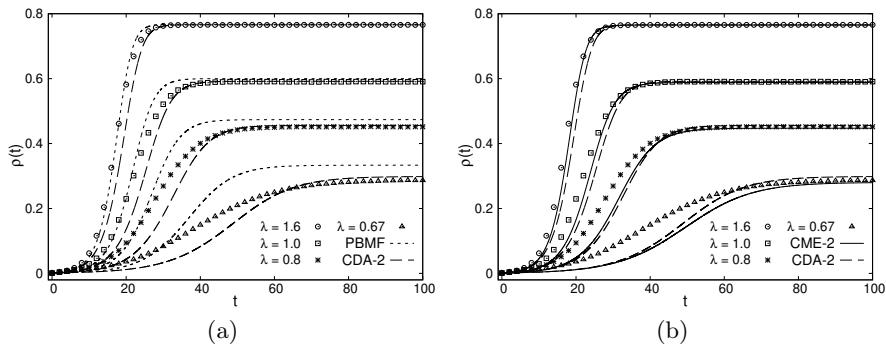


Figure L.1: Dynamics of the SIS model for epidemic propagation. The graphs show the temporal dependence of the average probability of finding an infectious individual. The epidemic starts from a single node in a random regular graph with connectivity  $c = 3$  and size  $N = 1000$ . The same value of the SIS model parameter  $\beta$  was used ( $\beta = 0.4$ ). Four different values of the parameter  $\mu = (0.25, 0.4, 0.5, 0.6)$  were chosen. The corresponding infection rates  $\lambda = \beta/\mu$  are  $\lambda = (1.6, 1.0, 0.8, 0.67)$ . Each point is the average of  $s = 10000$  Monte Carlo simulations. Results for *CDA-2* are shown with thick dashed lines. **a)** PBMF results are included with thin dashed lines. **b)** *CME-2* results are included with thick solid lines.

## Appendix M

# Average case for the Conditional Dynamic Approximation

In this appendix, average equations for the *CDA-2* are written for the dynamics in a regular random hypergraph with homogeneous initial conditions:  $P(\sigma_i) = \frac{1}{2}$  for all  $i = 1, \dots, N$ . It will be demonstrated that the version of *CDA-2* for a given graph, presented in Chapter 6, has as a particular case the Dynamic Independent Neighbors Approximation (DINA) [21, 22] referred to in subsection 4.2.2 and in Section 5.4.

Below is the *CDA-2* as it appears in Section 6.1:

$$\begin{aligned} \frac{d}{dt} P(\sigma_i, \sigma_{\partial i}) &= -r_i(\sigma_i, \sigma_{\partial i}) P(\sigma_i, \sigma_{\partial i}) + r_i(-\sigma_i, \sigma_{\partial i}) P(-\sigma_i, \sigma_{\partial i}) - \\ &\quad - \sum_{b \subset \partial i} \sum_{j \in b \setminus i} \sum_{\sigma_{\partial j \setminus b}} r_j(\sigma_j, \sigma_{\partial j}) \frac{P(\sigma_j, \sigma_{\partial j})}{\sum_{\sigma_{\partial j \setminus b}} P(\sigma_j, \sigma_{\partial j})} P(\sigma_i, \sigma_{\partial i}) + \\ &\quad + \sum_{b \subset \partial i} \sum_{j \in b \setminus i} \sum_{\sigma_{\partial j \setminus b}} r_j(-\sigma_j, \sigma_{\partial j}) \frac{P(\sigma_j, \sigma_{\partial j})}{\sum_{\sigma_{\partial j \setminus b}} P(-\sigma_j, \sigma_{\partial j})} P(\sigma_i, F_j[\sigma_{\partial i}]) \quad (\text{M.1}) \end{aligned}$$

This equation simplifies considerably when the  $N$  spins interact on a regular random hypergraph where all nodes have the same number  $c$  of neighbors. If the initial conditions are the same for all sites, all local probabilities are governed by the same equation. Additionally, if the system is initially unmagnetized ( $m(0) = 0$ ), it will remain so for all time  $t$ . It is then possible to express everything in terms of a single important parameter related to the energy: the number  $u = \sum_{a \subset \partial i} \delta(\prod_{k \in a} \sigma_k, -1)$  of unsatisfied interactions between  $\sigma_i$  and its neighbors. Here  $\delta(x, y)$  is the Kronecker delta. Thus, equation (M.1) can be rewritten as:

$$\begin{aligned}
 \frac{d}{dt}P(u) = & -r(u)P(u) + r(c-u)P(c-u) - \\
 & -(p-1)u\left[\sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}P(\hat{u}+1)\right]^{-1}\sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}r(\hat{u}+1)P(\hat{u}+1)P(u) \\
 & +(p-1)u\left[\sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}P(\hat{u})\right]^{-1}\sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}r(\hat{u})P(\hat{u})P(u-1) \\
 & -(p-1)(c-u)\left[\sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}P(\hat{u})\right]^{-1}\sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}r(\hat{u})P(\hat{u})P(u) \\
 & +(p-1)(c-u)\left[\sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}P(\hat{u}+1)\right]^{-1}\sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}r(\hat{u}+1)P(\hat{u}+1)P(u+1)
 \end{aligned} \tag{M.2}$$

where it has been assumed that the dynamic rules  $r_i$  depend only on the number of unsatisfied interactions  $u$  and the connectivity  $c$ .

For simplicity, let us introduce the probabilities

$$\hat{P}(u) = \binom{c}{u}P(u) \tag{M.3}$$

Equation (M.2) can be expressed entirely in terms of  $\hat{P}$  by simply rearranging the terms. The sums over  $\hat{u}$ , ranging from  $\hat{u} = 0$  to  $\hat{u} = c-1$ , should be modified as in the following example:

$$\begin{aligned}
 \sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}P(\hat{u}+1) &= \sum_{\hat{u}=0}^{c-1}\frac{(c-1)!}{\hat{u}!(c-1-\hat{u})!}P(\hat{u}+1) \\
 \sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}P(\hat{u}+1) &= \sum_{\hat{u}=0}^{c-1}\frac{(c-1)!}{\hat{u}!(c-1-\hat{u})!}\frac{(\hat{u}+1)!(c-\hat{u}-1)!}{c!}\hat{P}(\hat{u}+1) \\
 \sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}P(\hat{u}+1) &= \frac{1}{c}\sum_{\hat{u}=0}^{c-1}(\hat{u}+1)\hat{P}(\hat{u}+1) = \frac{1}{c}\sum_{u=0}^cu\hat{P}(u) \\
 \sum_{\hat{u}=0}^{c-1}\binom{c-1}{\hat{u}}P(\hat{u}+1) &= \frac{1}{c}\langle u \rangle
 \end{aligned} \tag{M.4}$$

The definition of  $\hat{P}$  was used in the second line of Equation (M.4) to express  $P(\hat{u}+1)$  in terms of the average probability  $\hat{P}(\hat{u}+1)$ . In the last line, the notation  $\langle \cdot \rangle \equiv \sum_{u=0}^c [\cdot] \hat{P}(u)$  was introduced for averages over the distribution  $\hat{P}(u)$ .

The following identities can be analogously deduced:

$$\sum_{\hat{u}=0}^{c-1} \binom{c-1}{\hat{u}} P(\hat{u}) = \frac{1}{c} \langle c - u \rangle \quad (\text{M.5})$$

$$\sum_{\hat{u}=0}^{c-1} \binom{c-1}{\hat{u}} r(\hat{u}+1) P(\hat{u}+1) = \frac{1}{c} \langle u r(u) \rangle \quad (\text{M.6})$$

$$\sum_{\hat{u}=0}^{c-1} \binom{c-1}{\hat{u}} r(\hat{u}) P(\hat{u}) = \frac{1}{c} \langle (c-u) r(u) \rangle \quad (\text{M.7})$$

Inserting Equations (M.4), (M.5), (M.6), and (M.7) into Equation (M.2) yields the differential equation:

$$\begin{aligned} \frac{d}{dt} P(u) &= -r(u) P(u) + r(c-u) P(c-u) - \\ &\quad -(p-1) \frac{\langle u r(u) \rangle}{\langle u \rangle} [u P(u) - (c-u) P(u+1)] \\ &\quad -(p-1) \frac{\langle (c-u) r(u) \rangle}{\langle c-u \rangle} [(c-u) P(u) - u P(u-1)] \end{aligned} \quad (\text{M.8})$$

The final average equations are obtained by multiplying Equation (M.8) by  $\binom{c}{u}$ , using the definition (M.3) of  $\hat{P}(u)$  and the relations:

$$\binom{c}{u} (c-u) P(u+1) = (u+1) \hat{P}(u+1) \quad (\text{M.9})$$

$$\binom{c}{u} u P(u-1) = (c-u+1) \hat{P}(u-1) \quad (\text{M.10})$$

The result is:

$$\begin{aligned} \frac{d}{dt} \hat{P}(u) &= -r(u) \hat{P}(u) + r(c-u) \hat{P}(c-u) - \\ &\quad -(p-1) \frac{\langle u r(u) \rangle}{\langle u \rangle} [u \hat{P}(u) - (u+1) \hat{P}(u+1)] \\ &\quad -(p-1) \frac{\langle (c-u) r(u) \rangle}{\langle c-u \rangle} [(c-u) \hat{P}(u) - (c-u+1) \hat{P}(u-1)] \end{aligned} \quad (\text{M.11})$$

These equations are a simplified version of the equation for a given graph (*CDA-2*) in the p-spin model. On the other hand, the DINA equations presented in Ref. [22] are written for pairwise spin interactions in a random graph. However, it is possible to recognize some similarities. In fact, one can take the version of the *CDA-2* for pairwise interactions:

$$\begin{aligned}
\frac{d}{dt} P(\sigma_i, \sigma_{\partial i}) = & -r_i(\sigma_i, \sigma_{\partial i}) P(\sigma_i, \sigma_{\partial i}) + r_i(-\sigma_i, \sigma_{\partial i}) P(-\sigma_i, \sigma_{\partial i}) - \\
& - \sum_{j \in \partial i} \sum_{\sigma_{\partial j \setminus i}} r_j(\sigma_j, \sigma_{\partial j}) \frac{P(\sigma_j, \sigma_{\partial j})}{\sum_{\sigma_{\partial j \setminus i}} P(\sigma_j, \sigma_{\partial j})} P(\sigma_i, \sigma_{\partial i}) + \\
& + \sum_{j \in \partial i} \sum_{\sigma_{\partial j \setminus i}} r_j(-\sigma_j, \sigma_{\partial j}) \frac{P(\sigma_j, \sigma_{\partial j})}{\sum_{\sigma_{\partial j \setminus i}} P(-\sigma_j, \sigma_{\partial j})} P(\sigma_i, F_j[\sigma_{\partial i}]) \quad (\text{M.12})
\end{aligned}$$

and then follow a similar path to average equations. The only difference to consider in order to obtain the DINA is that if the initial condition is not  $P(\sigma_i) = 1/2$ , the magnetization will not be zero in the dynamics. This introduces a second relevant parameter: the value of the central spin  $\sigma_i \equiv \sigma$ . Aside from this detail, the procedure is analogous and one obtains:

$$\begin{aligned}
\frac{d}{dt} \hat{P}(\sigma, u) = & -r(u) \hat{P}(\sigma, u) + r(c-u) \hat{P}(-\sigma, c-u) - \\
& - \frac{\langle u r(u) \rangle_{-\sigma}}{\langle u \rangle_{-\sigma}} [u \hat{P}(\sigma, u) - (u+1) \hat{P}(\sigma, u+1)] \\
& - \frac{\langle (c-u) r(u) \rangle_\sigma}{\langle c-u \rangle_\sigma} [(c-u) \hat{P}(\sigma, u) - (c-u+1) \hat{P}(\sigma, u-1)] \quad (\text{M.13})
\end{aligned}$$

where  $\langle \cdot \rangle_\sigma \equiv \sum_{u=0}^c [\cdot] P(\sigma, u)$ .

Thus, the DINA as presented in [22] has been rederived. One can think of the *CDA-2* as a general case applicable to a given graph of the DINA interactions. A version that can be easily adapted to other models and graph architectures.

## Appendix N

# Cavity Master Equation for the spin glass dynamics

In Chapter 6, a methodology is presented to study the dynamics of spin glasses. The limitations of the techniques presented in Chapter 4 and in Appendix K in this case necessitated the development of new equations.

This appendix contains the adaptation of *CME-2* (Chapter 4) and the formulation of *iCME-2* (Appendix K) for the ferromagnetic p-spin model defined on regular random hypergraphs. Results are shown for temperatures  $T$  below the transition temperature  $T_d$  to the spin glass phase, and it is verified that *CDA-2* (Chapter 6) is a significantly more accurate alternative.

*CME-2* is easily generalizable to hypergraphs, yielding:

$$\begin{aligned} \frac{d}{dt} p(\sigma_{\partial i \setminus a}, \sigma_i \| \sigma_{a \setminus i}) = & -r_i(\sigma_i, \sigma_{\partial i}) p(\sigma_{\partial i \setminus a}, \sigma_i \| \sigma_{a \setminus i}) + \\ & + r_i(-\sigma_i, \sigma_{\partial i}) p(\sigma_{\partial i \setminus a}, -\sigma_i \| \sigma_{a \setminus i}) - \\ & - \sum_{b \subset \partial i \setminus a} \sum_{k \in b \setminus i} \sum_{\sigma_{\partial k \setminus b}} r_k(\sigma_k, \sigma_{\partial k}) p(\sigma_{\partial k \setminus b}, \sigma_k \| \sigma_{b \setminus k}) \frac{p(\sigma_{\partial i \setminus a}, \sigma_i \| \sigma_{a \setminus i})}{p(\sigma_k \| \sigma_{b \setminus k})} + \\ & + \sum_{b \subset \partial i \setminus a} \sum_{k \in b \setminus i} \sum_{\sigma_{\partial k \setminus b}} r_k(-\sigma_k, \sigma_{\partial k}) p(\sigma_{\partial k \setminus b}, -\sigma_k \| \sigma_{b \setminus k}) \frac{p(F_k[\sigma_{\partial i \setminus a}], \sigma_i \| \sigma_{a \setminus i})}{p(-\sigma_k \| \sigma_{b \setminus k})} \end{aligned} \quad (\text{N.1})$$

where  $r_i(\sigma_i, \sigma_{\partial i})$  is the probability per unit time that the spin  $\sigma_i$  changes its value given the configuration of its neighbors and its state at time  $t$ . The cavity conditional probability  $p(\sigma_{\partial i \setminus a}, \sigma_i \| \sigma_{a \setminus i})$  is a fundamental auxiliary quantity for obtaining a closure of the Master Equation.

Similarly to what was done in Appendix M, when all sites have the same connectivity  $c$  and the initial conditions are  $P(\sigma_i) = 1/2$  for all  $i = 1, \dots, N$ , a single equation for the cavity probabilities can be written:

$$\begin{aligned} \frac{d}{dt} \hat{p}(\tilde{u} \parallel \Omega) = & -r(\tilde{u} + \delta_{\Omega,U}) \hat{p}(\tilde{u} \parallel \Omega) + \\ & + r(c - \tilde{u} - \delta_{\Omega,U}) \hat{p}(c - 1 - \tilde{u} \parallel S\delta_{\Omega,U} + U\delta_{\Omega,S}) - \\ & (p - 1) \langle r(\tilde{u}) \rangle_U [\tilde{u} \hat{p}(\tilde{u} \parallel \Omega) - (\tilde{u} + 1) \hat{p}(\tilde{u} + 1 \parallel \Omega)] \\ & (p - 1) \langle r(\tilde{u}) \rangle_S [(c - 1 - \tilde{u}) \hat{p}(\tilde{u} \parallel \Omega) - (c - \tilde{u}) \hat{p}(\tilde{u} - 1 \parallel \Omega)] \end{aligned} \quad (N.2)$$

where  $\tilde{u} = 1, \dots, c - 1$  is defined in this case as the number of unsatisfied interactions  $\tilde{u} = \sum_{b \in \partial i \setminus a} \delta(\prod_{k \in b} \sigma_k, -1)$  between the central spin  $\sigma_i$  and all its neighbors, except those in the plaquette  $a$ . Here  $\delta(x, y) \equiv \delta_{x,y}$  is simply the well-known Kronecker delta. In the cavity conditional probabilities  $p(\sigma_{\partial i \setminus a}, \sigma_i \parallel \sigma_{a \setminus i})$ , the condition  $\sigma_{a \setminus i}$  has been replaced by the variable  $\Omega = U, S$ . When  $\prod_{j \in a} \sigma_j = 1$  then  $\Omega = S$  and when  $\prod_{j \in a} \sigma_j = -1$  then  $\Omega = U$ . Finally, the averages are introduced

$$\langle r(\tilde{u}) \rangle_\Omega \equiv \frac{\sum_{\tilde{u}=0}^{c-1} r(\tilde{u} + \delta_{\theta,U}) \hat{p}(\tilde{u} \parallel \Omega)}{\sum_{\tilde{u}=0}^{c-1} \hat{p}(\tilde{u} \parallel \Omega)} \quad (N.3)$$

These equations are complemented by an average equation for the local probability  $\hat{P}(u)$  which has a similar form. The numerical results of integrating the *CME-2* at low temperatures in the ferromagnetic p-spin model are included in Fig. (N.1a).

On the other hand, the *iCME-2* is written for the probabilities  $p(\sigma_{\partial i \setminus a}, \sigma_i, \sigma_{a \setminus i})$  when the nodes  $a \setminus i$  are *end nodes*. This means using an auxiliary graph where they have only one neighbor: site  $i$ . The equations are:

$$\begin{aligned} \frac{d}{dt} p(\sigma_{\partial i \setminus a}, \sigma_i, \sigma_{a \setminus i}) = & -r_i(\sigma_i, \sigma_{a \setminus i}, \sigma_{\partial i \setminus a}) p(\sigma_{\partial i \setminus a}, \sigma_i, \sigma_{a \setminus i}) + \\ & + r_i(-\sigma_i, \sigma_{a \setminus i}, \sigma_{\partial i \setminus a}) p(\sigma_{\partial i \setminus a}, -\sigma_i, \sigma_{a \setminus i}) - \\ & - \sum_{b \in \partial i \setminus a} \sum_{k \in b \setminus i} \sum_{\sigma_{\partial k \setminus b}} r_k(\sigma_k, \sigma_{b \setminus k}, \sigma_{\partial k \setminus b}) p(\sigma_{\partial k \setminus b}, \sigma_k, \sigma_{b \setminus k}) \frac{p(\sigma_{\partial i \setminus a}, \sigma_i, \sigma_{a \setminus i})}{p(\sigma_k, \sigma_{b \setminus k})} + \\ & + \sum_{b \in \partial i \setminus a} \sum_{k \in b \setminus i} \sum_{\sigma_{\partial k \setminus b}} r_k(-\sigma_k, \sigma_{b \setminus k}, \sigma_{\partial k \setminus b}) p(\sigma_{\partial k \setminus b}, -\sigma_k, \sigma_{b \setminus k}) \frac{p(F_k[\sigma_{\partial i \setminus a}], \sigma_i, \sigma_{a \setminus i})}{p(-\sigma_k, \sigma_{b \setminus k})} \end{aligned} \quad (N.4)$$

Its average version is also easily obtained for regular random hypergraphs with homogeneous initial conditions:

$$\begin{aligned} \frac{d}{dt} \hat{p}(\tilde{u}, \Omega) = & -r(\tilde{u} + \delta_{\Omega,U}) \hat{p}(\tilde{u}, \Omega) + \\ & + r(c - \tilde{u} - \delta_{\Omega,U}) \hat{p}(c - 1 - \tilde{u}, S\delta_{\Omega,U} + U\delta_{\Omega,S}) - \\ & (p - 1) \langle\langle r(\tilde{u}) \rangle\rangle_U [\tilde{u} \hat{p}(\tilde{u}, \Omega) - (\tilde{u} + 1) \hat{p}(\tilde{u} + 1, \Omega)] \\ & (p - 1) \langle\langle r(\tilde{u}) \rangle\rangle_S [(c - 1 - \tilde{u}) \hat{p}(\tilde{u}, \Omega) - (c - \tilde{u}) \hat{p}(\tilde{u} - 1, \Omega)] \end{aligned} \quad (N.5)$$

where

$$\langle\langle r(\tilde{u}) \rangle\rangle_{\Omega} \equiv \frac{\sum_{\tilde{u}=0}^{c-1} r(\tilde{u} + \delta_{\theta,U}) \hat{p}(\tilde{u}, \Omega)}{\sum_{\tilde{u}=0}^{c-1} \hat{p}(\tilde{u}, \Omega)} \quad (\text{N.6})$$

These equations can also be complemented by an average equation for  $\hat{P}(u)$ . Its numerical integration is feasible without significant computational costs.

Fig. (N.1a) shows the results of the *iCME-2* for spin glass dynamics in the ferromagnetic p-spin model. The predicted time scales for the first relaxation are far from those observed in Monte Carlo simulations. The plateau energy value obtained with the *iCME-2* is well above the correct one.

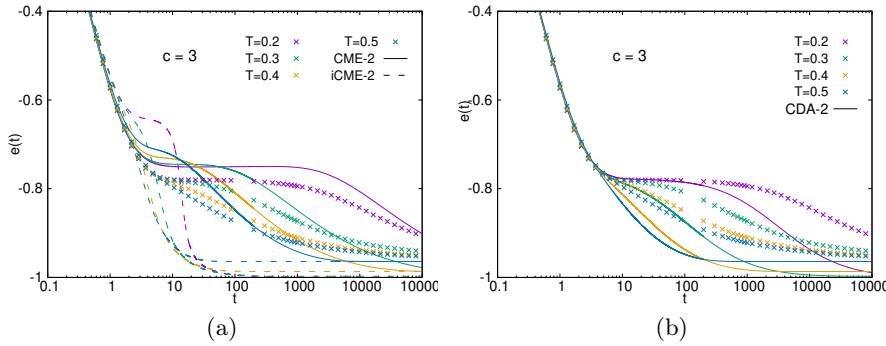


Figure N.1: Temporal dependence of the energy density  $e(t)$  for the ferromagnetic p-spin model with  $p = 3$  defined on regular random hypergraphs with connectivity  $c = 3$ . Each point is an average of  $10^4$  Monte Carlo simulations with system size  $N = 10^4$ . In each case, the system starts in a random configuration and is brought into contact with a thermal bath at a fixed temperature  $T$ . **a)** The results of the numerical integration of *CME-2* (solid lines) and *iCME-2* (dashed lines) in their average versions (see Eqs. (N.2) and (N.5)). **b)** Reference is made to the *CDA-2* predictions already presented in Chapter 6.

The *CME-2* also does not provide a good description of the simulations. Although the time scales are closer to the real ones, the predicted value for the plateau energy still overestimates what is obtained with Monte Carlo. For reference, Fig. (N.1b) presents the *CDA-2* results from Section 6.1. It is clear that the *CDA-2* represents a significant improvement over *CME-2* and *iCME-2*. This is why this technique was chosen for the entire study in Chapter 6.



## Appendix O

# The average CME-2 in K-SAT

In this appendix, the average formulations of the *CME-2* for K-SAT will be derived:

$$\begin{aligned} \frac{d}{dt} p(\sigma_{\partial i \setminus a}, \sigma_i \| \sigma_{a \setminus i}) &= -r_i(\sigma_i, \sigma_{\partial i}) p(\sigma_{\partial i \setminus a}, \sigma_i \| \sigma_{a \setminus i}) + \\ &\quad + r_i(-\sigma_i, \sigma_{\partial i}) p(\sigma_{\partial i \setminus a}, -\sigma_i \| \sigma_{a \setminus i}) - \\ &\quad - \sum_{b \subset \partial i \setminus a} \sum_{k \in b \setminus i} \sum_{\sigma_{\partial k \setminus b}} r_k(\sigma_k, \sigma_{\partial k}) p(\sigma_{\partial k \setminus b}, \sigma_k \| \sigma_{b \setminus k}) \frac{p(\sigma_{\partial i \setminus a}, \sigma_i \| \sigma_{a \setminus i})}{p(\sigma_k \| \sigma_{b \setminus k})} + \\ &\quad + \sum_{b \subset \partial i \setminus a} \sum_{k \in b \setminus i} \sum_{\sigma_{\partial k \setminus b}} r_k(-\sigma_k, \sigma_{\partial k}) p(\sigma_{\partial k \setminus b}, -\sigma_k \| \sigma_{b \setminus k}) \frac{p(F_k[\sigma_{\partial i \setminus a}], \sigma_i \| \sigma_{a \setminus i})}{p(-\sigma_k \| \sigma_{b \setminus k})} \end{aligned} \quad (\text{O.1})$$

where  $r_i(\sigma_i, \sigma_{\partial i})$  is the rate of flipping the variable  $\sigma_i$  given a specific configuration of its neighbors.

For most K-SAT algorithms, it is important to know at each step how many unsatisfied clauses contain a given variable. Recall that the value of a clause  $a$  can be written as:

$$E_a = \prod_{k \in a} \frac{1 - J_k^a \sigma_k}{2} \quad (\text{O.2})$$

where the parameters  $J_k^a = \pm 1$  are random couplings of the variables  $\sigma_k$  with the clauses.

In the following, it will be assumed that the rules  $r_i$  for the dynamics depend only on the number of unsatisfied clauses to which  $\sigma_i$  belongs and on the number of clauses that will become unsatisfied by flipping this variable. To parameterize the local probability corresponding to the variables  $\sigma_{\partial i \setminus a}$ ,  $\sigma_i$ , and the corresponding couplings, a symbol will be assigned to each possible configuration of a clause containing  $\sigma_i$  and  $K - 1$  other variables. This clause will be denoted by the letter  $b$ .

Two important features of the clause will be taken into account: i) the value of the product  $\sigma_i J_i^b$ , which determines whether the variable  $\sigma_i$  satisfies its coupling  $J_i^b$  or

not, and ii) the number  $z$  of satisfied couplings by the remaining  $K - 1$  variables:  $z = \sum_{k \in b \setminus i} (1 - J_k^b \sigma_k)/2$ . Thus, a particular configuration of the clause can be represented with the symbol  $u_z^{(l)}$  if:

- the variable  $\sigma_i$  does not satisfy its coupling:  $(1 - J_i^b \sigma_i)/2 = 1$
- $\sum_{k \in b \setminus i} (1 - J_k^b \sigma_k)/2 = z$
- $l$  is an index representing each of the  $\binom{K-1}{z}$  different configurations that have the same  $z$ . That is,  $l = 1, \dots, \binom{K-1}{z}$ . If  $z = 0$  or  $z = K - 1$ , the value of  $l$  will be omitted as there is only one configuration for these  $z$ .

Similarly, the symbol  $s_z^{(l)}$  is used when the variable  $\sigma_i$  satisfies the clause:  $(1 - J_i^b \sigma_i)/2 = 0$ . The numbers  $z$  and  $l$  have the same interpretations.

The notation for the conditional probability  $p(\sigma_{\partial i \setminus a}, \sigma_i \mid \sigma_{a \setminus i})$  will also be simplified by introducing the symbol  $\Omega$  to capture the relevant configurations of clause  $a$ . There are three important cases:

- $a$  is unsatisfied and is assigned  $\Omega = U$ .
- $a$  will become unsatisfied when the value of  $\sigma_i$  is flipped. This means that  $2^{-K} (1 + J_i^a \sigma_i) \prod_{k \in a \setminus i} (1 - J_k^a \sigma_k) = 1$ . In this case  $\Omega = S_0$
- $a$  is satisfied and at least one of the other  $K - 1$  variables satisfies its coupling with  $a$ . That is, even if  $\sigma_i$  is flipped, the clause remains satisfied. In this case  $\Omega = S$ .

Now the probability  $p(\sigma_{\partial i \setminus a}, \sigma_i \mid \sigma_{a \setminus i})$  can be rewritten as:

$$p(\sigma_{\partial i \setminus a}, \sigma_i \mid \sigma_{a \setminus i}) = p(u_0, \vec{u}_1, \dots, \vec{u}_{K-2}, u_{K-1}, s_0, \vec{s}_1, \dots, \vec{s}_{K-2}, s_{K-1} \mid \Omega) \quad (O.3)$$

where  $\vec{u}_z$  and  $\vec{s}_z$  are vectors with components  $u_z^{(l)}$  and  $s_z^{(l)}$ ,  $l = 1, \dots, \binom{K-1}{z}$ .

The new notation can be somewhat misleading. It should be noted that:

$$p_\gamma(\Omega) = \sum_{u_0} \dots \sum_{u_{K-1}} \sum_{s_0} \dots \sum_{s_{K-1}} p(u_0, \vec{u}_1, \dots, \vec{u}_{K-2}, u_{K-1}, s_0, \vec{s}_1, \dots, \vec{s}_{K-2}, s_{K-1} \mid \Omega) \neq 1 \quad (O.4)$$

This sum is a marginal of  $p(\sigma_{\partial i \setminus a}, \sigma_i \mid \sigma_{a \setminus i})$  which corresponds to the local probability  $p(\sigma_i \mid \sigma_{a \setminus i})$ , chosen so that  $\sigma_i$  and  $\sigma_{a \setminus i}$  are compatible with the value of  $\Omega$ . If equation (O.3) is substituted into equation (N.1) and summed over all configurations with the same values of  $u_0$  and  $s_0$ , one obtains:

$$\begin{aligned}
\frac{d}{dt} p_\gamma(u_0, s_0 \parallel \Omega) = & -r(u_0 + \delta_{\Omega, U_0}, s_0 + \delta_{\Omega, S_0}) p_\gamma(u_0, s_0 \parallel \Omega) + \\
& + r(s_0 + \delta_{\Omega, S_0}, u_0 + \delta_{\Omega, U_0}) p_\gamma(s_0, u_0 \parallel \mathcal{I}[\Omega]) - \\
& - (K-1) \langle r \rangle_U \left[ u_0 p_\gamma(u_0, s_0 \parallel \Omega) - (\gamma - u_0 - s_0) p_\gamma(u_0 + 1, s_0 \parallel \Omega) \right] - \\
& - \frac{(K-1)}{2^K - 2} \langle r \rangle_{S_0} \left[ (\gamma - u_0 - s_0) p_\gamma(u_0, s_0 \parallel \Omega) - u_0 p_\gamma(u_0 - 1, s_0 \parallel \Omega) \right] \\
& - (K-1) \langle r \rangle_S \left[ s_0 p_\gamma(u_0, s_0 \parallel \Omega) - (\gamma - u_0 - s_0) p_\gamma(u_0, s_0 + 1 \parallel \Omega) \right] - \\
& - \frac{(K-1)}{2^K - 2} \langle r \rangle_S \left[ (\gamma - u_0 - s_0) p_\gamma(u_0, s_0 \parallel \Omega) - s_0 p_\gamma(u_0, s_0 - 1 \parallel \Omega) \right]
\end{aligned} \tag{O.5}$$

where  $\gamma$  is related to the connectivity of the variable node ( $\gamma = c-1$ ) and the operator  $\mathcal{I}[\Omega] = U_0 \delta_{\Omega, S_0} + S_0 \delta_{\Omega, U_0} + S \delta_{\Omega, S}$  was introduced. The averages were also defined:

$$\langle r \rangle_\Omega = \sum_{\gamma'=0}^{\infty} \sum_{u'_0=0}^{\gamma'} \sum_{s'_0=0}^{\gamma'-u'_0} q(\gamma', u'_0, s'_0) r(u'_0 + \delta_{\Omega, U}, s'_0 + \delta_{\Omega, S_0}) \frac{p_{\gamma'}(u'_0, s'_0 \parallel \Omega)}{p_{\gamma'}(\Omega)} \tag{O.6}$$

Following equation (O.4), the probability  $p_{\gamma'}(\Omega)$  can also be written in terms of  $p_{\gamma'}(u'_0, s'_0 \parallel \Omega)$  as

$$p_{\gamma'}(\Omega) = \sum_{u''_0=0}^{\gamma'} \sum_{s''_0=0}^{\gamma'-u''_0} \frac{\gamma'!}{(u''_0)!(s''_0)!(\gamma' - u''_0 - s''_0)!} p_{\gamma'}(u''_0, s''_0 \parallel \Omega) \tag{O.7}$$

In Erdös-Rényi hypergraphs, the distribution  $q(\gamma', u'_0, s'_0)$  is simply a Poisson distribution:

$$q(\gamma', u'_0, s'_0) = e^{-\alpha K} \frac{(\alpha K)^{\gamma'!}}{\gamma'!} \frac{\gamma'!}{(u'_0)!(s'_0)!(\gamma' - u'_0 - s'_0)!} \tag{O.8}$$

The complexity of the numerical integration of equation (O.5) can be further reduced by defining the average probabilities  $\hat{p}_\gamma(u_0, s_0 \parallel \Omega) = q(\gamma, u_0, s_0) p_\gamma(u_0, s_0 \parallel \Omega)$ . One then has:

$$\begin{aligned}
\frac{d}{dt} \hat{p}_\gamma(u_0, s_0 \parallel \Omega) = & -r(u_0 + \delta_{\Omega, U_0}, s_0 + \delta_{\Omega, S_0}) \hat{p}_\gamma(u_0, s_0 \parallel \Omega) + \\
& + r(s_0 + \delta_{\Omega, S_0}, u_0 + \delta_{\Omega, U_0}) \hat{p}_\gamma(s_0, u_0 \parallel \mathcal{I}[\Omega]) - \\
& - (K-1) \langle r \rangle_U \left\{ u_0 \hat{p}_\gamma(u_0, s_0 \parallel \Omega) - (u_0 + 1) \hat{p}_\gamma(u_0 + 1, s_0 \parallel \Omega) \right\} - \\
& - \frac{(K-1)}{2^K - 2} \langle r \rangle_{S_0} \left\{ (\gamma - u_0 - s_0) \hat{p}_\gamma(u_0, s_0 \parallel \Omega) - (\gamma - u_0 - s_0 + 1) \hat{p}_\gamma(u_0 - 1, s_0 \parallel \Omega) \right\} \\
& - (K-1) \langle r \rangle_S \left\{ s_0 \hat{p}_\gamma(u_0, s_0 \parallel \Omega) - (s_0 + 1) \hat{p}_\gamma(u_0, s_0 + 1 \parallel \Omega) \right\} - \\
& - \frac{(K-1)}{2^K - 2} \langle r \rangle_S \left\{ (\gamma - u_0 - s_0) \hat{p}_\gamma(u_0, s_0 \parallel \Omega) - (\gamma - u_0 - s_0 + 1) \hat{p}_\gamma(u_0, s_0 - 1 \parallel \Omega) \right\}
\end{aligned} \tag{O.9}$$

Most importantly, the averages in equation (O.6) can also be expressed in terms of  $\hat{p}$ , simplifying the calculations:

$$\langle r \rangle_{\Omega} = \sum_{\gamma'=0}^{\infty} \sum_{u'_0=0}^{\gamma'} \sum_{s'_0=0}^{\gamma'-u'_0} e^{-\alpha K} \frac{(\alpha K)^{\gamma'}}{\gamma'!} r(u'_0 + \delta_{\Omega,U}, s'_0 + \delta_{\Omega,S_0}) \frac{\hat{p}_{\gamma'}(u'_0, s'_0 \| \Omega)}{\hat{p}_{\gamma'}(\Omega)} \quad (\text{O.10})$$

With this, it is straightforward to find the differential equations for the joint probabilities:

$$\begin{aligned} \frac{d\hat{P}_c(u_0, s_0)}{dt} &= -r(u_0, s_0) \hat{P}_c(u_0, s_0) + r(s_0, u_0) \hat{P}_c(s_0, u_0) \\ &\quad - (K-1) \langle r(u'_0 + 1, s'_0) \rangle_{U_0} \left\{ u_0 \hat{P}_c(u_0, s_0) - (u_0 + 1) \hat{P}_c(u_0 + 1, s_0) \right\} - \\ &\quad - \frac{(K-1)}{2^K - 1} \langle r(u'_0, s'_0 + 1) \rangle_{S_0} \left\{ (c - u_0 - s_0) \hat{P}_c(u_0, s_0) - (c - u_0 - s_0 + 1) \hat{P}_c(u_0 - 1, s_0) \right\} \\ &\quad - (K-1) \langle r(u'_0, s'_0) \rangle_S \left\{ s_0 \hat{P}_c(u_0, s_0) - (s_0 + 1) \hat{P}_c(u_0, s_0 + 1) \right\} - \\ &\quad - \frac{(K-1)}{2^K - 1} \langle r(u'_0, s'_0) \rangle_S \left\{ (c - u_0 - s_0) \hat{P}_c(u_0, s_0) - (c - u_0 - s_0 + 1) \hat{P}_c(u_0, s_0 - 1) \right\} \end{aligned} \quad (\text{O.11})$$

where  $c$  is the connectivity of the variable node and the numbers  $u_0, s_0$  can take any integer value between 0 and  $c$ .

The algorithms used in K-SAT generally initialize the variables with random values such that  $P(\sigma_i) = 1/2$ . In this case, Poisson distributions are simply chosen for the joint probabilities:

$$\hat{p}_{\gamma}(\tilde{u}_0, \tilde{s}_0 \| \Omega) \Big|_{t=0} = \frac{1}{2} \frac{e^{\alpha K} (\alpha K)^{\gamma}}{(u_0)!(s_0)!(\gamma' - u_0 - s_0)!} \left( \frac{1}{2^K} \right)^{u_0+s_0} \left( 1 - \frac{1}{2^{K-1}} \right)^{\gamma' - u_0 - s_0} \quad (\text{O.12})$$

$$\hat{P}_c(u_0, s_0) \Big|_{t=0} = \frac{e^{\alpha K} (\alpha K)^c}{(u_0)!(s_0)!(c - u_0 - s_0)!} \left( \frac{1}{2^K} \right)^{u_0+s_0} \left( 1 - \frac{1}{2^{K-1}} \right)^{c - u_0 - s_0} \quad (\text{O.13})$$

Everything is then ready to numerically integrate the average *CME-2* (Eqs. (O.9) and (O.11)). As in the deduction of the average *CDA-2* from Section 7.1, the key distinction here is which conditional probability is used to compute the averages (O.10). In the case of *CME-2*, the cavity probabilities  $p_{\gamma'}(u'_0, s'_0 \| \Omega)/p_{\gamma'}(\Omega)$  obtained from the numerical integration of the auxiliary equation (O.9) are inserted.

## Appendix P

### From the average CDA-2 to the DINA in K-SAT

Equation (7.3) from Chapter 7 approximately represents the mean behavior of *CDA-2* in K-SAT. In this appendix, we show how to derive the Dynamic Approximation of Independent Neighbors (DINA) [21], which has been used as a reference technique in Chapters 5 and 7.

It is useful to recall the average *CDA-2*:

$$\begin{aligned} \frac{d\hat{P}_c(u, s_0)}{dt} = & -r(u, s_0)\hat{P}_c(u, s_0) + r(s_0, u)\hat{P}_c(s_0, u) - \\ & -(K-1)\langle r \rangle_U \left( u\hat{P}_c(u, s_0) - (u+1)\hat{P}_c(u+1, s_0) \right) - \\ & -\frac{(K-1)}{2^K - 2} \langle r \rangle_{S_0} \left( (c-u-s_0)\hat{P}_c(u, s_0) - (c-u-s_0+1)\hat{P}_c(u-1, s_0) \right) - \\ & -(K-1)\langle r \rangle_S \left( s_0\hat{P}_c(u, s_0) - (s_0+1)\hat{P}_c(u, s_0+1) \right) - \\ & -\frac{(K-1)}{2^K - 2} \langle r \rangle_S \left( (c-u-s_0)\hat{P}_c(u, s_0) - (c-u-s_0+1)\hat{P}_c(u, s_0-1) \right) \end{aligned} \quad (\text{P.1})$$

where

$$\langle r \rangle_\Omega = \sum_{\gamma=0}^{\infty} e^{-\alpha K} \frac{(\alpha K)^\gamma}{\gamma!} \sum_{\tilde{u}=0}^{\gamma} \sum_{\tilde{s}_0=0}^{\gamma-\tilde{u}} r(\tilde{u} + \delta_{\Omega, U}; \tilde{s}_0 + \delta_{\Omega, S_0}) \hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | \Omega) \quad (\text{P.2})$$

The most important point of this closure is in the expression for the conditional probabilities in Equation (P.2). For *CDA-2*, the following choices are made:

$$\hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | U) = \frac{(\tilde{u} + 1) \hat{P}_{\gamma+1}(\tilde{u} + 1, \tilde{s}_0)}{\sum_{u'=0}^{\gamma} \sum_{s'_0=0}^{\gamma-u'} (u' + 1) \hat{P}_{\gamma+1}(u' + 1, s'_0)} \quad (\text{P.3})$$

$$\hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | S_0) = \frac{(\tilde{s}_0 + 1) \hat{P}_{\gamma+1}(\tilde{u}, \tilde{s}_0 + 1)}{\sum_{u'=0}^{\gamma} \sum_{s'_0=0}^{\gamma-u'} (s'_0 + 1) \hat{P}_{\gamma+1}(u', s'_0 + 1)} \quad (\text{P.4})$$

$$\hat{P}_\gamma(\tilde{u}, \tilde{s}_0 | S) = \frac{(\gamma - \tilde{u} - \tilde{s}_0 + 1) \hat{P}_{\gamma+1}(\tilde{u}, \tilde{s}_0)}{\sum_{u'=0}^{\gamma} \sum_{s'_0=0}^{\gamma-u'} (\gamma - u' - s'_0 + 1) \hat{P}_{\gamma+1}(u', s'_0)} \quad (\text{P.5})$$

On the way to DINA, two important steps must be taken. The first one is to further simplify the equations. As observed in Equation (5.11), the probabilities in DINA depend only on the number of unsatisfied clauses  $u$ . The authors of Ref. [21] assume that:

$$\hat{P}_c(u, s_0) = \hat{P}_c(u) \binom{c-u}{s_0} \left(\frac{1}{2^K - 1}\right)^{s_0} \left(1 - \frac{1}{2^K - 1}\right)^{c-u-s_0} \quad (\text{P.6})$$

That is, all configurations that satisfy the clause are equiprobable. If the approximation (P.6) is substituted into Equation (P.1) and summed over all values of  $s_0$ , we obtain:

$$\begin{aligned} \frac{d}{dt} \hat{P}_c(u) &= -\hat{P}_c(u) \sum_{k=0}^{c-u} \binom{c-u}{k} r(u, k) \left(\frac{1}{2^K - 1}\right)^k \left(1 - \frac{1}{2^K - 1}\right)^{c-u-k} + \\ &\quad + \sum_{k=0}^{c-u} \binom{c-k}{u} r(k, u) \left(\frac{1}{2^K - 1}\right)^u \left(1 - \frac{1}{2^K - 1}\right)^{c-k-u} \hat{P}_c(k) - \\ &\quad - (K-1) \langle r \rangle_U \left\{ u \hat{P}_c(u) - (u+1) \hat{P}_c(u+1) \right\} - \\ &\quad - \frac{(K-1)}{2^K - 1} \langle r \rangle_{S_0} \left\{ (c-u) \hat{P}_c(u) - (c-u+1) \hat{P}_c(u-1) \right\} \end{aligned} \quad (\text{P.7})$$

where

$$\begin{aligned} \langle r \rangle_\Omega &= \sum_{\gamma=0}^{\infty} e^{-\alpha K} \frac{(\alpha K)^\gamma}{\gamma!} \sum_{u=0}^{\gamma} \hat{P}_\gamma(u \| \Omega) \times \\ &\quad \times \sum_{k=0}^{\gamma-u} \binom{\gamma-u}{k} r(u + \delta_{\Omega,U}, k + \delta_{\Omega,S_0}) \left(\frac{1}{2^K - 1}\right)^k \left(1 - \frac{1}{2^K - 1}\right)^{\gamma-u-k} \end{aligned} \quad (\text{P.8})$$

Once again, the relevant issue is which expression to use for the conditional probability  $\hat{P}_\gamma(u \| \Omega)$ . Following the same line as Equations (P.3), (P.4), and (P.5), we can write:

$$\hat{P}_\gamma(u \| U) = \frac{(u+1) \hat{P}_{\gamma+1}(u+1)}{\sum_{u'=0}^{\gamma} (u'+1) \hat{P}_{\gamma+1}(u'+1)} \quad (\text{P.9})$$

$$\hat{P}_\gamma(u \| S_0) = \frac{(\gamma-u+1) \hat{P}_{\gamma+1}(u)}{\sum_{u'=0}^{\gamma} (\gamma-u'+1) \hat{P}_{\gamma+1}(u')} \quad (\text{P.10})$$

At this point, we have averaged equations that can be integrated numerically. If Equation (P.7) is complemented with expressions (P.8), (P.9), and (P.10), it is possible to obtain the time dependence of  $\hat{P}_c(u)$ . However, this new average of CDA-2 is not equivalent to DINA. To continue, further simplifications will be needed, which will be discussed later.

By inserting Equation (P.9) into Equation (P.8) for  $\Omega = U$ , we obtain:

$$\langle r \rangle_U = \sum_{\gamma=0}^{\infty} e^{-\alpha K} \frac{(\alpha K)^\gamma}{\gamma!} \frac{\sum_{u=0}^{\gamma} (u+1) r_{\gamma+1}(u+1) \hat{P}_{\gamma+1}(u+1)}{\sum_{u=0}^{\gamma} (u+1) \hat{P}_{\gamma+1}(u+1)} \quad (\text{P.11})$$

where a new notation is used for the average over  $k$  of the transition probability per unit time  $r(u, k)$ .

$$r_{\gamma+1}(u+1) = \sum_{k=0}^{\gamma-u} \binom{\gamma-u}{k} r(u + \delta_{\Omega, U}, k + \delta_{\Omega, S_0}) \left( \frac{1}{2^K - 1} \right)^k \left( 1 - \frac{1}{2^K - 1} \right)^{\gamma-u-k} \quad (\text{P.12})$$

If we apply the change of variable  $c = \gamma + 1$  in Equation (P.11), it takes the form:

$$\langle r \rangle_U = \sum_{c=0}^{\infty} \frac{c}{\alpha K} e^{-\alpha K} \frac{(\alpha K)^c}{c!} \frac{\sum_{u=0}^c u r_c(u) \hat{P}_c(u)}{\sum_{u=0}^c u \hat{P}_c(u)} \quad (\text{P.13})$$

To arrive at DINA, a final approximation is needed. In Equation (P.13), an average is taken over all connectivities of a product of three factors: the fraction  $c/\alpha K$ , the sum  $\sum_{u=0}^c u r_c(u) \hat{P}_c(u)$ , and the sum  $\sum_{u=0}^c u \hat{P}_c(u)$ . It will be assumed that these quantities can be averaged independently:

$$\begin{aligned} \langle r \rangle_U &= \left( \sum_{c=0}^{\infty} \frac{c}{\alpha K} e^{-\alpha K} \frac{(\alpha K)^c}{c!} \right) \frac{\sum_{c=0}^{\infty} e^{-\alpha K} \frac{(\alpha K)^c}{c!} \sum_{u=0}^c u r_c(u) \hat{P}_c(u)}{\sum_{c=0}^{\infty} e^{-\alpha K} \frac{(\alpha K)^c}{c!} \sum_{u=0}^c u \hat{P}_c(u)} \\ \langle r \rangle_U &= \frac{\langle u r(u) \rangle_t}{\langle u \rangle_t} \end{aligned} \quad (\text{P.14})$$

where  $\langle \cdot \rangle_t = \sum_{c=0}^{\infty} e^{-\alpha K} \frac{(\alpha K)^c}{c!} \sum_{u=0}^c [\cdot] \hat{P}_c(u)$ .

One can proceed in the same way with the averages  $\langle r \rangle_{S_0}$ , which reduce to  $\langle (c-u) r(c-u) \rangle_t / \langle (c-u) \rangle_t$ . Finally, by redefining the average local probability  $\hat{P}(u, s) \equiv e^{-\alpha K} \frac{(\alpha K)^c}{c!} \hat{P}_{u+s}(u)$ , we obtain the equations for the Dynamic Independent Neighbors Approximation (DINA) as presented in subsection 5.4.1 and in Ref. [21]:

$$\begin{aligned} \frac{d\hat{P}^t(u, s)}{dt} = & -r(u, s)\hat{P}^t(u, s) + \\ & + \left(\frac{1}{2^K - 1}\right)^u \sum_{k=0}^s \binom{u+k}{k} \left(1 - \frac{1}{2^K - 1}\right)^k r(s-k, u+k) \hat{P}^t(s-k, u+k) - \\ & - \frac{(K-1)}{2^K - 1} \frac{\langle sr(u, s) \rangle_t}{\langle s \rangle_t} [s\hat{P}^t(u, s) - (s+1)\hat{P}^t(u-1, s+1)] - \\ & - (K-1) \frac{\langle ur(u, s) \rangle_t}{\langle u \rangle_t} [u\hat{P}^t(u, s) - (u+1)\hat{P}^t(u+1, s-1)] \end{aligned}$$

where  $\langle \cdot \rangle = \sum_u \sum_s [\cdot] \hat{P}(u, s)$

It has been stated that if Equation (P.7) is complemented with averages (P.8), numerical results can be obtained. To derive these equations to DINA, the factorization of Equation (P.14) was necessary. In Fig. (P.1), it is observed that the predictions of the two average equations for the algorithmic transition of FMS are almost indistinguishable. This indicates that the approximation (P.14) does not have a major impact on the accuracy of the equations.

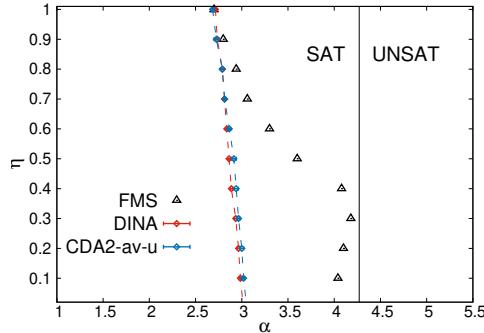


Figure P.1: Comparison between the phase diagram of the FMS algorithm in 3-SAT and the theoretical predictions of the average equations. Red points represent the results of DINA. Blue points represent the results of Equation (P.7) with averages calculated according to Equation (P.8). The boundary between the SAT and UNSAT regions of the algorithm was obtained by running it on 100 instances of the problem for different values of  $\alpha$  for a time  $t = 10^5 N$ , with  $N = 50000$ . For each  $\eta$ , the value  $\alpha_{FMS}(\eta)$  was determined at which half of the runs solve the problem (triangles). The algorithmic transition corresponding to the average equations was obtained by plotting the theoretical curves on a logarithmic scale for different values of  $\alpha$  and  $\eta$ . This determined the region where the curves quickly converge to zero. For more details on the procedure, see Appendix J.

## Part V

### Bibliography and acronyms



# Bibliography

- [1] L. Onsaguer. Cristal statistics. i. a two-dimensional model with an order-disorder transition. *Physical Review, Series II*, 65:117–149, 1944.
- [2] K. Huang. *Statistical Mechanics*. Wiley, 1987.
- [3] S. F. Edwards and P. W. Anderson. Theory of spin glasses. *Journal of Physics F: Metal Physics*, 5(5):965, 1975.
- [4] D. Sherrington and S. Kirkpatrick. Solvable model of a spin-glass. *Physical Review Letters*, 35(26):1792–1796, 1975.
- [5] M. Mezard, G. Parisi, and M. Virasoro. *Spin Glass Theory and Beyond*. World Scientific, Singapore, 1987.
- [6] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.
- [7] Y. Fu and P. W. Anderson. Application of statistical mechanics to np-complete problems in combinatorial optimisation. *Journal of Physics A: Mathematical and General*, 19(9):1605, jun 1986.
- [8] M. Mézard and G. Parisi. A replica analysis of the travelling salesman problem. *Journal De Physique*, 47(8):1285–1296, 1986.
- [9] D. Amit. *Modelling brain function: the world of attractor neural networks*. Cambridge University Press, 1989.
- [10] M. Mézard and G. Parisi. The Bethe lattice spin glass revisited. *The European Physical Journal B*, 20(2):217–233, 2001.
- [11] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.
- [12] M. Mezard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297(5582):812–815, 2002.

- [13] S. Cook. The complexity of theorem-proving procedures. In *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing*, page 151. ACM, 1971.
- [14] M. Garey and D. S. Johnson. *Computers and Intractability; A guide to the Theory of NP-completeness*. Freeman, San Francisco, 1979.
- [15] R. Mulet, A. Pagnani, M. Weigt, and R. Zecchina. Coloring random graphs. *Physical Review Letters*, 89:207206, 2002.
- [16] S. Mertens, M. Mézard, and R. Zecchina. Threshold values of random k-sat from the cavity method. *Random Structures & Algorithms*, 28(3):340–373, 2006.
- [17] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proceedings of the National Academy of Sciences*, 104(25):10318–10323, 2007.
- [18] J. Ardelius and E. Aurell. Behavior of heuristics on large and hard satisfiability problems. *Physical Review E*, 74:037702, Sep 2006.
- [19] M. Alava, J. Ardelius, E. Aurell, P. Kaski, S. Krishnamurthy, P. Orponen, and S. Seitz. Circumspect descent prevails in solving random constraint satisfaction problems. *Proceedings of the National Academy of Sciences*, 105(40):15253–15257, 2008.
- [20] R. Lemoy, M. Alava, and E. Aurell. Local search methods based on variable focusing for random  $k$ -satisfiability. *Physical Review E*, 91:013305, Jan 2015.
- [21] W. Barthel, A. K. Hartmann, and M. Weigt. Solving satisfiability problems by fluctuations: The dynamics of stochastic local search algorithms. *Physical Review E*, 67:066104, 2003.
- [22] G. Semerjian and M. Weigt. Approximation schemes for the dynamics of diluted spin models: the ising ferromagnet on a bethe lattice. *Journal of Physics A: Mathematical and General*, 37(21):5525–5546, may 2004.
- [23] J. Hatchett, I. Pérez-Castillo, A. Coolen, and N. S. Skantzos. Dynamical replica analysis of disordered ising spin systems on finitely connected random graph. *Physical Review Letters*, 95:117204, 2005.
- [24] A. Mozeika and A. Coolen. Dynamical replica analysis of processes on finitely connected random graphs: I. vertex covering. *Journal of Physics A: Mathematical and Theoretical*, 41(11):115003, 2008.

- [25] B. Karrer and M. E. J. Newman. Message passing approach for general epidemic models. *Physical Review E*, 82:016101, 2010.
- [26] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200, 2001.
- [27] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925, 2015.
- [28] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics*, 81:591–646, May 2009.
- [29] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010.
- [30] T. M. Pham, J. Korbel, R. Hanel, and S. Thurner. Empirical social triad statistics can be explained with dyadic homophylic interactions. *Proceedings of the National Academy of Sciences*, 119(6):e2121103119, 2022.
- [31] M. Mezard and A. Montanari. *Information, physics, and computation*. Oxford University Press, 2009.
- [32] A. S. Mata and S. C. Ferreira. Pair quenched mean-field theory for the susceptible-infected-susceptible model on complex networks. *EPL (Europhysics Letters)*, 103(4):48003, 2013.
- [33] I. Neri and D. Bollé. The cavity approach to parallel dynamics of Ising spins on a graph. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(08):P08009, 2009.
- [34] T. Barthel, C. De Bacco, and S. Franz. Matrix product algorithm for stochastic dynamics on networks applied to nonequilibrium Glauber dynamics. *Physical Review E*, 97:010104, Jan 2018.
- [35] E. Aurell, G. D. Ferraro, E. Domínguez, and R. Mulet. A cavity master equation for the continuous time dynamics of discrete spins models. *Physical Review E*, 95:052119, 2017.
- [36] E. Aurell, E. Domínguez, D. Machado, and R. Mulet. Exploring the diluted ferromagnetic p-spin model with a cavity master equation. *Physical Review E*, 97:05103(R), 2018.
- [37] E. Domínguez, D. Machado, and R. Mulet. The cavity master equation: average and fixed point of the ferromagnetic model in random graphs. *Journal of Statistical Mechanics: Theory and Experiment*, 2020:073304, 2020.

- [38] E. Ortega, D. Machado, and A. Lage-Castellanos. Dynamics of epidemics from cavity master equations: Susceptible-infectious-susceptible models. *Physical Review E*, 105:024308, Feb 2022.
- [39] D. Machado and R. Mulet. From random point processes to hierarchical cavity master equations for stochastic dynamics of disordered systems in random graphs: Ising models and epidemics. *Physical Review E*, 104:054303, Nov 2021.
- [40] E. Aurell, E. Domínguez, D. Machado, and R. Mulet. A theory of non-equilibrium local search on random satisfaction problems. *Physical Review Letters*, 123:230602, 2019.
- [41] E. Aurell, D. Machado, and R. Mulet. A closure for the master equation starting from the dynamic cavity method. *Journal of Physics A: Mathematical and Theoretical*, 56(17):17LT02, apr 2023.
- [42] D. Machado, R. Mulet, and F. Ricci-Tersenghi. Improved mean-field dynamical equations are able to detect the two-step relaxation in glassy dynamics at low temperatures. *Journal of Statistical Mechanics: Theory and Experiment*, 2023(12):123301, dec 2023.
- [43] S. Franz, M. Mézard, and F. Ricci-Tersenghi. A ferromagnet with a glass transition. *Europhysics Letters*, page 465, 2001.
- [44] F. Ricci-Tersenghi, M. Weigt, and R. Zecchina. Simplest random k-satisfiability problem. *Physical Review E*, 63:026702, Jan 2001.
- [45] S. Franz, M. Leone, F. Ricci-Tersenghi, and R. Zecchina. Exact solutions for diluted spin glasses and optimization problems. *Physical Review Letters*, 87:127209, Aug 2001.
- [46] M. Newman and G. Barkema. *Monte Carlo Methods in Statistical Physics*. Clarendon Press, 1999.
- [47] J.-C. Walter and G. Barkema. An introduction to monte carlo methods. *Physica A: Statistical Mechanics and its Applications*, 418:78–87, 2015. Proceedings of the 13th International Summer School on Fundamental Problems in Statistical Physics.
- [48] A. Montanari and G. Semerjian. On the dynamics of the glass transition on bethe lattices. *Journal of Statistical Physics*, 124:103–189, 2006.
- [49] M. Mézard and A. Montanari. Reconstruction on trees and spin glass transition. *Journal of Statistical Physics*, 124:1317–1350, 2006.

- [50] A. Montanari and G. Semerjian. Rigorous inequalities between length and time scales in glassy systems. *Journal of Statistical Physics*, 125:23–54, 2006.
- [51] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian. Clusters of solutions and replica symmetry breaking in random k-satisfiability. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(04):P04004, apr 2008.
- [52] S. Seitz, M. Alava, and P. Orponen. Focused local search for random 3-satisfiability. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(06):P06006, jun 2005.
- [53] L. Budzynski, F. Ricci-Tersenghi, and G. Semerjian. Biased landscapes for random constraint satisfaction problems. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(2):023302, feb 2019.
- [54] N. van Kampen. *Stochastic Processes in Physics and Chemistry*, volume 1. Elsevier, Amsterdam, 1992.
- [55] H. Sompolinsky and A. Zippelius. Dynamic theory of the spin-glass phase. *Physical Review Letters*, 47:359–362, Aug 1981.
- [56] H. Sompolinsky. Time-dependent order parameters in spin-glasses. *Physical Review Letters*, 47:935–938, Sep 1981.
- [57] K. Binder and A. P. Young. Spin glasses: Experimental facts, theoretical concepts, and open questions. *Reviews of Modern Physics*, 58:801–976, Oct 1986.
- [58] M. Alba, M. Ocio, and J. Hammann. Ageing process and response function in spin glasses: An analysis of the thermoremanent magnetization decay in ag:mn (2.6%). *Europhysics Letters*, 2(1):45, jul 1986.
- [59] M. Alba, J. Hammann, M. Ocio, P. Refregier, and H. Bouchiat. Spin-glass dynamics from magnetic noise, relaxation, and susceptibility measurements (invited). *Journal of Applied Physics*, 61(8):3683–3688, 1987.
- [60] L. Cugliandolo and J. Kurchan. Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model. *Physical Review Letters*, 71(1):173, 1993.
- [61] L. F. Cugliandolo and J. Kurchan. On the out-of-equilibrium relaxation of the sherrington-kirkpatrick model. *Journal of Physics A: Mathematical and General*, 27(17):5749, sep 1994.

- [62] L. F. Cugliandolo and D. S. Dean. Full dynamical solution for a spherical spin-glass model. *Journal of Physics A: Mathematical and General*, 28(15):4213, 1995.
- [63] J.-P. Bouchaud, L. F. Cugliandolo, J. Kurchan, and M. Mézard. Out of equilibrium dynamics in spin-glasses and other glassy systems. *Spin glasses and random fields*, 12, 1998.
- [64] J. P. Bouchaud. Weak ergodicity breaking and aging in disordered systems. *Journal de Physique I France*, 2(9):1705–1713, 1992.
- [65] M. Bernaschi, A. Billoire, A. Maiorano, G. Parisi, and F. Ricci-Tersenghi. Strong ergodicity breaking in aging of mean-field spin glasses. *Proceedings of the National Academy of Sciences*, 117(30):17522–17527, 2020.
- [66] B. Derrida, E. Gardner, and A. Zippelius. An exactly solvable asymmetric neural network model. *EPL (Europhysics Letters)*, 4(2):167, 1987.
- [67] A. Crisanti and H. Sompolinsky. Dynamics of spin systems with randomly asymmetric bonds: Ising spins and glauber dynamics. *Physical Review A*, 37:4865, 1988.
- [68] H. Rieger, M. Schreckenberg, and J. Zittartz. Glauber dynamics of the asymmetric sk-model. *Z. Phys. B - Condensed Matter*, 74:527–538, 1989.
- [69] J. P. L. Hatchett, B. Wemmenhove, I. P. Castillo, T. Nikoletopoulos, N. S. Skantzos, and A. C. C. Coolen. Parallel dynamics of disordered ising spin systems on finitely connected random graphs. *Journal of Physics A: Mathematical and General*, 37(24):6201, jun 2004.
- [70] R. J. Glauber. Time-dependent statistics of the ising model. *Journal of Mathematical Physics*, 4:294, 1963.
- [71] A. Coolen and D. Sherrington. Order-parameter flow in the sk spin glass. i. replica symmetry. *Journal of Physics A: Mathematical and General*, 27(23):7687, 1994.
- [72] A. Coolen, S. Laughton, and D. Sherrington. Dynamical replica theory for disordered spin systems. *Physical Review B*, 53:8184, 1996.
- [73] S. Laughton, A. Coolen, and D. Sherrington. Order-parameter flow in the sk spin-glass: II. inclusion of microscopic memory effects. *Journal of Physics A: Mathematical and General*, 29(4):763, 1996.

- [74] A. Mozeika and A. Coolen. Dynamical replica analysis of processes on finitely connected random graphs: II. dynamics in the griffiths phase of the diluted ising ferromagnet. *Journal of Physics A: Mathematical and Theoretical*, 42(19):195006, 2009.
- [75] H. Nishimori and M. Yamana. Dynamical probability distribution function of the sk model at high temperatures. *Journal of the Physical Society of Japan*, 65(1):3–6, 1996.
- [76] A. Coolen, S. Laughton, and D. Sherrington. Perturbative large deviation analysis of non-equilibrium dynamics. *Journal of the Physical Society of Japan*, 93:084001, 2014.
- [77] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos. Epidemic spreading in real networks: an eigenvalue viewpoint. In *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.*, pages 25–34, 2003.
- [78] E. Cator and P. Van Mieghem. Second-order mean-field susceptible-infected-susceptible epidemic threshold. *Physical review E*, 85(5):056111, 2012.
- [79] E. Aurell and H. Mahmoudi. A message-passing scheme for non-equilibrium stationary states. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(04):P04014, 2011.
- [80] E. Aurell and H. Mahmoudi. Dynamic mean-field and cavity methods for diluted ising systems. *Physical Review E*, 85(3):031119, 2012.
- [81] R. R. Wilkinson and K. J. Sharkey. Message passing and moment closure for susceptible-infected-recovered epidemics on finite networks. *Physical Review E*, 89:022808, Feb 2014.
- [82] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina. Optimizing spread dynamics on graphs by message passing. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(09):P09011, sep 2013.
- [83] G. Del Ferraro and E. Aurell. Dynamic message-passing approach for kinetic spin models with reversible dynamics. *Physical Review E*, 92(1):010102, 2015.
- [84] S. Crotti and A. Braunstein. Matrix product belief propagation for reweighted stochastic dynamics over graphs. *Proceedings of the National Academy of Sciences*, 120(47):e2307935120, 2023.
- [85] F. Behrens, B. Hudcová, and L. Zdeborová. Backtracking dynamical cavity method. *Physical Review X*, 13:031021, Aug 2023.

- [86] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes Volume I: Elementary Theory and Methods*. Springer Science & Business Media, Berlin, 2002.
- [87] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems*, volume 320. Springer Verlag, 1999.
- [88] R. G. Gallager. Low-density parity-check codes. *IRE Transactions on Information Theory*, 8(1):21–28, 1962.
- [89] L. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. In *Proceedings of the IEEE*, volume 77, pages 257–286, 1989.
- [90] A. Dembo and A. Montanari. Ising models on locally tree-like graphs. *The Annals of Applied Probability*, 20(2):565 – 592, 2010.
- [91] E. Cuevas. An agent-based model to evaluate the covid-19 transmission risks in facilities. *Computers in Biology and Medicine*, 121:103827, 2020.
- [92] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.
- [93] R. S. Sander, G. S. Costa, and S. C. Ferreira. Sampling methods for the quasistationary regime of epidemic processes on regular and complex networks. *Physical Review E*, 94:042308, Oct 2016.
- [94] L. Viana and A. J. Bray. Phase diagrams for dilute spin glasses. *Journal of Physics C: Solid State Physics*, 18(15):3037–3051, may 1985.
- [95] G. Semerjian and R. Monasson. Relaxation and metastability in a local search procedure for the random satisfiability problem. *Physical Review E*, 67:066103, 2003.
- [96] P. Van Mieghem, J. Omic, and R. Kooij. Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17(1):1–14, 2009.
- [97] C. Castellano and R. Pastor-Satorras. Thresholds for epidemic spreading in networks. *Physical Review Letters*, 105:218701, Nov 2010.
- [98] D. H. Silva, S. C. Ferreira, W. Cota, R. Pastor-Satorras, and C. Castellano. Spectral properties and the accuracy of mean-field approaches for epidemics on correlated power-law networks. *Physical Review Research*, 1:033024, Oct 2019.

- [99] D. H. Silva, F. A. Rodrigues, and S. C. Ferreira. High prevalence regimes in the pair-quenched mean-field theory for the susceptible-infected-susceptible model on networks. *Physical Review E*, 102:012313, Jul 2020.
- [100] M. F. P. B. Aspvall and R. E. Tarjan. A linear-time algorithm for testing the truth of certain quantified boolean formulas. *Information Processing Letters*, 8:121–123, 1979.
- [101] C. Papadimitriou. On selecting a satisfying truth assignment. In N. Y. IEEE Computer Society, editor, *Proceedings of the 32nd IEEE Symposium on the Foundations of Computer Science*, volume FOCS-91, page 163–169, 1991.
- [102] B. Selman, H. Kautz, and B. Cohen. Local search strategies for satisfiability testing. In D. S. Johnson and M. A. Trick, editors, *Coloring, and Satisfiability: Second DIMACS Implementation Challenge*, volume 26, pages 521–532, 1996.
- [103] J. Bouchaud and M. Mézard. Self induced quenched disorder: a model for the glass transition. *Journal de Physique I France*, 4(8):1109–1114, 1994.
- [104] M. Bellitti, F. Ricci-Tersenghi, and A. Scardicchio. Entropic barriers as a reason for hardness in both classical and quantum algorithms. *Physical Review Research*, 3(4):043015, 2021.



# Acronyms

BP .....	<i>Belief Propagation.</i>
CDA .....	<i>Conditional Dynamics Approximation.</i>
CME .....	<i>Cavity Master Equation.</i>
DCM .....	<i>Dynamic Cavity Method.</i>
DINA .....	<i>Dynamic Independent Neighbors Approximations.</i>
DMP .....	<i>Dynamic Message Passing.</i>
DRT .....	<i>Dynamical Replica Theory.</i>
ER .....	<i>Erdős-Rényi</i> random graph.
FMS .....	<i>Focused Metropolis Search.</i>
IBMF .....	<i>Individual-Based Mean Field</i> , mean-field approximation for epidemic spreading.
K-SAT .....	Random <i>K-Satisfiability</i> , combinatorial optimization problem.
MPBP .....	<i>Matrix Product Belief Propagation</i> , closure for the dynamic cavity method.
PBMF .....	<i>Pair-Based Mean Field</i> , mean-field approximation for epidemic spreading.
RRG .....	<i>Random Regular Graph.</i>
RSB .....	<i>Replica Symmetry Breaking.</i>
SA .....	<i>Simulated Annealing.</i>
SID .....	<i>Survey Inspired Decimation.</i>
SIR .....	<i>Susceptible-Infectious-Recovered</i> , compartmental model for epidemic spreading.

- SIS ..... *Susceptible-Infectious-Susceptible*, compartmental model for epidemic spreading.
- SK ..... *Sherrington-Kirkpatrick* model of spin glasses.
- TRPP ..... *Theory of Random Point Processes*.