A statistical mechanics model of the adaptive immune system static and dynamical analysis in different connectivity regimes

Bartolucci, Silvia

Awarding institution:
King's College London

A statistical mechanics model of the adaptive immune system: static and dynamical analysis in different connectivity regimes

Silvia Bartolucci
Department of Mathematics
King's College London

In partial fulfilment of the requirements for the degree of
Doctor of Philosophy

To myself.

And my family.

Abstract

In this thesis project, we model a sub-part of the adaptive immune system, composed of B and T lymphocytes, which interact to produce a suitable immune response against antigens.

From a statistical mechanics perspective, this system can be modelled as a bipartite network with sparse links where the nodes represent B and T cells respectively, signalling via particular proteins called cytokines. Assuming that B lymphocytes evolve on a faster timescale than T cells, we study the dynamics of an effective mono-partite graph of T cells only where the B cells have been integrated out. Interestingly, this system can be mapped into a Hopfield-like associative network, which is able to retrieve and perform multiple immune strategies simultaneously.

Using techniques such as Kramers-Moyal expansions for master equations, we carry out a dynamical analysis of the network evolving via Glauber sequential update. We derive equations quantifying the evolution in time of the immune response strength, analysing the nature and the stability of the stationary solutions in different regimes of dilution and network connectivity via linear stability analysis and Monte Carlo simulations.

The model has also been extended to include the effect of receptors promiscuity, sampling B-T interactions locally from heterogeneous degree distributions

and the effect of the antigens. Finally, we introduce interactions between B lymphocytes, called idiotypic interactions, studying their effect on the system's dynamics.

We also analyse the effect of idiotypic interactions in the high storage and finite connectivity regime at equilibrium, using the cavity method to derive equations for the distributions of observables of the system. In particular, we obtain the B clone size distribution, studying its behaviour in different regions of the phase space.

Acknowledgements

A PhD is a lot about patience and perfectionism. That is what I learned during these years, especially thanks to my supervisor Dr. Alessia Annibale and her constant advice and support during the whole journey.

I wish to thank the research groups that welcomed me: the Disordered Systems Group, the IMMB and the MABRA group. In particular, I wish to thanks Prof. Ton Coolen for his friendship and help, Dr. Mozeika for the lively discussions and the interesting scientific collaboration, Prof. D. Dunn-Walters for her insights on the secrets of the immune systems. It has been a challenging and pleasant experience collaborating with them.

One of the most precious things I gained from this PhD was also the possibility to meet amazing people, with whom I hope to continue sharing pieces of life. My friend, flatmate and colleague Aleksandra Aloric for being able to stand my presence 24/7, wishing her that all the efforts will be repaid by a bright and gratifying career and life, wherever in the world. Silvia Grigolon, because everything is easier, when you have a friend that you can always count on. Barbara Bravi, for the pleasant time spent together. Ada Altieri, for the beautiful times in Beg Rohu and her friendship.

There is one more fundamental person completing this picture, who saw me

crying in despair, who celebrated with me any little achievement and who guided me when I was lost: I made it, and that is thanks to you as well.

Last but not least, my beloved family in Italy, hoping that somehow the huge price in distance I let you pay over these year is compensated by this little achievement that I dedicate to you.

Previously published work

(1) S. Bartolucci, A. Annibale
Associative networks with diluted patterns: dynamical analysis at low and medium load
J. Phys. A: Math. Theor. **47(41)**, 415001 (2014).

(2) S. Bartolucci, A. Annibale
A dynamical model of the adaptive immune system: effects of cells promiscuity, antigens and BB interactions
J. Stat. Mech. Theor. Exp. **2015(8)**, P08017 (2015).

(3) S. Bartolucci, A. Mozeika and A. Annibale
The role of idiotypic interactions in the adaptive immune system: a belief-propagation approach
Preprint arXiv:1605.01290 (Accepted for publication in *J. Stat. Mech. Theor. Exp.*) (2016).

(4) E. Agliari , A. Barra, S. Bartolucci, A. Galluzzi, F. Guerra and F. Moauro
Parallel processing in immune networks
Phys. Rev. E **87(4)**, 042701 (2013).

The papers (1), (2) and (3) cover the content of chapters 2 , 3 and 4 respectively. Paper (4) has mainly been used for the introduction in chapter 1.

Contents

List of Figures

Glossary

- **Antigen**: substance or molecule that induces an immune response and the antibodies production.
- **Epitope**: a specific piece of an antigen that is recognised by the immune system, e.g. by antibodies, B cells, or T cells.
- **BCR**: a B-cell receptor is a transmembrane protein located on the external surfaces of B-cells. It has a unique antigen binding site, through which B cells recognise external pathogens.
- **TCR**: a T-cell receptor is a transmembrane protein located on the external surface of T-cells. This receptor can recognise fragments of antigen as peptides bound to the so-called major histocompatibility complex (MHC) molecules.
- **Clone**: group of identical cells derived from the same progenitor sharing the same type of antigen receptor.
- **Cytokine**: signalling protein that is secreted by lymphocytes and affects the behaviour of nearby cells with appropriate receptors.

I find it astonishing that the immune system embodies a degree of complexity which suggests some more or less superficial though striking analogies with human language, and that this cognitive system has evolved and functions without assistance of the brain.

N. K. Jerne, Nobel Lecture.

1 Introduction

The immune system is a complex collection of organs, tissues and cells which is present in all vertebrates and protects the organism from external pathogens and infections [1, 2]. The constituents of this system coordinate themselves, reciprocally sending signals, to produce an *immune response* targeted at defeating the infection.

The initial protection against external microbes is provided by the so-called *innate immunity*. Epithelial barriers, which contain natural antibiotics, can block the body invasion, while so-called natural killer (NK) cells and phagocytes destroy the microbes that managed to elude the first defence barrier and entered blood circulation or tissues. Despite the fact that the innate immune response can potentially deal with many types of microbes and infections, the microbial evolution has also created pathogens able to resist it. To deal with this class of pathogens, the body takes advantage of the *adaptive immune system*, whose main constituents are *lymphocytes* and their products, such as antibodies. The adaptive immune response is, generally, enhanced and triggered by the components

of the innate immune system, but is characterised by a higher *specificity*: lymphocytes have specific receptors that can bind to particular substances produced by the microbes, called *antigens*. The main constituents of the adaptive immune system are:

- *B cells*: the only lymphocytes that are able to produce *antibodies*, i.e. Y-shaped proteins that can recognise, tag and eventually kill external invaders.
- *T cells*: they are responsible for the so called *cell-mediated immunity* and can recognise fragments of antigens if bound to particular cells called Antigen Presenting Cells (APC).

In the following, we will introduce the details of the interaction process between T and B lymphocytes, which is responsible for the onset of an immune response (sec. 1.1). In addition, in sec. 1.2 we will discuss the mathematical approach to the adaptive immune system modelling and the statistical mechanics tools that we will use in this thesis.

1.1 Behind the scenes of the adaptive immune response

When a pathogen enters the host overcoming the first defence barrier of the innate immune system, it is (hopefully) recognised and attacked by lymphocytes. B cells, in particular, have specific receptors called BCR to recognise the antigen binding to it. The antigen is then internalised, broken into pieces and displayed on the B cell surface: B cells behave as an antigen presenting cell (APC), showing parts of the recognised antigen and waiting for an extra confirmation signal to be activated.

At this stage, T cells come into play. They are not able to directly identify and bind to an antigen, but can recognise it if displayed on APC cells. Indeed,

T cells have a receptor, the TCR, which can bind to the B cells and trigger their activation or suppression sending confirmatory signals in the form of excitatory or inhibitory signalling proteins, the *cytokines*.

This response (when activated) consists of the secretion of *antibodies* (proteins able to chemically bind and neutralise the antigen) by B cells, thus (possibly) avoiding the propagation of the infection. This two-signals mechanism prevents erroneous B cells activations from mismatching antigens or other cells of the organism.

In this process, both activated T and B cells proliferate, undergoing the so-called *clonal expansion*. This generates the amount of cells required to eliminate the infection from a small pool of naïve lymphocytes, which have never encountered an antigen before. They form *clones* or groups of identical cells sharing the same type of receptors, either TCR or BCR, and therefore responding to the same type of antigens. The newly generated B and T cells also differentiate into *effector* cells, which effectively deal with the infection attacking the antigens, and *memory* cells, which keep information and memory of the infection in the immune system. The memory cells remain in an inactive state in the system but are able to respond faster to a new encounter of the same pathogen: this immune response, which is more rapid and effective than the first one, is called *secondary immune response*. Once the infection has been eliminated, the stimulatory signals triggering the lymphocytes are removed and cells die during the *apoptosis* process.

To summarise, *specificity* and *memory* are two features of the adaptive immune response that allow for a targeted and tailored reaction to pathogens at the first encounter with them and a faster and stronger response in case of repeated infections.

1.2 Immunology and Mathematics

The 1950s witnessed several important progresses in our theoretical understanding of the adaptive immune system. In particular, two important theories were formulated:

- *The clonal selection theory* by F. Macfarlane Burnet [3], stating that each lymphocyte expresses only one receptor type, *specific* for a particular antigen.
- *The idiotypic network theory* by N. K. Jerne [4], stating that lymphocytes can interact via the so-called *idiotypes*, unique determinants of cell receptors [5], forming a network.

These theories, formulated by experimentalists, were originally presented in a qualitative form but paved the way to a mathematical formulation. For instance, models of clonal selection theory include the first work by Bell [6] describing the B cells clonal proliferation and antibodies production in the presence of antigens, followed by several modifications and extensions by Perelson and collaborators [7–9]. In these models, the evolution in time of the B clones system was described by a set of coupled differential equations, solved via numerical integration. In addition, the clonal selection theory raised important questions about the size of the repertoire, i.e. the size of the receptors space, needed to recognise any antigen that one may encounter. This repertoire may include also self-reactive cells attacking the organism, which survived the selection process [2]. Quantifying the receptor space in both size and specificity was important to understand the mechanisms of self/non-self recognition and antigen identification [10]. The repertoire completeness issue is also subject of several more recent investigations, based on Bayesian inference and maximum entropy models including comparison with available data [11–13].

For what concerns the idiotypic network theory, several models of interactions between cells have been proposed to investigate their effect on the immune response. We will summarise them in more details in sec. 3.3, where we will also introduce our approach to the modelling of these interactions. For an excellent review of previous literature on the subject we refer to [14].

Mathematical modelling also includes T clones dynamics and activation via stochastic and probabilistic models. One of the main issues analysed in this context is the T clones activation threshold, i.e. the minimum number of triggered TCR needed to activate their response [15–17].

Finally, there is a general tendency to gradually abandon the reductionist approach, based on single cells or molecules analysis, and rather focus on the collective behaviour and emergent properties of the adaptive immune system as a whole [12, 18–24].

The first model to pave the way to a statistical mechanics approach to the immune system modelling was introduced in 1990 by G. Parisi [18]. In this work, antibodies concentrations c_i, $i = 1, \dots, N$, in the absence of antigens, are modelled as binary variables $0, 1$ evolving in time in discrete steps τ according to the simple rule

$$c_i(t + \tau) = \theta[h_i(t)] \ , \tag{1.2.1}$$

where $\theta[x]$ is zero for $x < 0$ and 1 otherwise, and $h_i(t) = \sum_{j=1}^{N} J_{ij} c_j(t)$. The matrix J_{ij} represents the interactions between antibodies i and j. The state of the immune system is, then, determined by the values of the concentrations of all possible N antibodies.

The main problem consists in understanding how the network decides which antibodies should be produced and are physiologically relevant. To this purpose, one assumes that only a fraction α of the N antibodies can be predefined and

studies what the highest number $M = \alpha N$ of antibodies concentrations (compatible with (1.2.1)) in a system of N possible antibodies is. In this framework, M can be mapped into the *storage capacity* of the network and the maximum capacity can be explicitly computed. Hence, the problem can be translated into finding one or more equilibrium states with M predetermined concentrations where the maximum value of M ($M_c = \alpha_c N$) for which it is still possible to find equilibrium states represents the storage capacity of the system.

In this model, the interactions are assumed to be symmetric ($J_{ij} = J_{ji}$), with zero diagonal terms ($J_{ii} = 0$) and randomly sampled from the interval $[-1, 1]$: under these conditions it reduces to the well-known infinite-range spin glass model [25]. Indeed, in this work the common features between the immune system, spin glasses and neural networks were discussed for the first time.

Under this assumption one can easily compute the value of α_c using standard techniques in statistical mechanics. Indeed, for spin glasses it has been shown that the number of stationary configurations is proportional to $2^{\lambda N}$ with $\lambda \sim 0.3$ [25]. Assuming that the average number of solutions can be also seen as the most probable, then there exists an equilibrium state only for $\alpha < \alpha_c = \lambda$. Hence, using a simple framework one can show that the number of selected antibodies to be produced or inhibited is proportional to the size of the whole repertoire.

In more recent years, the system composed of interacting T and B lymphocytes has been looked at through the prism of statistical mechanics to understand its global features and functionalities and a systematic connection between neural networks and immune system has been made [20–24].

Fully connected models of idiotypic interactions have been analysed in [20–22] to understand the mechanism of clonal anergy in self-directed lymphocytes, which managed to escape from the selection process during the initial stages of development [1], and the inhibitory effects of the idiotypic network. These models also include important basic phenomena characterising real immune systems such as

low-dose tolerance, i.e. unresponsiveness to antigens presented in low concentrations, memory of the encountered pathogens and self/non-self discrimination, i.e. the ability to distinguish external invaders from the host's cells.

The signalling process between B and T lymphocytes using statistical mechanics techniques has been introduced in [23], where the basic features of the model on a fully connected topology are discussed and the mapping to an associative memory with "cytokine" patterns representing immunological strategies is presented. The full connectivity assumption was, then, relaxed in [24] where associative memories with diluted patterns were first introduced in the context of the immune system modelling. The features of these types of networks, in particular the ability to simultaneously recall multiple patterns or strategies, have been studied in *statics* in different regimes of dilution and at low, medium storage and near saturation in [24, 26, 27].

In particular, in the low and medium load [24, 26], the retrieval properties of the network, i.e. the number of patterns successfully recognised, and the different types of activation, i.e. symmetric (all patterns retrieved in the same fashion) or hierarchical (one patterns is prioritised with respect to the others), have been studied varying the parameter of the system, in particular noise and dilution. In the high storage regime [27], using graph theory and the replica method it has been shown that N T cells can orchestrate the immune responses to an extensive number of antigenic attacks simultaneously, coordinating $P = \alpha N$ B cells in parallel. Increasing α, the system undergoes a second order transition to a phase characterised by clonal cross-talk, where each B clone feels the presence and interfere with all others B clones in the system, possibly affecting its activation. However, in the clonal cross-talk region the system's performance in terms of patterns retrieval is not drastically compromised, but degrades smoothly.

In the following, we will analyse the *dynamics* of these B-T interactions models and we will extend them to incorporate other important mechanisms of the

immune response. In particular, the analysis of the dynamics of the system will allow us to find the region of stability of the various types of activation patterns (symmetric or hierarchical) in the different regions of the phase space in the low and medium storage regime, as we will show in ch. 2. Moreover, we will include in the model idiotypic interactions, receptors promiscuity and antigens to analyse their effects on the retrieval capabilities of the system (3). Finally, we will also study the system in the saturation regime, including B-T and B-B interactions and using the cavity method to detect the clonal cross-talk region in this framework, varying the system's parameters including noise and B-B interaction strength (ch. 4).

1.3 The statistical mechanics approach

We summarise below this model, which constitutes the starting point of our thesis. We consider an immune repertoire composed of P different B clones whose log-concentration is $b_\mu,\ \mu = 1, \dots, P$ and N T cells $\sigma_i,\ i = 1, \dots, N$. B cells can be modelled as real variables: in the absence of interactions with T cells, i.e. at rest, we assume the b_μ's to be Gaussian. When a B clone receives signals from T clones, it can either expand (growing by several orders of magnitude with respect to the typical size at rest), i.e. $b_\mu \gg b_0$, or be inhibited and reduce its size, i.e. $b_\mu \ll b_0$. T clones are modelled as binary variables or "spins" $\sigma_i = \pm 1$, $i = 1, \dots, N$, which can be active *i.e.* secreting cytokines $(+1)$ or quiescent (-1). B and T clones communicate with each other exchanging signalling proteins, the *cytokines*, which are assumed to be quenched discrete random variables. A cytokine ξ_i^μ sent from a T clone i to a B clone μ can be excitatory $\xi_i^\mu = +1$, inhibitory $\xi_i^\mu = -1$ or there might be no connection between the specific clones i and μ, hence $\xi_i^\mu = 0$. In fig. 1.1 we provide a schematic summary of the main features of this model.

The interacting system of B-T cells can be phenomenologically described by

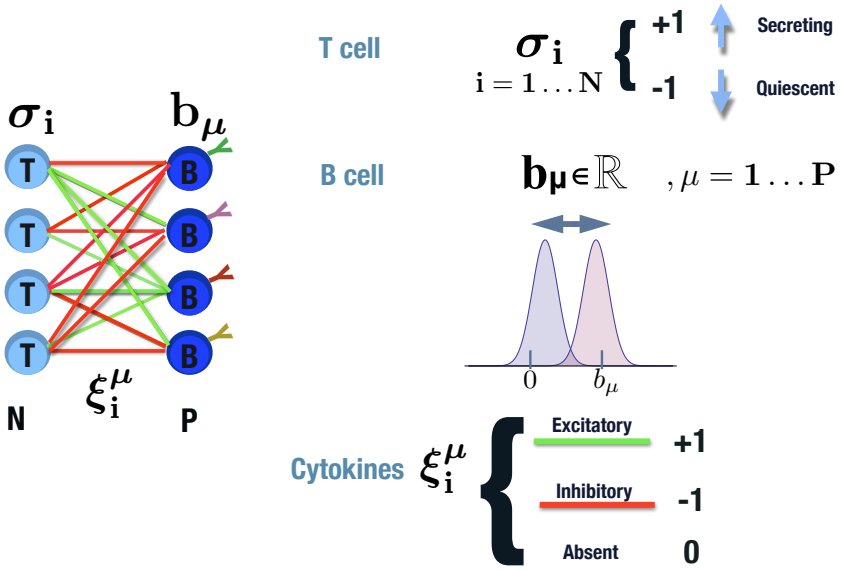

Figure 1.1: Schematic representation of the B-T cells interactions via cytokines and summary of the assumptions of the model.

the Hamiltonian

$$\mathcal{H}(\boldsymbol{\sigma}, \mathbf{b}|\boldsymbol{\xi}) = -N^\theta \sum_{i=1}^{N} \sum_{\mu=1}^{P} \xi_i^\mu \sigma_i b_\mu + \frac{1}{2} \sum_{\mu=1}^{P} b_\mu^2 \ , \tag{1.3.2}$$

where the exponent θ ensures the correct scaling of the Hamiltonian with the system size N. In order to better understand its meaning, let us assume that the B clones variables evolve according to a gradient descent on the Hamiltonian (1.3.2), yielding

$$\frac{\mathrm{d}b_\mu}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial b_\mu} + \eta_\mu(t) = N^\theta \sum_{i=1}^{N} \xi_i^\mu \sigma_i - b_\mu + \eta_\mu(t) \ . \tag{1.3.3}$$

B clones concentration increases (decreases) in presence of excitatory (inhibitory) interactions with T clones, as modelled by the first term in (1.3.3). Moreover, B clones may die at a rate proportional to the log population size (second term of (1.3.3)), which for simplicity we assume to be the same for all clones. In the absence of T clones and assuming a Gaussian white noise with $\langle \eta_\mu(t) \rangle = 0$ and

$\langle \eta_\mu(t)\eta_\nu(t')\rangle = 2T\delta_{\mu,\nu}\delta(t-t')$, where $T = 1/\beta$ is the strength of the noise, eq. (1.3.3) represents an Ornstein-Uhlenbeck process with stationary distribution

$$p(b_\mu) = \frac{1}{\sqrt{2\pi T}} \mathrm{e}^{-\frac{b_\mu^2}{2T}} \ . \tag{1.3.4}$$

The partition function $Z_N(\beta|\boldsymbol{\xi})$ at inverse noise level β, which is consistent with our assumption $b_\mu = \mathcal{N}(0,T)$, is given by

$$Z_N(\beta|\boldsymbol{\xi}) = \sum_{\boldsymbol{\sigma}} \int \mathrm{d}b_1 \dots \mathrm{d}b_P \exp\left[-\beta\mathcal{H}(\boldsymbol{\sigma},\mathbf{b}|\boldsymbol{\xi})\right] \ . \tag{1.3.5}$$

Integrating over b_μ yields

$$Z_N(\beta|\boldsymbol{\xi}) = \sum_{\boldsymbol{\sigma}} \exp\left[-\beta\mathcal{H}(\boldsymbol{\sigma}|\boldsymbol{\xi})\right] \ , \tag{1.3.6}$$

where

$$\mathcal{H}(\boldsymbol{\sigma}|\boldsymbol{\xi}) = -\frac{1}{2N^\theta} \sum_{i,j=1}^{N} J_{ij}\sigma_i\sigma_j \ , \tag{1.3.7}$$

$$J_{ij} = \sum_{\mu=1}^{P} \xi_i^\mu \xi_j^\mu . \tag{1.3.8}$$

Interestingly, eq. (1.3.7) represents the Hamiltonian of the Hopfield model for associative networks [28–30], whose features will be introduced in next section (sec. 1.3.1). Therefore, the interacting system of B and T cells via cytokines as described by the Hamiltonian $\mathcal{H}(\boldsymbol{\sigma},\mathbf{b}|\boldsymbol{\xi})$ is thermodynamically equivalent to a Hopfield-like network [28–30] encoded in an effective Hamiltonian $\mathcal{H}(\boldsymbol{\sigma}|\boldsymbol{\xi})$ containing interacting T cells only via the cytokines. A schematic representation of the marginalisation process is shown in fig. 1.2.

1.3.1 The Hopfield model in a nutshell

The Hopfield model was introduced to model neurons interactions via synaptic connections in the brain. In particular, the idea was to model the human memory

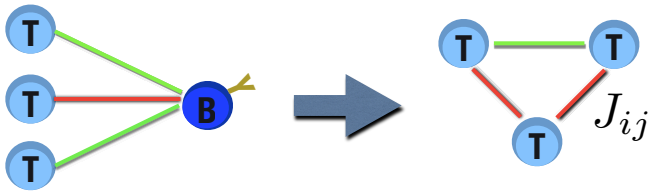

Figure 1.2: Scheme of the marginalisation process that leads from a bipartite network of interacting B and T clones (left) to an effective T clones-only system (right).

as a collective phenomenon involving large networks of connected neurons. The N neurons in the brain are equivalent to the T cells in the immune system model: they can, indeed, be in two states, firing (+1) or quiescent (-1) and can be modelled as spin variables $\sigma_i = \pm 1,\ i = 1, \dots, N$. During the learning process [31], we store P *patterns* of information, each as a N-bit vector $\boldsymbol{\xi}^\mu = (\xi_1^\mu, \xi_2^\mu, \dots, \xi_N^\mu)$, $\mu = 1, \dots, P$ with entries $\xi_i^\mu \in \{+1, -1\}$. The state of the network at each time t is defined by the configuration of all neurons $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_N)$.

The dynamics of the network state is determined by the interactions J_{ij} between neurons, defined in eq. (1.3.8) according to the so-called *Hebb Rule* [30]. The interactions directly depend on the information stored in the system: when a new pattern is stored, J_{ij}'s are linearly modified by the addition of a new term in the sum. The learning process, therefore, corresponds to a modification of the connections between neurons. In the Hopfield model, patterns are randomly and independently drawn from the distribution

$$\mathbb{P}(\xi_i^\mu) = \frac{1}{2}\left(\delta_{\xi_i^\mu,+1} + \delta_{\xi_i^\mu,-1}\right) \ . \tag{1.3.9}$$

The stored patterns, hence, act as a *quenched disorder* in the system. The latter is described by the Hamiltonian,

$$\mathcal{H}(\boldsymbol{\sigma}|\boldsymbol{\xi}) = -\frac{1}{2N}\sum_{i,j=1}^{N} J_{ij}\sigma_i\sigma_j \ . \tag{1.3.10}$$

This system works differently from other types of memories, in particular for what concerns the process of information retrieval; Hopfield-like memories, indeed, work by *association* using the actual content of the information patterns rather than the physical location where they are stored to recall the information. The human memory is a good example of associative memory as it associates a new input with the memories it contains, stored during the learning process. When seeing a face or listening to a part of a song we can reconstruct by *association* a previously learnt piece of information.

Hence, this system works as an *associative memory* when it manages to retrieve one of the stored patterns of information: this happens when the network configuration exactly corresponds to one of the stored patterns, namely when $\sigma_i = \xi_i^\mu,\ \forall i = 1, \dots, N$. To monitor the retrieval properties one introduces the so-called *magnetisations* or *overlaps*, which are defined as follows

$$m_\mu = \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \sigma_i \ , \tag{1.3.11}$$

and measure the similarity between the network state $\boldsymbol{\sigma}$ and the μ-th stored pattern. A magnetisation different from zero would indicate the retrieval of the corresponding pattern.

It has long been known that away from saturation, i.e. for $P \ll N$, the patterns $\{\boldsymbol{\xi}^\mu\}$ are stable attractors of the system dynamics for any $T < 1$. The specific attractor approached under dynamical evolution depends on the system's initialisation. In addition, even/odd mixtures of n patterns are also solutions, although they do not correspond to any stored information and this retrieval is generally undesired. At low temperature, odd mixtures of n patterns are stable and they destabilise at n-dependent temperatures T_n, which increase with decreasing n. This implies that, in the vicinity of $T = 1$ from below, only pure states ($n = 1$) are stable [28, 29]. Even mixtures of n patterns are instead always unstable. Hence, this system is able to retrieve *one* pattern at a time.

1.3.2 The importance of being diluted

One of the strongest non-biological assumption of the Hopfield model lies in the full connectivity choice, meaning that all neurons send signals to all others. To improve this choice, several *diluted* models have been proposed, where for each neuron a fraction c of randomly selected links has been removed. Different choices of the scaling of c with the system size N have been analysed, leading to different behaviours of the system:

- $c = N$, which corresponds to the case with no dilution [29, 32];
- $\lim_{N\to\infty} c = +\infty$ and $\lim_{N\to\infty} \frac{c}{N} = 0$ or *extreme dilution* case [33, 34];
- $c = \mathcal{O}(1)$, *finite connectivity* with dilution independent of the system size N [35].

The goal was mostly to probe the network resilience to dilution and its capability to work as an associative memory. Still the information processing remained unchanged, with the system being able to handle and retrieve one pattern at a time.

But this is not the only possible way to introduce the dilution in such networks. Using the mapping introduced in sec. 1.3, one could think of cutting links in the bipartite graph first and only then transform it into a Hopfield-like associative memory. Interestingly, this process leads to a drastic change in the system's performance, making it able to parallel process the information and simultaneously retrieve multiple patterns. In practical terms, this choice corresponds to introducing "zeros" in the patterns entries, whose distribution is now chosen of the form

$$\mathbb{P}(\{\boldsymbol{\xi}\}) = \prod_{i\mu} \left[d\left(\delta_{\xi_i^\mu,1} + \delta_{\xi_i^\mu,-1}\right) + (1-d)\delta_{\xi_i^\mu,0} \right] , \tag{1.3.12}$$

where d tunes the dilution level in the network. In the retrieval phase, when the system state $\boldsymbol{\sigma}$ tries to converge and align with one particular pattern, for example $\boldsymbol{\xi}^1$, not all the entries $\boldsymbol{\sigma}$ are involved: the entries corresponding to the zeros of $\boldsymbol{\xi}^1$ can be aligned with a second pattern, e. g. $\boldsymbol{\xi}^2$, and so forth as shown in fig. 1.3. A detailed analysis of the properties of such systems in equilibrium can be found in [36], where multitasking associative networks were first introduced.

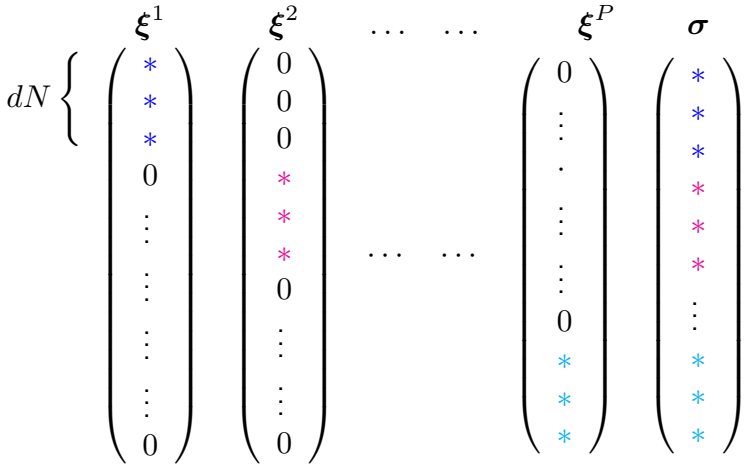

Figure 1.3: Illustrative explanation of the parallel retrieval. During the retrieval process the network state $\boldsymbol{\sigma}$ entries are first aligned with the dN non-zero entries of $\boldsymbol{\xi}^1$ (asterisks represent ± 1 entries); the remaining $\boldsymbol{\sigma}$ entries are aligned with the non-zero entries of the second available pattern vector $\boldsymbol{\xi}^2$, etc.

Intuitively, it is clear that the retrieval of an extensive number of patterns $P = \mathcal{O}(N)$ will require finite connectivity, i.e. $d = \frac{c}{N}$ with $c = \mathcal{O}(N^0)$. A sub-extensive number of patterns $P = \mathcal{O}(N^\delta)$, with $\delta < 1$, can be instead recalled in the extremely diluted regime, where $d = \frac{c}{N^\gamma}$ and $\gamma < 1$. In chapters 2,3 we will solve the dynamics of this model and its extensions in the extremely diluted regime. In chapter 4 we will consider the static of this model in the finitely connected case.

The truth is rarely pure and never simple.

Oscar Wilde, The Importance of Being Earnest

2

Dynamics of associative memories and B-T clones interactions

2.1 Introduction

Associative memories are neural networks able to retrieve and recognise previously stored information. The pioneering work on brain memory is due to J. J. Hopfield, where the neural network performs the retrieval of one information pattern at a time. Recently, new types of associative memories with diluted patterns have been introduced [36] as a generalisation of the standard Hopfield model [28–30,32]. Interestingly, the introduction of dilution in the stored patterns of information drastically changes the collective behaviour of the system, switching its operational behaviour from sequential to parallel processing of multiple pieces of information at the same time.

This feature has potentially sound applications in artificial intelligence [31,37, 38], where existing models can only accomplish a sequential processing of information. In addition, it has been shown that these models are thermodynamically equivalent to bipartite networks [36] useful to describe systems with two interacting parties of agents. In particular, they have recently been used to model signalling processes between B and T lymphocytes in immune networks, which are able to orchestrate several immune responses in parallel [24].

Despite the wide range of possible applications, these types of memories have been studied analytically only in statics so far, where several dilution and storage regimes were investigated [26,27,39]. Dynamical studies are limited to numerical simulations and are restricted to the low storage regime, where the number of stored patterns is finite [40].

In this chapter, we will study the dynamics of the associative memories with diluted patterns in different dilution and storage regimes via non-equilibrium statistical mechanics techniques, first developed in [41–43] for standard Hopfield networks and spin glasses.

2.2 Introducing the dynamics

In this section, we study the dynamics of diluted associative memories with interactions of the form in eq. 1.3.8. Patterns are quenched random variables, independently and identically distributed with probability distribution

$$\mathbb{P}(\{\boldsymbol{\xi}\}) = \prod_{i\mu} \left[\frac{c}{2N^{\gamma}} \left(\delta_{\xi_i^{\mu},1} + \delta_{\xi_i^{\mu},-1} \right) + \left(1 - \frac{c}{N^{\gamma}} \right) \delta_{\xi_i^{\mu},0} \right] , \tag{2.2.1}$$

where $\{\boldsymbol{\xi}\}$ involves a collection of PN random variables ξ_i^{μ}. The fraction of non-zero entries in the stored patterns will determine the degree of dilution of the system: for $\gamma = 0$ the system is finitely diluted, whereas for $\gamma > 0$ the system is extremely diluted. Indicating the number of stored patterns with $P = \alpha N^{\delta}$,

with $\alpha \in \mathbb{R}^+$ and tuning $\delta \in [0, 1)$ one may explore different storage regimes. For $\delta > 0$, the number of patterns stored in the system diverges in the thermodynamic limit, however as long as $\delta < 1$, it grows sub-linearly with the system size, and the system is away from saturation. We assume that the T clones σ_i evolve minimising the same energy function as the B clones

$$\mathcal{H}(\boldsymbol{\sigma}, \mathbf{b}|\boldsymbol{\xi}) = -N^{\gamma-1} \sum_{i=1}^{N} \sum_{\mu=1}^{P} \xi_i^\mu \sigma_i b_\mu + \frac{1}{2} \sum_{\mu=1}^{P} b_\mu^2 \ , \tag{2.2.2}$$

where the prefactor $N^{\gamma-1}$ ensures that the Hamiltonian is $\mathcal{O}(N)$. We can, therefore, write their dynamics in the presence of noise in discrete time with steps Δ as

$$\sigma_i(t + \Delta) = \text{sgn}\left(\sum_{\mu=1}^{P} \xi_i^\mu b_\mu(t) + \zeta_i(t) \right) \ , \tag{2.2.3}$$

where $\zeta_i(t)$ represents a zero-average noise drawn from $P(\zeta) = \frac{\beta}{2}(1 - \tanh^2(\beta\zeta))$ with inverse temperature β. In addition, assuming that B clones evolve on faster timescales than T clones, we can insert the stationary solution of (1.3.3) in (2.2.3), obtaining

$$\sigma_i(t + \Delta) = \text{sgn}\left(\sum_{\mu=1}^{P} \sum_{j=1}^{N} \xi_i^\mu \xi_j^\mu \sigma_j(t) + \zeta_i(t) \right) \ . \tag{2.2.4}$$

For $\Delta = \frac{1}{N}$ with $N \gg 1$, the σ_i's follow a random sequential Glauber dynamics [31] via an effective local field

$$h_i(\boldsymbol{\sigma}) = \frac{1}{N^{1-\gamma}} \sum_{j \neq i}^{N} J_{ij} \sigma_j \ . \tag{2.2.5}$$

The interactions J_{ij} are defined in (1.3.8) and the normalisation $N^{1-\gamma}$ ensures that $h_i = \mathcal{O}(N^0)$. The probability $p_t(\boldsymbol{\sigma})$ of finding the system in a state $\boldsymbol{\sigma} =$

$(\sigma_1, \ldots, \sigma_N)$ at time t, evolves according to the master equation

$$\frac{\partial p_t(\boldsymbol{\sigma})}{\partial t} = \sum_{i=1}^{N} \left[p_t(F_i(\boldsymbol{\sigma})) w_i(F_i(\boldsymbol{\sigma})) - p_t(\boldsymbol{\sigma}) w_i(\boldsymbol{\sigma}) \right] , \tag{2.2.6}$$

where F_i is the i-th spin-flip operator $F_i(\sigma_1, .., \sigma_i, ..., \sigma_N) = (\sigma_1, ..., -\sigma_i, ...\sigma_N)$ and transition rates between $\boldsymbol{\sigma}$ and $F_i(\boldsymbol{\sigma})$ have the Glauber form

$$w_i(\boldsymbol{\sigma}) = \frac{1}{2}[1 - \sigma_i \tanh(\beta h_i(\boldsymbol{\sigma}))] , \tag{2.2.7}$$

where $\beta = 1/T$ is the inverse temperature of the system, giving the rate of spontaneous spin flips.

For symmetric couplings, such as those given in (1.3.8), the dynamics obeys detailed balance and the process evolves towards equilibrium, described by the Boltzmann distribution, with Hamiltonian:

$$\mathcal{H}(\boldsymbol{\sigma}|\boldsymbol{\xi}) = -\frac{1}{2N^{1-\gamma}} \sum_{i,j=1}^{N} \sum_{\mu=1}^{P} \xi_i^\mu \xi_j^\mu \sigma_i \sigma_j . \tag{2.2.8}$$

We note, however, that our approach does not require this symmetry restriction and paves the way to a broader range of applications where symmetric interactions would not be realistic, particularly in theoretical immunology [24]. One can show that the prefactor provides the correct normalisation to the Hamiltonian [26]. A heuristic argument is that, since the number of non-zero entries in a given pattern is $\mathcal{O}(N^{1-\gamma})$, one has, for condensed [1] patterns, $\sum_i \xi_i^\mu \sigma_i = \mathcal{O}(N^{1-\gamma})$ with the expected number of condensed patterns being $\mathcal{O}(N/N^{1-\gamma}) = \mathcal{O}(N^\gamma)$, so their contribution to the Hamiltonian is $\mathcal{O}(N)$. Non-condensed patterns, for which $\sum_i \xi_i^\mu \sigma_i = \mathcal{O}(N^{(1-\gamma)/2})$, are present only for $\delta > \gamma$ and are $\mathcal{O}(N^\delta)$, thus providing a contribution which is always sub-leading away from saturation, as expected.

[1] Set of patterns that have been retrieved, hence with non zero associated magnetisation.

2.2.1 Macroscopic dynamics

For large N, finding a solution for the set of 2^N coupled, non-linear differential equations (2.2.6) is very hard, and a convenient approach is to use the microscopic stochastic laws for the T clones dynamics (2.2.6) to derive dynamical equations (2.2.24) for the probability distribution of a suitable set of macroscopic observables. As for the standard Hopfield model, it is convenient to choose these as (suitably normalised) overlaps between the microscopic configurations of the system and the stored patterns $\boldsymbol{\xi}^\mu$, $\mu = 1, \ldots, P$

$$m_\mu(\boldsymbol{\sigma}) = \frac{1}{cN^{1-\gamma}} \sum_{j=1}^{N} \xi_j^\mu \sigma_j \ , \tag{2.2.9}$$

with the prefactor chosen to make these quantities of order $\mathcal{O}(1)$. These order parameters quantify the pattern retrieval of the system, as in the absence of retrieval they vanish in the thermodynamic limit, whereas an order $\mathcal{O}(1)$ overlap m_μ indicates that the system has retrieved the pattern $\boldsymbol{\xi}^\mu$. We note that for large N and $\gamma < 1$, the effective fields (2.2.5) depend on $\boldsymbol{\sigma}$ only through the overlaps, i.e. $h_i(\boldsymbol{\sigma}) = c \sum_\mu \xi_i^\mu m_\mu(\boldsymbol{\sigma})$, where c controls the dilution.

In the context of the immune system modelling, the overlap m_μ quantifies the activation level of the μ-th B clone through the signals sent from the T clones σ_j according to the stored "immune strategies" or cytokine patterns ξ_j^μ.

2.2.1.1 Master equation and Kramers-Moyal expansion

The probability of finding the system, at time t, in a state where the macroscopic parameters (2.2.9) take values $\mathbf{m} = (m_1, ..., m_P)$, is

$$P_t(\mathbf{m}) = \sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma}) \delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma})) \ . \tag{2.2.10}$$

The latter evolves in time, due to equation (2.2.6) governing the microscopic probability $p_t(\boldsymbol{\sigma})$, as

$$\frac{\mathrm{d}P_t(\mathbf{m})}{\mathrm{d}t} = \sum_{\boldsymbol{\sigma}} \sum_{i=1}^{N} \delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma})) \left[p_t(F_i(\boldsymbol{\sigma})) w_i(F_i(\boldsymbol{\sigma})) - p_t(\boldsymbol{\sigma}) w_i(\boldsymbol{\sigma}) \right] , \quad (2.2.11)$$

where relabelling $F_i(\boldsymbol{\sigma}) \to \boldsymbol{\sigma}$ in the first term on the RHS gives

$$\frac{\mathrm{d}P_t(\mathbf{m})}{\mathrm{d}t} = \sum_{\boldsymbol{\sigma}} \sum_{i=1}^{N} p_t(\boldsymbol{\sigma}) w_i(\boldsymbol{\sigma}) [\delta(\mathbf{m} - \mathbf{m}(F_i(\boldsymbol{\sigma}))) - \delta(\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma}))] . \quad (2.2.12)$$

By introducing the increments

$$\Delta_{i\mu} = m_\mu(F_i(\boldsymbol{\sigma})) - m_\mu(\boldsymbol{\sigma}) = \frac{2}{cN^{1-\gamma}} \xi_i^\mu \sigma_i \quad (2.2.13)$$

and Taylor expanding (2.2.12) in powers of $\Delta_{i\mu}$, we obtain the so-called Kramers-Moyal (KM) expansion. In order for the expansion to hold, increments $\Delta_{i\mu}$ must be small. This condition is satisfied in the thermodynamic limit $N \to \infty$ for any $\gamma < 1$. The case $\gamma = 1$, where each T clone is finitely connected to other T clones, violates the condition necessary to perform the KM expansion and will not be considered here. The finitely connected regime ($\gamma = 1$) for a similar model will be studied in chapter 4 at equilibrium. By Taylor expanding the δ-function in powers of $\Delta_{i\mu}$, we obtain

$$\begin{aligned} \frac{\mathrm{d}P_t(\mathbf{m})}{\mathrm{d}t} = & \sum_{i=1}^{N} \sum_{\boldsymbol{\sigma}} w_i(\boldsymbol{\sigma}) p_t(\boldsymbol{\sigma}) \times \\ & \times \sum_{\ell \geq 1} \frac{(-1)^\ell}{\ell!} \sum_{\mu_1=1}^{P} \cdots \sum_{\mu_\ell=1}^{P} \frac{\partial^\ell}{\partial m_{\mu_1} \cdots \partial m_{\mu_\ell}} \left[\delta[\mathbf{m} - \mathbf{m}(\boldsymbol{\sigma})] [\Delta_{i\mu_i} \cdots \Delta_{i\mu_\ell}] \right] . \end{aligned} \quad (2.2.14)$$

Defining

$$F^{(\ell)}_{\mu_1 \ldots \mu_\ell}(\mathbf{m}, t) = \left\langle \sum_{i=1}^{N} w_i(\boldsymbol{\sigma}) \Delta_{i\mu_1} \cdots \Delta_{i\mu_\ell} \right\rangle_{\mathbf{m},t} , \quad (2.2.15)$$

with

$$\langle(\dots)\rangle_{\mathbf{m},t} = \frac{\sum_{\boldsymbol{\sigma}} \delta(\mathbf{m}-\mathbf{m}(\boldsymbol{\sigma}))(\dots)p_t(\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}} \delta(\mathbf{m}-\mathbf{m}(\boldsymbol{\sigma}))p_t(\boldsymbol{\sigma})} \ , \tag{2.2.16}$$

we can write the equation for the time evolution of $P_t(\mathbf{m})$ as

$$\frac{\partial P_t(\mathbf{m})}{\partial t} = \sum_{\ell\geq 1} \frac{(-1)^\ell}{\ell!} \sum_{\mu_1=1}^{P} \cdots \sum_{\mu_\ell=1}^{P} \frac{\partial^\ell}{\partial m_{\mu_1} \cdots \partial m_{\mu_\ell}} [P_t(\mathbf{m}) F^{(\ell)}_{\mu_1 \dots \mu_\ell}(\mathbf{m}, t)] \ . \tag{2.2.17}$$

Keeping only terms up to the first order in the $\Delta_{i\mu}$'s leads to the Liouville equation, corresponding to a deterministic evolution of the order parameters $\mathbf{m}$. Second order terms lead to the Fokker-Planck equation, which includes diffusive processes [44]. Hence, a sufficient condition for the observables $\mathbf{m}(\boldsymbol{\sigma})$ to evolve deterministically, in the limit $N \to \infty$, is

$$\lim_{N\to\infty} \sum_{\ell\geq 2} \frac{(-1)^\ell}{\ell!} \sum_{\mu_1=1}^{P} \cdots \sum_{\mu_\ell=1}^{P} \sum_{i=1}^{N} \frac{\partial^\ell}{\partial m_{\mu_1} \cdots \partial m_{\mu_\ell}} \langle |\Delta_{i\mu_1} \cdots \Delta_{i\mu_\ell}| \rangle_{\mathbf{m},t} = 0 \ . \tag{2.2.18}$$

Each $\Delta_{i\mu}$, as defined in (2.2.13), contributes to the sums an order $\mathcal{O}(N^{-1})$ because the entries ξ_i^μ are non-zero with probability $cN^{-\gamma}$. Recalling that $P = \alpha N^\delta$, condition (2.2.18) reduces to

$$\lim_{N\to\infty} N^{2\delta-1} = 0 \ , \tag{2.2.19}$$

which is satisfied for $\delta < \frac{1}{2}$. For this range of values of the storage load δ, higher order terms in the KM expansion vanish in the thermodynamic limit and the equation for $P_t(\mathbf{m})$ reduces, to leading order in N, to the Liouville equation

$$\frac{\partial P_t(\mathbf{m})}{\partial t} = -\sum_{\mu=1}^{P} \frac{\partial}{\partial m_\mu} \left[P_t(\mathbf{m}) F^{(1)}_\mu(\mathbf{m}) \right] \ , \tag{2.2.20}$$

where $F_\mu^{(1)}$ results from insertion of (2.2.7) and (2.2.13) in (2.2.15) and is independent of time. Setting $\hat{\beta} = \beta c$, we have

$$F_\mu^{(1)}(\mathbf{m}) = \frac{1}{cN^{1-\gamma}} \sum_{i=1}^{N} \xi_i^\mu \tanh\left(\hat{\beta} \sum_{\nu=1}^{P} \xi_i^\nu m_\nu\right) - m_\mu \ . \tag{2.2.21}$$

Inserting $1 = \sum_{\boldsymbol{\xi}} \delta_{\boldsymbol{\xi},\boldsymbol{\xi}_i}$ and rearranging the terms we obtain

$$F_\mu^{(1)}(\mathbf{m}) = \frac{N^\gamma}{c} \left\langle \xi^\mu \tanh\left(\hat{\beta} \sum_{\nu=1}^{P} \xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} - m_\mu \ . \tag{2.2.22}$$

Indeed, for large N we can define $\mathbb{P}(\boldsymbol{\xi}) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{\boldsymbol{\xi},\boldsymbol{\xi}_i}$ and introduce the average $\langle f(\boldsymbol{\xi})\rangle_{\boldsymbol{\xi}} = \sum_{\boldsymbol{\xi}} \mathbb{P}(\boldsymbol{\xi}) f(\boldsymbol{\xi})$ hence obtaining (2.2.22). In particular, in this step the average over all sites for one specific realisation of $\boldsymbol{\xi}$ has been replaced by the average over the distribution of patterns at one particular site.

If the number N of vectors $\boldsymbol{\xi}_i$ is much larger than the number 2^P of vectors $\boldsymbol{\xi}$, this replacement is justified by appealing to the law of large numbers. However, for larger P this statement is non-trivial and relies on the assumption of self-averageness, whereby a quantity does not depend on the particular realisation of the disorder $\{\xi_i^\mu\}$ but only on the distribution the disorder is sampled from [29,31]. The distribution $\mathbb{P}(\boldsymbol{\xi})$

$$\mathbb{P}(\boldsymbol{\xi}) = \prod_\mu \left[\frac{c}{2N^\gamma}\left(\delta_{\xi^\mu,1} + \delta_{\xi^\mu,-1}\right) + \left(1 - \frac{c}{N^\gamma}\right)\delta_{\xi^\mu,0}\right] \ , \tag{2.2.23}$$

then involves a collection of P random variables $\{\xi^\mu\}$ instead of the original NP random variables $\{\xi_i^\mu\}$. We note that $F_\mu^{(1)}$ is $\mathcal{O}(1)$ as each ξ^μ is non-zero with probability $cN^{-\gamma}$. The Liouville equation corresponds to a deterministic evolution for the m_μ's, given by the set of ordinary, coupled, non-linear differential equations

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \frac{N^\gamma}{c} \left\langle \xi^\mu \tanh\left(\hat{\beta} \sum_{\nu=1}^{P} \xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} - m_\mu \ . \tag{2.2.24}$$

In the following sections we will investigate the nature and stability of the steady state solutions of (2.2.24) in different regions of the tunable parameters (noise, dilution, storage load).

2.2.2 The steady state: possible classes of solutions

The steady state of the dynamics, $\mathrm{d}\mathbf{m}/\mathrm{d}t = \mathbf{0}$, is given by the self-consistency equation for $\mathbf{m}$

$$\mathbf{m} = \frac{N^\gamma}{c} \left\langle \boldsymbol{\xi} \tanh\left(\hat{\beta}(\mathbf{m}\cdot\boldsymbol{\xi})\right)\right\rangle_{\boldsymbol{\xi}} \ . \tag{2.2.25}$$

One can show that the system undergoes a phase transition at $T_c = c$, with the equilibrium phase at $T > T_c$ characterised by $\mathbf{m} = \mathbf{0}$ and retrieval occurring at $T < T_c$, where $\mathbf{m} \neq \mathbf{0}$. Taking the scalar product of (2.2.25) with $\mathbf{m}$ and using the inequality $|\tanh x| \leq |x|$, gives

$$\mathbf{m}^2 \leq \frac{N^\gamma}{c}\hat{\beta}\left\langle (\mathbf{m}\cdot\boldsymbol{\xi})^2\right\rangle_{\boldsymbol{\xi}} = \hat{\beta}\mathbf{m}^2 \ , \tag{2.2.26}$$

which implies $\mathbf{m} = \mathbf{0}$ for $\hat{\beta} > 1$. By linearising the system (2.2.24) about $\mathbf{m} = \mathbf{0}$,

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \left(\hat{\beta} - 1\right) m_\mu \ , \tag{2.2.27}$$

we find that $\mathbf{m} = \mathbf{0}$ is stable for $\hat{\beta} < 1$ and reached exponentially fast with rate $\tau = 1/(1-\hat{\beta})$. For $\hat{\beta} = 1$, i.e. at the critical temperature, the first order terms vanish and an expansion of (2.2.24) up to third order

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = m_\mu\left(\left(\frac{c}{N^\gamma} - \frac{1}{3}\right)m_\mu^2 - \frac{c}{N^\gamma}\mathbf{m}^2\right) + \mathcal{O}(m_\mu^5) \tag{2.2.28}$$

shows a power law decay $m_\mu \sim t^{-1/2}$ towards $\mathbf{m} = \mathbf{0}$. Hence, for $T > c$ there is no retrieval. On the other hand, for $T < c$ the solution $\mathbf{m} = \mathbf{0}$ becomes unstable and we expect nonzero solutions $\mathbf{m} \neq \mathbf{0}$ to bifurcate continuously from the $\mathbf{m} = \mathbf{0}$ solution as we decrease the temperature below T_c. We will inspect the stability

of the non-zero steady state solutions $\mathbf{m}^\star$ below T_c by means of linear stability analysis.

What is, then, the structure of the first class of solutions $\mathbf{m}^\star$ to bifurcate away from $\mathbf{m}=\mathbf{0}$ below T_c? To answer this question, we Taylor expand (2.2.25) for small $|\mathbf{m}|$ in powers of $\epsilon=\hat{\beta}-1$:

$$\begin{aligned} m_\mu &\simeq \frac{N^\gamma}{c}\left[\sum_{\nu=1}^P \hat{\beta}\langle\xi^\nu\xi^\mu\rangle_{\boldsymbol{\xi}} m_\nu - \frac{\hat{\beta}^3}{3}\sum_{\nu,\rho,\lambda=1}^P \langle\xi^\nu\xi^\mu\xi^\rho\xi^\lambda\rangle_{\boldsymbol{\xi}} m_\nu m_\rho m_\lambda + \mathcal{O}(\mathbf{m}^5)\right] \\ &= (1+\epsilon)m_\mu - \frac{m_\mu}{3}\left(m_\mu^2 + \frac{3c}{N^\gamma}\sum_{\rho\neq\mu}^P m_\rho^2\right) + \mathcal{O}(\mathbf{m}^5,\epsilon\mathbf{m}^3)\ , \end{aligned} \tag{2.2.29}$$

where we have used $\hat{\beta}^3=(1+\epsilon)^3\simeq 1+3\epsilon+\ldots$ and anticipated that $m_\mu=\mathcal{O}(\epsilon^{1/2})$. Note that from the definition of $\mathbb{P}(\boldsymbol{\xi})$ in (2.2.23) it follows that $\langle\xi^\mu\rangle_{\boldsymbol{\xi}}=0$ and $\langle\xi^\mu\xi^\nu\rangle_{\boldsymbol{\xi}}=\frac{c}{N^\gamma}\delta_{\mu\nu}$. This yields

$$0 = m_\mu\left[\epsilon - \frac{1}{3}\left(m_\mu^2 + \frac{3c}{N^\gamma}\left(\mathbf{m}^2 - m_\mu^2\right)\right)\right]\ , \tag{2.2.30}$$

which gives either the trivial solution $m_\mu=0$ or

$$m_\mu^2 = \frac{3\epsilon - \frac{3c}{N^\gamma}\mathbf{m}^2}{1-\frac{3c}{N^\gamma}}\ . \tag{2.2.31}$$

For $\gamma>0$ and large N we get

$$m_\mu^2 = 3\epsilon - \frac{3c}{N^\gamma}\mathbf{m}^2\ , \tag{2.2.32}$$

whereas for $\gamma=0$ one has

$$m_\mu^2 = 3\frac{\epsilon - c\mathbf{m}^2}{1-3c}\ . \tag{2.2.33}$$

Since each component m_μ only depends on the whole vector $\mathbf{m}$, it is clear that, close to criticality, we have, for each component, $m_\mu\in\{-m,0,m\}$, where m is the amplitude of the magnetisations. Using the invariance of the dynamical equations under $m_\mu\to -m_\mu$, we can from now on focus on non-negative solutions, which

we can write, up to permutations of pattern indices, as $\mathbf{m} = m(1, \dots, 1, 0, \dots, 0)$, with

$$\mathbf{m}^2 = m^2 n, \tag{2.2.34}$$

where n represents the number of non-null entries of the vector of magnetisations $\mathbf{m}$. We will call this class of solution *symmetric*, as condensed patterns are retrieved all with the *same intensity*.

Different storage and dilution regimes will yield different scaling for the number n of condensed patterns, as we shall detail below. In the following sections, we will derive the stability region of this class of solutions and we will discuss how the system breaks the symmetry at low temperature.

2.2.3 Noise distribution

Interestingly, the nature and stability of the non-trivial solution below criticality depend on the values of the parameters δ and γ, controlling the storage load and the dilution in our system, respectively. To investigate this, it is convenient to manipulate (2.2.24)

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \left\langle \tanh\left(\hat{\beta}\left(m_\mu + \sum_{\nu\neq\mu}^{P} \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} - m_\mu \ , \tag{2.2.35}$$

where we already performed the average over $\xi^\mu \neq 0$[2]. Inserting $1 = \int_{-\infty}^{+\infty} \mathrm{d}z \delta\big(z - \sum_{\nu\neq\mu}^{P} \xi^\nu m_\nu\big)$, we can express the dynamical equations

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \int_{-\infty}^{+\infty} \mathrm{d}z \tanh\left(\hat{\beta}\left(m_\mu + z\right)\right) P_\mu(z|\{m_\nu\}) - m_\mu \tag{2.2.36}$$

[2]In the following, for the sake of simplicity, the notation $\langle(\cdots)\rangle_{\boldsymbol{\xi}}$ will denote either the full average over all ξ's, or the average over the remaining ξ's after performing the average over non-zero values of one (or more) patterns ξ^μ. The use of one meaning or the other should be clear from the context.

in terms of the noise distribution

$$P_\mu(z|\{m_\nu\}) = \left\langle \delta\left(z - \sum_{\nu\neq\mu}^{P} \xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} . \tag{2.2.37}$$

The latter can be calculated using the Fourier representation of the δ-function

$$P_\mu(z|\{m_\nu\}) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}\hat{z}}{2\pi} \mathrm{e}^{\mathrm{i}z\hat{z}} \left\langle \prod_{\nu\neq\mu}^{P} \mathrm{e}^{-\mathrm{i}\hat{z}\xi^\nu m_\nu} \right\rangle_{\boldsymbol{\xi}} \tag{2.2.38}$$

and performing the average over the disorder

$$\langle \mathrm{e}^{-\mathrm{i}\hat{z}\xi^\nu m_\nu}\rangle_{\xi^\nu} = 1 + \frac{c}{N^\gamma}(\cos(\hat{z}m_\nu) - 1) \simeq \mathrm{e}^{\frac{c}{N^\gamma}(\cos(\hat{z}m_\nu)-1)} , \tag{2.2.39}$$

where the last approximation holds for large N. This leads to

$$P_\mu(z|\{m_\nu\}) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}\hat{z}}{2\pi} \mathrm{e}^{\mathrm{i}z\hat{z}} \mathrm{e}^{\sum_{\nu\neq\mu}^{P} \frac{c}{N^\gamma}(\cos(\hat{z}m_\nu)-1)} . \tag{2.2.40}$$

Different choices of the parameters δ, γ, confer different properties to the distribution (2.2.40), thus leading to different dynamical behaviours. In particular, when $\gamma > 0$, extending the sum to all ν in (2.2.40) gives a negligible contribution in the thermodynamic limit, so the distribution $P_\mu(z|\{m_\nu\})$ converges, for large N, to the distribution

$$P(z|\mathbf{m}) = \left\langle \delta\left(z - \sum_{\nu=1}^{P} m_\nu \xi^\nu\right)\right\rangle_{\boldsymbol{\xi}} = \int_{-\infty}^{+\infty} \frac{\mathrm{d}\hat{z}}{2\pi} \mathrm{e}^{\mathrm{i}z\hat{z} + \sum_{\nu=1}^{P} \frac{c}{N^\gamma}(\cos(\hat{z}m_\nu)-1)} , \tag{2.2.41}$$

which depends on the whole vector $\mathbf{m} = (m_1, \ldots, m_P)$. This highlights the possibility of having symmetric solutions as fixed points of the dynamics in large regions of the phase diagram, meaning that all patterns are retrieved with the same intensity. In contrast, for $\gamma = 0$, the distribution in (2.2.40) retains its dependence on μ in the thermodynamic limit and one expects the symmetry

of patterns to be broken and that the system retrieves patterns with different intensities, i.e. *hierarchically*.

In addition, we note that for $\gamma > \delta$, (and $\gamma > 0$, since $\delta \geq 0$), using $P = \alpha N^{\delta}$, the probability distribution $P(z|\mathbf{m})$ reduces to a δ-function in the thermodynamic limit

$$P(z|\mathbf{m}) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}\hat{z}}{2\pi} \mathrm{e}^{\mathrm{i}z\hat{z}} \mathrm{e}^{\sum_{\nu=1}^{\alpha N^{\delta}} \frac{c}{N^{\gamma}} (\cos(\hat{z}m_{\nu})-1)} \simeq \delta(z) \ . \tag{2.2.42}$$

This leads to a decoupling of the dynamical equations (2.2.36) and the model reduces to P independent Hopfield models with a single stored pattern, or, equivalently, to a set of P independent Curie-Weiss ferromagnets,

$$\frac{\mathrm{d}m_{\mu}}{\mathrm{d}t} = \tanh\left(\hat{\beta}m_{\mu}\right) - m_{\mu} \ , \qquad \mu = 1, \ldots, P \ , \tag{2.2.43}$$

with a rescaled critical temperature $\hat{T}_c = T_c/c$, due to the presence of dilution in the patterns. Each overlap approaches exponentially (one of) the non-zero (gauge-symmetric) solutions of $m = \tanh(\hat{\beta}m)$ and the stable steady state of the system is given by $\mathbf{m} = m(1, \ldots, 1)$, for any temperature below criticality. This shows that the system is able to retrieve *all* the stored patterns in parallel and that the retrieval is symmetric. This behaviour is confirmed by Monte Carlo simulations, shown in figure 2.1, where the trajectories of the overlaps are seen to approach (up to finite size effects) the expected steady states, represented by the symbols. Details of the Monte Carlo simulation algorithm are summarised in appendix A.

Conversely, for any $\gamma \leq \delta$, the dynamical equations (2.2.36) are coupled, which is a signature of *interference* between patterns. This holds true even when the number of patterns is finite, i.e. for $\delta = \gamma = 0$. In the sections below, we will study the nature and stability of the stationary points of the dynamics for $\gamma \leq \delta$ and show that despite the presence of pattern interference, the system is able to retrieve all stored patterns in parallel, although symmetric solutions will no

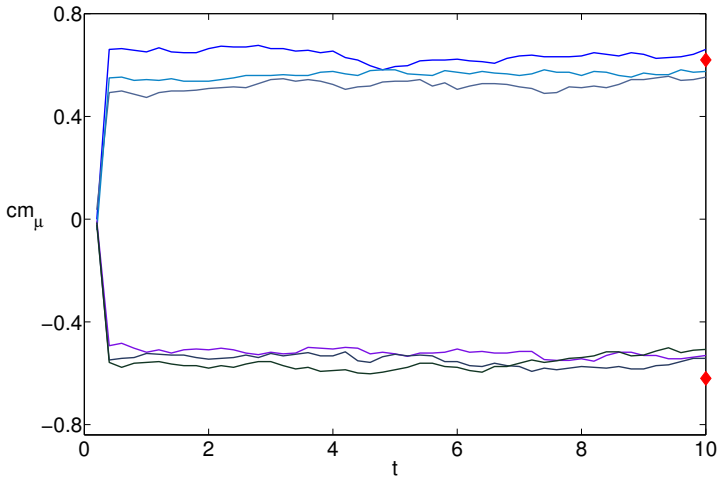

Figure 2.1: Results of Monte Carlo simulations with $N = 10^4$ spins at $c = 0.8$, $\hat{T} = T/c = 0.6$, $\gamma = 0.3$ and $\delta = 0.25$. We plot the overlaps m_μ with different patterns $\boldsymbol{\xi}^\mu$, drawn from (2.2.23), as a function of time t. Red symbols represent the magnetisations at the steady state predicted by (2.2.43). Deviations from the predicted values are small and compatible with finite size effects $\mathcal{O}(N^{-\gamma})$.

longer be stable at all temperatures below criticality.

We will analyse, separately, the dynamics of the system for $\delta = 0$, i.e. the so-called *low storage regime* (sec. 2.3), where the number of stored patterns is finite, and for $\delta > 0$, i.e. the so-called *medium storage regime* (sec. 2.4), where the number of stored patterns is sub-linear in the system size.

2.2.4 Linear stability analysis

In order to determine the regions where different classes of solutions are stable we can inspect the eigenvalues of the Jacobian of the dynamical system (2.2.24). By definition, the Jacobian matrix is given by

$$\mathcal{J}_{\mu\nu} = \left.\frac{\partial F_\mu^{(1)}(\mathbf{m})}{\partial m_\nu}\right|_{\mathbf{m}=\mathbf{m}^\star}, \quad F_\mu^{(1)}(\mathbf{m}) = \frac{N^\gamma}{c}\left\langle \xi^\mu \tanh\left(\hat{\beta}\sum_{\nu=1}^{P}\xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} - m_\mu \ , \tag{2.2.44}$$

evaluated at the steady state $\mathbf{m}^\star$. When the system recalls a number n of patterns, we have, up to permutations of pattern indices, $\mathbf{m}^\star = (m_1, \ldots, m_n, 0 \ldots, 0)$ with $m_\mu \neq 0 \; \forall \; \mu \leq n$, and the resulting matrix is

$$\mathcal{J}_{\mu\nu} = \frac{N^\gamma}{c}\hat{\beta}\left\langle \xi^\mu \xi^\nu \left(1 - \tanh^2\left(\hat{\beta}\mathbf{m}\cdot\boldsymbol{\xi}\right)\right)\right\rangle_{\boldsymbol{\xi}} - \delta_{\mu\nu} \; . \tag{2.2.45}$$

$\mathcal{J}_{\mu\nu}$ is a block matrix, whose diagonal elements are

$$\mathcal{J}_{\mu\mu} = \hat{\beta}(1 - r_\mu) - 1, \quad \text{for } \mu \leq n \; , \tag{2.2.46}$$

$$\mathcal{J}_{\mu\mu} = \hat{\beta}\,(1 - r) - 1, \quad \text{for } \mu > n \; , \tag{2.2.47}$$

with

$$r_\mu = \left\langle \tanh^2\left(\hat{\beta}\left(m_\mu + \sum_{\nu\neq\mu}^{n}\xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} \; , \tag{2.2.48}$$

$$r = \left\langle \tanh^2\left(\hat{\beta}\sum_{\nu=1}^{n}\xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} \; . \tag{2.2.49}$$

Off-diagonal elements are

$$\mathcal{J}_{\mu\nu} = -\hat{\beta}\frac{c}{N^\gamma}R \; , \quad \text{for } \mu,\nu \leq n \; , \tag{2.2.50}$$

$$\mathcal{J}_{\mu\nu} = 0 \; , \quad \text{otherwise} \; , \tag{2.2.51}$$

with

$$R = \left\langle \xi^\mu \xi^\nu \tanh^2\left(\hat{\beta}\mathbf{m}\cdot\boldsymbol{\xi}\right)\right\rangle_{\boldsymbol{\xi}} \; . \tag{2.2.52}$$

Note that the matrix becomes diagonal for $N \to \infty$ as off-diagonal terms $\mathcal{J}_{\mu\nu} \to 0$. Exploiting the block structure of the matrix, we can easily compute its eigenvalues

$$\begin{array}{rl} \text{Eigenspaces} & \text{Eigenvalues} \\ \mathbf{x} = (1, \dots, 1, 0, \dots, 0) & \lambda_1 = \hat{\beta}(1 - r_\mu) - 1 - \hat{\beta}(n-1)\dfrac{c}{N^\gamma} R \ , \\ \mathbf{x} = (x_1, \dots, x_n, 0, \dots, 0), \quad \sum_\mu x_\mu = 0 & \lambda_2 = \hat{\beta}(1 - r_\mu) - 1 + \hat{\beta}\dfrac{c}{N^\gamma} R \ , \\ \mathbf{x} = (0, \dots, 0, x_{n+1}, \dots, x_P) & \lambda_3 = \hat{\beta}(1 - r) - 1 \ . \end{array} \tag{2.2.53}$$

Fixed points of the dynamical systems will be stable in the region of the parameter space (T, c) where the largest eigenvalue of the Jacobian is negative, i.e. all eigenvalues are negative.

2.3 Low Storage and Finite Dilution: $\gamma = \delta = 0$

In the next sections we analyse the regime $\delta \geq \gamma$, where, as explained in sec. 2.2.3, a cross-talk between patterns appears.

2.3.1 A toy model: $P = 2$

Let us first illustrate the main features of the dynamics at low storage (finite number of patterns, $\delta = 0$) and with finite dilution ($\gamma = 0$), by considering the simple toy model with $P = 2$ patterns. In this case, the dynamical equations (2.2.24) reduce to

$$\begin{aligned} \frac{\mathrm{d}m_1}{\mathrm{d}t} &= (1-c)\tanh\left(\hat{\beta}m_1\right) + \frac{c}{2}\left[\tanh\left(\hat{\beta}(m_1+m_2)\right) + \tanh(\hat{\beta}(m_1-m_2))\right] - m_1 \ , \\ \frac{\mathrm{d}m_2}{\mathrm{d}t} &= (1-c)\tanh\left(\hat{\beta}m_2\right) + \frac{c}{2}\left[\tanh\left(\hat{\beta}(m_1+m_2)\right) - \tanh\left(\hat{\beta}(m_1-m_2)\right)\right] - m_2 \ . \end{aligned} \tag{2.3.54}$$

Note that in this case the patterns distribution in (2.2.23) becomes $\mathbb{P}(\boldsymbol{\xi}) = \prod_\mu \left[\frac{c}{2}\left(\delta_{\xi^\mu,+1} + \delta_{\xi^\mu,-1}\right) + (1-c)\delta_{\xi^\mu,0}\right]$. Close to criticality ($\hat{\beta} \simeq 1$), we have from (2.2.33), $m_1 = m_2 = m$, with amplitudes

$$m^2 = \frac{3\epsilon - 3c\mathbf{m}^2}{1 - 3c} \tag{2.3.55}$$

and $\mathbf{m}^2 = 2m^2$, yielding

$$m^2 = \frac{3\epsilon}{1 + 3c} \ . \tag{2.3.56}$$

We analyse the stability of the symmetric mixtures studying the eigenvalues of the Jacobian (2.2.44) evaluated at the fixed point $\mathbf{m}^\star = m(1,1)$. Thanks to the simple form of the matrix, eigenvalues are easily found as

$$\lambda_1 = \hat{\beta} - 1 - (1-c)\hat{\beta}\tanh^2\left(\hat{\beta}m\right) - c\hat{\beta}\tanh^2\left(2\hat{\beta}m\right) \ , \tag{2.3.57}$$

$$\lambda_2 = \hat{\beta} - 1 - (1-c)\hat{\beta}\tanh^2\left(\hat{\beta}m\right) \ . \tag{2.3.58}$$

We note that $\lambda_1 < \lambda_2$, hence the stability is given by the region where the largest eigenvalue λ_2 is negative. Analytically we can also study the behaviour of λ_2 near $T \simeq T_c$ and $T \simeq 0$.

Near the critical temperature, we expand (2.3.58) in powers of $\hat{\beta} = c\beta = 1 + \epsilon$ and using (2.3.56) we find

$$\lambda_2 = -\frac{2\epsilon(1-3c)}{1+3c} \ . \tag{2.3.59}$$

Hence, for $T \simeq T_c$ solutions are stable when $c < 1/3$.

In the opposite limit, $T \to 0$, i.e. $\beta \to \infty$, we have $\tanh^2(\hat{\beta}m) \to 1$, and

$$\lambda_2 \sim \hat{\beta}c - 1 > 0 \ . \tag{2.3.60}$$

Hence, symmetric solutions are unstable at low temperature for any value of

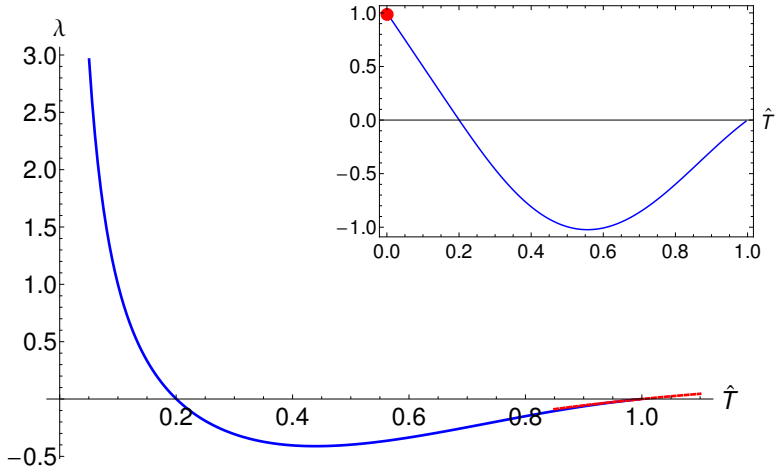

Figure 2.2: Eigenvalue λ_2 as a function of the temperature $\hat{T} = T/c$, for a fixed $c = 0.2$. The red dashed line gives the theoretical prediction near the critical temperature (2.3.59). The figure in the inset shows agreement with (2.3.60), which gives $\lambda_2/(\hat{T}c) \to 1$ as $\hat{T} \to 0$ (red marker).

$c > \sqrt{T}$. The full temperature-dependence of λ_2 can be obtained numerically and is shown in figure 2.2 for dilution $c = 0.2$. The limits for temperature close to zero and to the critical point are in agreement with the theoretical predictions (2.3.59) and (2.3.60), shown in the graph and in the graph inset as the dashed line and the symbol, respectively.

We highlight three regions: a paramagnetic region (**P**) for $T > T_c = c$ where the only solution is $\mathbf{m} = \mathbf{0}$; a region (**S**) for $T < c$ and for $c < 1/3$ where symmetric solutions are stable; finally a region (**H**) where $\mathbf{m} \neq \mathbf{0}$ and symmetric mixtures are unstable, so that a new class of asymmetric solutions takes over, the so-called *hierarchical solutions*. For $T \to 0$ these take the form

$$\mathbf{m} = (1, 1 - c) \ , \tag{2.3.61}$$

which is the steady state of the system (2.3.54). Calculating the eigenvalues λ_1, λ_2 of the stability matrix (2.2.44) evaluated at $\mathbf{m}^\star = (1, 1-c)$, we have $\lambda_1, \lambda_2 \to -1$ as $\beta \to \infty$, so $\mathbf{m} = (1, 1 - c)$ is an attractor at low temperature. This type of retrieval prioritises the activation of one particular B clone, e. g. b_1, with respect to the other.

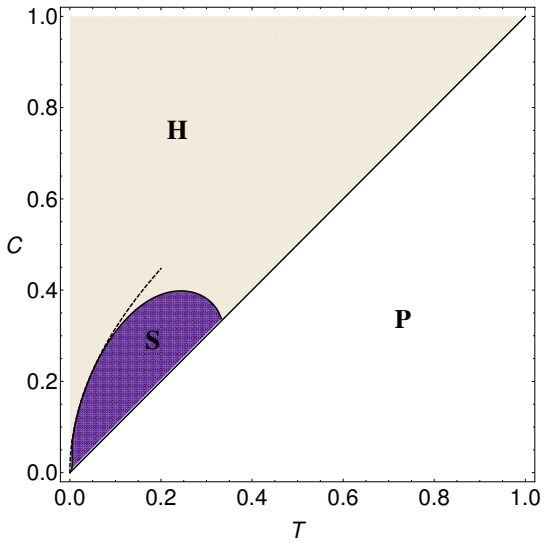

Figure 2.3: Phase diagram in the parameter space (T, c) for the case $\delta = \gamma = 0$ $(P = 2)$. The (**S**) area represents the region where the symmetric mixtures are stable; the (**H**) region is characterised by hierarchical states, while the paramagnetic state (**P**) is stable for $T > T_c = c$. The (**S**) region is obtained as the contour plot of the equation $\lambda_2 = 0$ solved numerically together with equations (2.3.54) at stationarity. The approach to zero is consistent with $c = \sqrt{T}$ predicted by the theory (dashed line).

This behaviour is confirmed by Monte Carlo simulations and by the phase portrait of the dynamical system (2.3.54), both shown in fig. 2.4. One can see that phase curves for $\hat{T} = T/c = 1.25$, i.e. $T > T_c$, evolve to the steady state $\mathbf{m} = \mathbf{0}$; phase curves at temperature $\hat{T} = 0.8$ and $c = 0.25$, (inside the (**S**) region), evolve to symmetric mixtures (where $m_1 = m_2$); at smaller temperature, $\hat{T} = 0.01$, and $c = 0.5$ (inside the (**H**) region), symmetric states become unstable with phase velocities pointing away from them, and non-symmetric stable steady states (with $m_1 \neq m_2$) appear. At $\hat{T} = 0.01$, the system evolves towards the low temperature hierarchical state $\mathbf{m} = (1, 1 - c)$. Results are in agreement with equilibrium analysis and simulations [24, 36, 40].

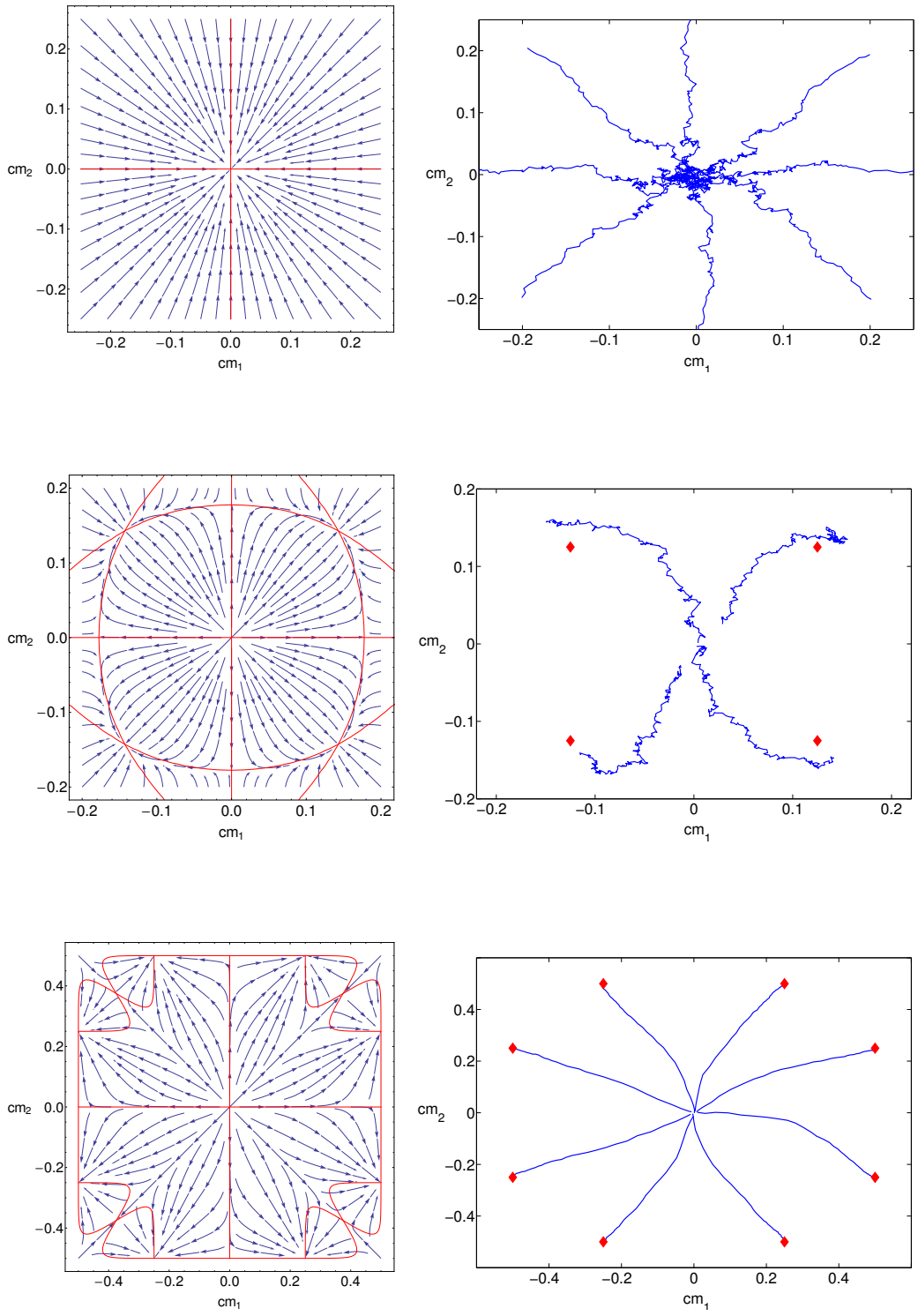

Figure 2.4: Left panels: Phase portraits of the dynamical system (2.3.54) in different regimes of noise T and dilution c for $P = 2$ patterns. Red lines represent the null-clines and stationary states are at the intersections of null-clines. Right panels: Monte Carlo simulations with $N = 10^4$ spins, with T and c as in the left panels; red markers represent the stationary states of the dynamical system. From top to bottom: $c = 0.4, \hat{T} = 1.25$; $c = 0.25, \hat{T} = 0.8$; $c = 0.5$, $\hat{T} = 0.01$: in the latter regime, stationary states are given by $\mathbf{m} = (1, 1 - c)$, as expected at low temperature.

2.3.2 Generalisation to $P > 2$ patterns

Generalising to P patterns, and assuming we have, close to criticality, $\mathbf{m} = m(1, \ldots, 1, 0, \ldots, 0)$ with $\mathbf{m}^2 = nm^2$, we get from (2.2.33)

$$m^2 = \frac{3\epsilon}{1 - 3c + 3cn} \ , \tag{2.3.62}$$

which coincides with (2.3.56) for $P = n = 2$. To study the stability of symmetric mixtures, we evaluate the Jacobian (2.2.44) at the steady state $\mathbf{m}^\star = m(1, \ldots, 1, 0, \ldots, 0)$. Diagonal elements are given by (2.2.46), (2.2.47) where, for symmetric states,

$$r_\mu = \left\langle \tanh^2 \left(\hat{\beta} m \left(1 + \sum_{\nu \neq \mu}^{n} \xi^\nu \right) \right) \right\rangle_{\boldsymbol{\xi}} \ , \tag{2.3.63}$$

$$r = \left\langle \tanh^2 \left(\hat{\beta} m \sum_{\nu=1}^{n} \xi^\nu \right) \right\rangle_{\boldsymbol{\xi}} \ . \tag{2.3.64}$$

In contrast, off-diagonal elements are, for $\gamma = 0$,

$$\mathcal{J}_{\mu\nu} = -\hat{\beta} c R \ , \qquad \text{for } \mu, \nu \leq n \ , \tag{2.3.65}$$

$$\mathcal{J}_{\mu\nu} = 0 \ , \qquad \text{otherwise} \ , \tag{2.3.66}$$

with

$$R = \left\langle \xi^\mu \xi^\nu \tanh^2 \left(\hat{\beta} m \sum_{\nu=1}^{n} \xi^\nu \right) \right\rangle_{\boldsymbol{\xi}} \ . \tag{2.3.67}$$

Due to the symmetry of the problem, we have three distinct eigenspaces, with $r_\mu = r_1 \ \forall \ \mu$:

$$\lambda_1 = \hat{\beta}(1 - r_1) - 1 - \hat{\beta} c (n-1) R \ , \tag{2.3.68}$$

$$\lambda_2 = \hat{\beta}(1 - r_1) - 1 + \hat{\beta} c R \ , \tag{2.3.69}$$

$$\lambda_3 = \hat{\beta}(1 - r) - 1 \ . \tag{2.3.70}$$

Near $T \simeq T_c$ we Taylor expand r_1, r, R in powers of $\epsilon = \hat{\beta} - 1$, obtaining respectively:

$$r_1 = \frac{3\epsilon + 3(n-1)c\epsilon}{1 - 3c + 3cn} \ , \tag{2.3.71}$$

$$r = \frac{3nc\epsilon}{1 - 3c + 3cn} \ , \tag{2.3.72}$$

$$R = \frac{6\epsilon}{1 - 3c + 3cn} \ . \tag{2.3.73}$$

Hence, eigenvalues near the critical temperature are given by

$$\lambda_1 = -\frac{-2\epsilon + 6c\epsilon - 6c\epsilon n}{1 - 3c + 3cn} \ , \tag{2.3.74}$$

$$\lambda_2 = \frac{-2\epsilon + 6c\epsilon}{1 - 3c + 3cn} \ , \tag{2.3.75}$$

$$\lambda_3 = \frac{\epsilon - 3c\epsilon}{1 - 3c + 3cn} \ . \tag{2.3.76}$$

We note that $\lambda_1 < 0$ for any c, and, as in the case $P = 2$, $\lambda_2 < 0$ for $c < 1/3$. On the other hand, $\lambda_3 < 0$ for $c > 1/3$. Since the eigenvalue λ_3 only comes into play for $P > n$, symmetric mixtures $\mathbf{m} = m(1, \dots, 1)$, where *all* patterns are recalled, i.e. with $n = P$, are stable for any $c < 1/3$ near criticality. In contrast, there is no region of the phase space where mixtures of the form $\mathbf{m} = m(1, \dots, 1, 0, \dots, 0)$, in which a number of patterns is not recalled, is stable. Results are confirmed by Monte Carlo simulations shown in fig. 2.5. Interestingly, the stability threshold $c < 1/3$ near T_c is independent of the number of patterns P. However, we note that for a general temperature T the eigenvalues and their stability will in general depend on the number of patterns involved [40].

2.3.2.1 Low-T instability

At low temperature, symmetric mixtures become unstable. In order to show this, it is convenient to introduce a discrete noise distribution. In analogy with

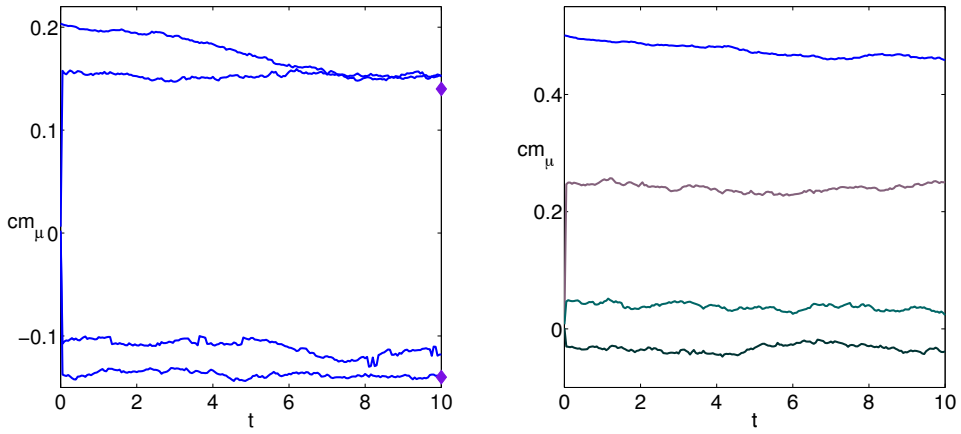

Figure 2.5: Simulations with $N = 10^4$ spins and $P = 4$ patterns of m_μ as a function of time. Left Panel: Symmetric solutions for $\hat{T} = 0.5$, $c = 0.2$. The markers represent the amplitude of the symmetric mixtures evaluated solving the self-consistency equation (2.3.80) for $P = 4$. Right Panel: hierarchical retrieval for $c = 0.5$ and $\hat{T} = 0.4$.

what we did in sec. 2.2.3, we consider

$$\begin{aligned} \tilde{P}_n(z) &= \left\langle \delta_{z,\sum_{\nu=1}^n \xi^\nu} \right\rangle_{\boldsymbol{\xi}} = \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega\hat{z}} \prod_{\nu=1}^{n} \langle \mathrm{e}^{\mathrm{i}\omega\xi^\nu} \rangle_{\boldsymbol{\xi}} \\ &= \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega z} \left[1 + c(\cos\omega - 1)\right]^n \ , \end{aligned} \tag{2.3.77}$$

where n is the number of retrieved patterns. The discrete noise distribution is correctly normalised, as one can check by summing over $z \in \{-n, \ldots, +n\}$, i.e. $\sum_{z=-n}^{+n} \tilde{P}_n(z) = 1$. We can express r_1 and r, as given in (2.2.48), (2.2.49), in terms of (2.3.77):

$$r_1 = \sum_z \tilde{P}_{n-1}(z) \tanh^2\left(\hat{\beta} m \,(1+z)\right) \ , \tag{2.3.78}$$

$$r = \sum_z \tilde{P}_n(z) \tanh^2\left(\hat{\beta} m z\right) \ , \tag{2.3.79}$$

where z runs over the support of $\tilde{P}_n(z)$ and $\tilde{P}_{n-1}(z)$ respectively. The self-consistency equation for the amplitude m of the symmetric mixtures (2.2.34)

becomes

$$m = \sum_z \tilde{P}_{n-1}(z) \tanh\left(\hat{\beta} m (1+z)\right) \ . \tag{2.3.80}$$

It is easy to see that $\tilde{P}_n(z)$ satisfies a useful recursion relation. Using the definition of $\tilde{P}_n(z)$ in(2.3.77), one has

$$\begin{aligned}
\tilde{P}_n(z) &= \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega z} \left[1 + c(\cos\omega - 1)\right] \left[1 + c(\cos\omega - 1)\right]^{n-1} \\
&= (1-c) \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega z} \left[1 + c(\cos\omega - 1)\right]^{n-1} + \frac{c}{2} \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega(z-1)} \left[1 + c(\cos\omega - 1)\right]^{n-1} + \\
&+ \frac{c}{2} \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega(z+1)} \left[1 + c(\cos\omega - 1)\right]^{n-1} \ .
\end{aligned} \tag{2.3.81}$$

Hence, this yields

$$\tilde{P}_n(z) = (1-c)\tilde{P}_{n-1}(z) + \frac{c}{2}\left(\tilde{P}_{n-1}(z+1) + \tilde{P}_{n-1}(z-1)\right) \ , \tag{2.3.82}$$

which represents a discrete-time lazy, symmetric random walker, taking right and left unit steps along the z-axis with equal probabilities $c/2$. All the $\tilde{P}_n(z)$'s in the hierarchy can then be determined from the knowledge of $\tilde{P}_0(z) \equiv \delta(z)$. The random walker is lazy for $c < 1$ because at each iteration n it has a finite probability to take no step.

When $\beta \to \infty$, $r_1 \to 1 - \tilde{P}_{n-1}(-1)$, $r \to 1 - \tilde{P}_n(0)$ and $R \to 0$, hence

$$\lambda_{1,2} \to \hat{\beta}\tilde{P}_{n-1}(1) - 1 \ , \tag{2.3.83}$$

$$\lambda_3 \to \hat{\beta}\tilde{P}_n(0) - 1 \ , \tag{2.3.84}$$

where we used $\tilde{P}_n(z) = \tilde{P}_n(-z)$. We note that for $c < 1$, we have a non-zero probability $\tilde{P}_n(0) > 0$ for the walker to be at the origin for any finite time $n \geq 0$, hence symmetric mixtures are unstable for any value of n at low temperature. In contrast, $c = 1$ introduces a periodicity in the return times of the walker, so that $\tilde{P}_n(0) = 0$ for n odd and $\tilde{P}_n(1) = 0$ for n even. This means that for n odd

(even) *all* eigenvalues are negative (positive), in agreement with the standard (i.e. undiluted) Hopfield model [29], where, at zero temperature, all odd mixtures are attractors and all even mixtures are unstable. In particular, for $n = 1$, $\lambda_{1,2} = -1$ and from (2.3.82) we have

$$\tilde{P}_1(z) = (1-c)\delta(z) + \frac{c}{2}\left(\delta(z+1) + \delta(z-1)\right) \ , \tag{2.3.85}$$

so $\lambda_3 = \hat{\beta}(1-c) - 1$. Hence, the so-called *pure* state, where the system recalls only one pattern, is stable at low temperature only for the undiluted case $c = 1$, but is unstable for any $c < 1$.

In [36,40] an *ansatz* was made for the form of the magnetisation vector in the hierarchical state at $T = 0$,

$$m_\mu = (1-c)^{\mu-1} \ , \quad \mu = 1, \ldots, n \ , \tag{2.3.86}$$

where $n \leq P$ is the number of condensed patterns and $m_\mu = 0 \ \forall \ \mu > n$. We shall denote briefly this state as $\mathbf{m}_H$. We note that (2.3.86), with $n = P$, are indeed fixed points of (2.3.54) for $\beta \to \infty$ and $P = 2$, and for larger P, they solve self-consistently (2.2.25) for $\beta \to \infty$ for wide regions of the dilution parameter c. The stability of this solution is given by the eigenvalues of the stability matrix (2.2.44) evaluated at $\mathbf{m}^\star = \mathbf{m}_H$. The matrix is again diagonal in the limit $N \to \infty$, with elements given in (2.2.46), (2.2.47) and

$$r_\mu = \left\langle \tanh^2\left(\hat{\beta}\left(m_\mu + \sum_{\nu\neq\mu}^{n} \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} \ , \tag{2.3.87}$$

$$r = \left\langle \tanh^2\left(\hat{\beta}\sum_{\nu=1}^{n} \xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} \ . \tag{2.3.88}$$

Upon introducing as in section 2.2.3, $P_\mu(z|\{m_\nu\}) = \left\langle \delta\left(z - \sum_{\nu\neq\mu}^{n} \xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}}$ and

$P(z|\mathbf{m}) = \langle \delta\left(z - \sum_{\nu=1}^{n} \xi^{\nu} m_{\nu}\right) \rangle_{\boldsymbol{\xi}}$, we can rewrite the expressions as

$$r_{\mu} = \int \mathrm{d}z \tanh^2\left(\hat{\beta}(m_{\mu} + z)\right) P_{\mu}(z|\{m_{\nu}\}) \ , \tag{2.3.89}$$

$$r = \int \mathrm{d}z \tanh^2\left(\hat{\beta} z\right) P(z|\mathbf{m}) \ . \tag{2.3.90}$$

For $\beta \to \infty$, the eigenvalues become

$$\lambda_{\mu} = \hat{\beta} P_{\mu}(-m_{\mu}|\{m_{\nu}\}) - 1 \ , \quad \text{for } \mu \leq n \ , \tag{2.3.91}$$

$$\lambda_{\mu} = \hat{\beta} P(0|\mathbf{m}) - 1 \ , \quad \text{for } \mu > n \ . \tag{2.3.92}$$

For $\gamma = 0$ and n condensed patterns, $P_{\mu}(z|\{m_{\nu}\})$ and $P(z|\mathbf{m})$ respectively read

$$P_{\mu}(z|\{m_{\nu}\}) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega z} \prod_{\nu \neq \mu}^{n} \langle \mathrm{e}^{\mathrm{i}\omega \xi^{\nu} m_{\nu}} \rangle_{\boldsymbol{\xi}} = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega z} \prod_{\nu \neq \mu}^{n} [1 + c(\cos(\omega m_{\nu}) - 1)] \tag{2.3.93}$$

and

$$P(z|\mathbf{m}) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} \, \mathrm{e}^{-\mathrm{i}\omega z} \prod_{\nu=1}^{n} [1 + c(\cos(\omega m_{\nu}) - 1)] \ . \tag{2.3.94}$$

Similarly to what was shown in sec. 2.3.2.1 (see eq. (2.3.81) for details), we can relate $P(z|\mathbf{m})$ and $P_{\mu}(z|\{m_{\nu}\})$, $\forall\ \mu$, as follows

$$P(z|\mathbf{m}) = (1-c) P_{\mu}(z|\{m_{\nu}\}) + c \frac{P_{\mu}(z + m_{\mu}|\{m_{\nu}\}) + P_{\mu}(z - m_{\mu}|\{m_{\nu}\})}{2} \ . \tag{2.3.95}$$

It is worth stressing that, differently from the case $\gamma \neq 0$ analysed in sec. 2.2.3, here for $\gamma = 0$ the distribution $P_{\mu}(z|\{m_{\nu}\})$ does *not* converge to $P(z|\mathbf{m})$ in the thermodynamic limit. Setting $z = 0$ in (2.3.95) to check the stability condition, as required by (2.3.92), we have

$$P(0|\mathbf{m}) = (1 - c) P_{\mu}(0|\{m_{\nu}\}) + c P_{\mu}(m_{\mu}|\{m_{\nu}\}) \ . \tag{2.3.96}$$

The stability along the directions $\mu = n + 1, \ldots, P$ requires $P(0|\mathbf{m}) = 0$. It

also implies the stability along the directions $\mu = 1 \dots, n$, i.e. $P_\mu(m_\mu|\{m_\nu\}) = P_\mu(0|\{m_\nu\}) = 0 \ \forall \ \mu$ as for $0 < c < 1$ the two terms on the RHS are both non-negative. However, $P(0|\mathbf{m}) = \langle \delta \left(\sum_{\nu=1}^{n} \xi^\nu m_\nu \right) \rangle_{\boldsymbol{\xi}}$ is bounded from below by the probability of drawing $\boldsymbol{\xi} = \mathbf{0}$, i.e. $P(0|\mathbf{m}) \geq (1-c)^n > 0$, so hierarchical mixtures can only be stable for $P = n$ and the stability condition is given by $\lambda_\mu < 0$ in (2.3.91). The condition for their stability is $P_\mu(m_\mu|\{m_\nu\}) = 0 \ \forall \ \mu \leq P$.

We can check this condition for the simple case with $P = 2$, where we also derived the form of the hierarchical ansatz (2.3.61). In particular, we have to compute $P_1(m_1 = 1|m_2 = 1-c)$ and $P_2(m_2 = 1-c|m_1 = 1)$ from (2.3.93). They respectively read

$$P_1(1|m_2) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega} [1 + c(\cos(\omega(1-c)) - 1)] \ , \tag{2.3.97}$$

$$P_2(1-c|m_1) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega(1-c)} [1 + c(\cos(\omega) - 1)] \ . \tag{2.3.98}$$

With simple manipulations we obtain

$$P_1(1|m_2) = c[\delta(c) + \delta(2-c)]/2 \ , \tag{2.3.99}$$

$$P_2(1-c|m_1) = (1-c)\delta(1-c) + c[\delta(c) + \delta(2-c)]/2 \ , \tag{2.3.100}$$

both vanishing for any $c \in (0, 1)$. This shows that the hierarchical mixtures, of the form $\mathbf{m} = (1, 1-c)$, are stable at low temperature, since their eigenvalues $\lambda_1 = \hat{\beta} P_1(1|m_2) - 1$ and $\lambda_2 = \hat{\beta} P_2(1-c|m_1) - 1$ are negative. Note that in sec. 2.3.1 we were able to check only when symmetric mixtures destabilise, but we did not analyse directly the stability of hierarchical mixtures.

2.4 Extreme dilution and medium storage regime

In the following sections, we analyse the dynamical behaviour of the system in the medium load regime (i.e. $\delta > 0$), where the number of patterns scales sub-linearly in the system size, when patterns are extremely diluted (i.e. $\gamma > 0$).

We will analyse the retrieval properties and obtain the phase diagram for the stability of *symmetric* or *hierarchical* solutions.

2.4.1 Cross-talk effect for $\delta = \gamma$

In this section we study the case $\delta = \gamma$. We have shown in sec. 2.2.1 that the first class of stationary states to bifurcate away from $\mathbf{m} = \mathbf{0}$ below criticality, are the symmetric mixtures $\mathbf{m} = m(1, \dots, 1, 0, \dots, 0)$, with $\mathbf{m}^2 = nm^2$, where n is the number of retrieved components.

For $\delta = \gamma > 0$, we assume that n is a fraction of the total number of patterns $P = \alpha N^\gamma$, hence we set $n = \phi N^\gamma$, with $0 \leq \phi \leq \alpha$. Using $\mathbf{m}^2 = \phi N^\gamma m^2$ in (2.2.32), we obtain the amplitude of symmetric mixtures near the critical temperature $\hat{\beta}_c = 1 + \epsilon$

$$m^2 \simeq \frac{3\epsilon}{1 + 3c\phi} \,. \tag{2.4.101}$$

We analyse the stability of symmetric mixtures by looking at the eigenvalues of the Jacobian $\mathcal{J}_{\mu\nu}$ (2.2.44) evaluated at $\mathbf{m} = m(1, \dots, 1, 0, \dots, 0)$. Its elements are given by (2.2.46), (2.2.47) and (2.2.51), where, for symmetric states

$$r_\mu = \left\langle \tanh^2 \left(\hat{\beta} m \left(1 + \sum_{\nu \neq \mu}^{\phi N^\gamma} \xi^\nu \right) \right) \right\rangle_{\boldsymbol{\xi}} \,, \tag{2.4.102}$$

$$r = \left\langle \tanh^2 \left(\hat{\beta} m \sum_{\nu=1}^{\phi N^\gamma} \xi^\nu \right) \right\rangle_{\boldsymbol{\xi}} \,. \tag{2.4.103}$$

Note that in this case $r_\mu = r_1 \; \forall \; \mu$. As before, it is convenient to consider the discrete distribution[3] $\tilde{P}_\phi(z) = \left\langle \delta_{z, \sum_{\nu=1}^{\phi N^\gamma} \xi^\nu} \right\rangle_{\boldsymbol{\xi}}$. In the thermodynamic limit it can

[3] In principle one should introduce a second distribution $Q_\mu(z) = \left\langle \delta_{z, \sum_{\nu \neq \mu}^{\phi N^\gamma} \xi^\nu} \right\rangle_{\boldsymbol{\xi}}$ but it can be easily shown that the two differ by an amount that is negligible, $\mathcal{O}(N^{-\gamma})$.

be written as

$$\tilde{P}_\phi(z) = \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega z + c\phi(\cos\omega - 1)} = \mathrm{e}^{-\phi c} I_z(\phi c) \ , \tag{2.4.104}$$

where $I_z(\phi c)$ is the modified Bessel function of the first kind [45]. Clearly, $\tilde{P}_\phi(z) = \tilde{P}_\phi(-z)$, and due to the property of the Bessel functions,

$$\frac{\mathrm{d}I_z(\phi c)}{\mathrm{d}(\phi c)} = \frac{1}{2}\left[I_{z-1}(\phi c) + I_{z+1}(\phi c)\right] \ , \tag{2.4.105}$$

$\tilde{P}_\phi(z)$ evolves according to

$$\frac{\mathrm{d}\tilde{P}_\phi(z)}{\mathrm{d}\phi} = \frac{c}{2}\left[\tilde{P}_\phi(z-1) + \tilde{P}_\phi(z+1) - 2\tilde{P}_\phi(z)\right] \ , \tag{2.4.106}$$

which describes a continuous-time symmetric random walker taking unit steps along the z-axis with rate $c/2$, as ϕ increases. We can rewrite r_1 and r in terms of $\tilde{P}_\phi(z)$, as:

$$r_1 = \sum_z \tilde{P}_\phi(z) \tanh^2\left(\hat{\beta} m(1+z)\right) \ , \tag{2.4.107}$$

$$r = \sum_z \tilde{P}_\phi(z) \tanh^2\left(\hat{\beta} m z\right) \ . \tag{2.4.108}$$

In the thermodynamic limit, $N \to \infty$, $\mathcal{J}_{\mu\nu} \to 0$ for $\mu \neq \nu$ and the stability matrix is diagonal, with two eigenvalues

$$\lambda_1 \simeq \hat{\beta}(1 - r_1) - 1 \ , \quad \text{with} \quad deg(\lambda_1) = \phi N^\gamma \ , \tag{2.4.109}$$

$$\lambda_2 \simeq \hat{\beta}(1 - r) - 1 \ , \quad \text{with} \quad deg(\lambda_2) = N^\gamma(\alpha - \phi) \ , \tag{2.4.110}$$

where $deg(\lambda_i)$ represents the degeneracy of eigenvalue λ_i. Near $T \simeq T_c = c$ we Taylor-expand r_1 and r in powers of $\epsilon = \hat{\beta} - 1$

$$r_1 \simeq \frac{3\epsilon(1 + c\phi)}{1 + 3c\phi} \ , \qquad r \simeq \frac{3\epsilon c\phi}{1 + 3c\phi} \ , \tag{2.4.111}$$

and compute the eigenvalues

$$\lambda_1 \simeq -\frac{2\epsilon}{1+3c\phi} < 0 \ , \qquad \lambda_2 \simeq \frac{\epsilon}{1+3c\phi} > 0 \ . \tag{2.4.112}$$

Near $T \simeq 0$, we have

$$r_1 \to 1 - \tilde{P}_\phi(-1) \ , \qquad \lambda_1 \to \hat{\beta} \mathrm{e}^{-\phi c} I_1(\phi c) - 1 \ , \tag{2.4.113}$$

$$r \to 1 - \tilde{P}_\phi(0) \ , \qquad \lambda_2 \to \hat{\beta} \mathrm{e}^{-\phi c} I_0(\phi c) - 1 \ , \tag{2.4.114}$$

where the asymptotic holds for large $\hat{\beta}$. The functions $I_0(x), I_1(x)$ are non-negative and $\mathcal{O}(1)$ for any finite $x > 0$, and become small only for $x \to \infty$, as $\mathrm{e}^{-x} I_n(x) \sim 1/\sqrt{2\pi x}$ $\forall$ n for large x. However, for $\delta = \gamma$, $\phi \leq \alpha$, hence $\lambda_1, \lambda_2 > 0$ and symmetric mixtures are unstable for any finite value of α, at low temperature.

From the analysis above it follows that $\lambda_2 > 0$ near $T \simeq 0$ and $T \simeq T_c$. Indeed, one can check numerically that λ_2 stays positive $\forall$ $T < c$, for any value of ϕ, by computing it from (2.4.108) and (2.4.110), with m solving the self-consistency equation

$$m = \sum_z \tilde{P}_\phi(z) \tanh\left(\hat{\beta} m (1+z)\right) \ . \tag{2.4.115}$$

Since λ_2 only comes into play for $\alpha > \phi$, symmetric mixtures bifurcating from $\mathbf{m} = \mathbf{0}$ must be in the form $\mathbf{m} = m_\alpha(1, \ldots, 1)$, i.e. with $\phi = \alpha$, indicating parallel retrieval of all patterns.

Their stability is controlled by λ_1. In fig. 2.6 we plot the eigenvalue λ_1 as a function of the scaled temperature $\hat{T} = T/c$ for a fixed value $\hat{\alpha} = \alpha c$, as well as the theoretical predictions near T_c (2.4.111) and $T \simeq 0$ (2.4.113). In fig. 2.7 we show the critical line in the parameter space $(\hat{T}, \hat{\alpha})$ separating the region (**S**) where

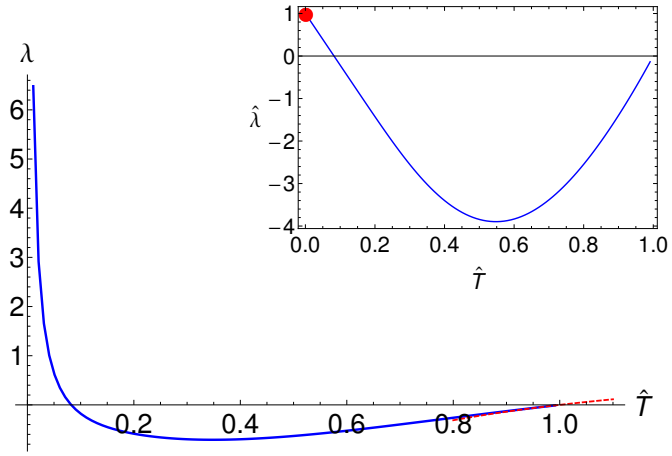

Figure 2.6: Eigenvalue λ_1 as a function of the temperature $\hat{T}$, for a fixed $\hat{\alpha} = 0.2$. The red dashed line gives the theoretical prediction near the critical temperature (eq. (2.4.111)). The figure in the inset is in agreement with eq. (2.4.113), which holds for $T \to 0$ and gives $\hat{\lambda} = \lambda / \mathrm{e}^{-\phi c} I_0(\phi c) \to 1$.

symmetric mixtures are stable, from the region (**H**) where they are unstable. The phase diagram shows that for large regions of the tunable parameters (noise and storage load) the system is able to recall all the stored patterns in parallel and symmetrically. However, for low temperature, the symmetry of patterns is broken and should not be assumed at zero temperature [26]. We suggest in the next subsection the hierarchical structure of the steady state at zero temperature. The dotted line in fig. 2.7 is given by the $\alpha \to 0$ limit of eq. (2.4.113) for $\phi = \alpha$

$$\lambda_1 \simeq \hat{\beta}(1 - \alpha c)\frac{\alpha c}{2} - 1 \ , \tag{2.4.116}$$

where we used $\mathrm{e}^{-\alpha c} I_1(\alpha c) \simeq (1 - \alpha c)\frac{\alpha c}{2} + \mathcal{O}(\alpha^3)$. The stability condition for symmetric mixtures for small α becomes

$$\hat{T} > \frac{\hat{\alpha}}{2} \ , \tag{2.4.117}$$

with $\hat{T} = \hat{\beta}^{-1} = T/c$ and $\hat{\alpha} = \alpha c$.

In fig. 2.8 (left panel) we show Monte Carlo simulations of a system with $N = 10^4$ spins, evolving according to a sequential Glauber dynamics with $\delta =$

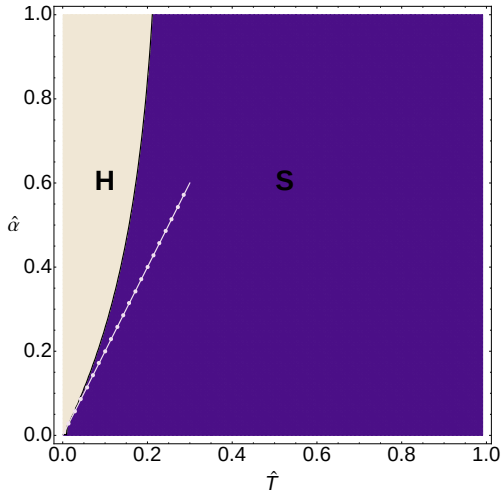

Figure 2.7: Phase diagram in the regime $\delta = \gamma$. In the parameter space $(\hat{T}, \hat{\alpha})$ the purple area represents the region where symmetric mixtures exist and are stable, which correspond to $\lambda_1 < 0$. The area denoted by (**H**) represents the region where we expect a hierarchical retrieval. The dotted line is given by the small α behaviour of λ_1 in eq. (2.4.117).

$\gamma = 0.25$, $\hat{T} = 0.75$, $c = 0.8$ and $\alpha = 1$. This choice of the parameters corresponds to the region in the phase space where symmetric mixtures are stable. Overlaps with different patterns evolve to symmetric, non-zero values in agreement, up to finite-size effects, with the values predicted by the theory (2.4.115).

In conclusion, pattern cross-talk appears in this regime as noise affecting the random-walk distribution (2.4.106), which has the effect of shrinking the amplitude of symmetric mixtures with respect to the regime without cross-talk. There $\tilde{P}_\phi(z)$ crosses over to a delta function centered in $z = 0$ and amplitudes are given by the Curie-Weiss equation (2.2.43).

Finally, let us briefly digress into the properties of the network analysed here. It is, indeed, interesting to compute its percolation threshold for $\delta = \gamma = 1$. In network theory we have a giant component in a graph $\mathcal{G}$ when the average number of second nearest neighbours is larger than the average number of first nearest

neighbours [46]. In this case, this condition is translated as follows

$$(\alpha c^2 N^{1-\gamma})^2 > \alpha c^2 N^{1-\gamma} \ , \tag{2.4.118}$$

as we have $P = \alpha N^\gamma$ patterns with probability $\frac{c}{N^\gamma}$ of having a non-zero link with N T clones. The condition reads

$$\alpha c^2 > N^{\gamma-1} \ . \tag{2.4.119}$$

In the regime $\gamma < 1$, this condition is always satisfied for $\alpha > 0$, hence the network is always above the percolation threshold for any value α. In contrast, for $\alpha = 0$, which corresponds to the case $\delta < \gamma$, the system is below the percolation threshold and, as discussed in sec. 2.2.3, it separates into a set of independent Curie-Weiss models. In chapter 4, we will instead analyse the case with $\delta = \gamma = 1$ where the percolation threshold is given by $\alpha c^2 = 1$. There, our approach will allow us to predict the structure of states in two cases: (i) at $T = 0$ for αc^2 below the percolation threshold (ii) at $T > 0$ and for values αc^2 below a T-dependent threshold (see sec. 4.6.2).

2.4.1.1 Extending the hierarchical ansatz to the medium storage regime

As explained above, symmetric mixtures are unstable at low temperature. It is not *a priori* clear how the symmetry is broken, however for $T \to 0$ one may expect that the system retrieves patterns in a hierarchical fashion similar to the one found in the low storage regime. Reasoning as in [24, 36, 40], we may assume that the system starts aligning its entries to the non-zero entries of the first pattern, which are $cN^{1-\gamma}$, sparing $N - cN^{1-\gamma}$ entries of $\boldsymbol{\sigma}$ to align with the $cN^{1-\gamma}$ non-zero entries of the second pattern and so forth. Hence, we have the following expressions for the non-normalised overlaps $M_\mu = cN^{1-\gamma} m_\mu$, with m_μ

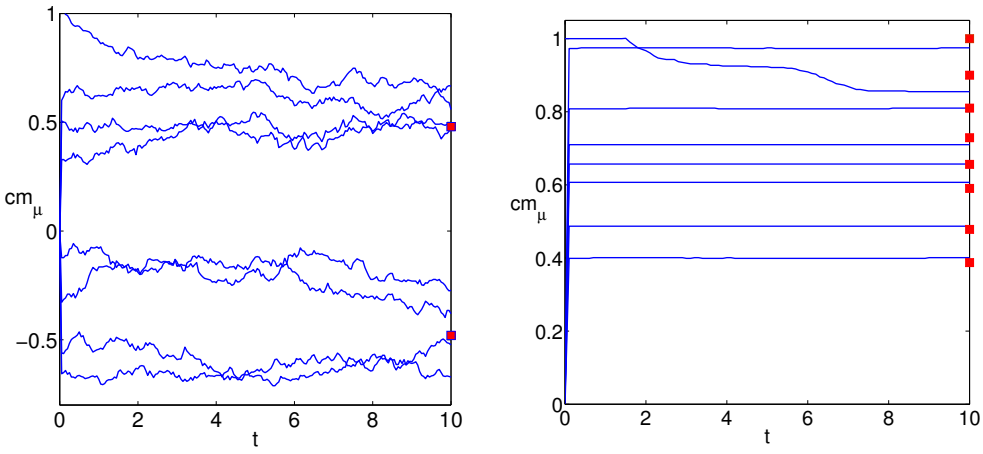

Figure 2.8: Monte Carlo simulations of a system with $N = 10^4$ spins, $\gamma = \delta = 0.25$ and $\alpha = 1$ evolving according to sequential Glauber dynamics. We plot the overlaps m_μ with different patterns $\boldsymbol{\xi}^\mu$ as a function of time t. Left: $\hat{T} = 0.75$, $c = 0.8$. Red markers represent the values theoretically predicted by (2.4.115). Right: $\hat{T} = 0.01$, $c = 1$. Red markers represent the values heuristically predicted by the ansatz (2.4.121) at $T = 0$.

as defined in 2.2.9:

$$\begin{aligned} M_1 &= cN^{1-\gamma} \,, \\ M_2 &= \frac{c}{N^\gamma}\left(N - cN^{1-\gamma}\right) = cN^{1-\gamma}\left(1 - \frac{c}{N^\gamma}\right) \,, \\ M_3 &= \frac{c}{N^\gamma}\left(N - cN^{1-\gamma} - cN^{1-\gamma} + c^2 N^{1-2\gamma}\right) = cN^{1-\gamma}\left(1 - \frac{c}{N^\gamma}\right)^2 \,. \\ &\dots \end{aligned} \tag{2.4.120}$$

Proceeding iteratively, we arrive at the following heuristic general expression for the normalised overlaps

$$m_\mu = \left(1 - \frac{c}{N^\gamma}\right)^{\mu-1} \,, \quad \mu = 1, \dots, \phi N^\gamma \,, \tag{2.4.121}$$

whose stability at $T = 0$ is given by the eigenvalues

$$\begin{aligned} \lambda_\mu &= \hat{\beta}\tilde{P}(m_\mu) \,, \quad \mu = 1, \dots, \phi N^\gamma \,, \\ \lambda_\nu &= \hat{\beta}\tilde{P}(0) \,, \quad \nu = \phi N^\gamma + 1, \dots, \alpha N^\gamma \,, \end{aligned} \tag{2.4.122}$$

where the distribution $\tilde{P}(z)$ is defined in (2.2.41). As in sec. (2.3.2), the probability $\tilde{P}(0)$ is bounded by the probability of drawing $\boldsymbol{\xi} = \mathbf{0}$, $\tilde{P}_\phi(0) \geq (1-cN^{-\gamma})^{\phi N^\gamma} \simeq \mathrm{e}^{-c\phi} > 0$, hence a necessary condition to have stability is $\alpha = \phi$. Monte Carlo simulations at small temperature support the validity of the ansatz (2.4.121) for $\gamma > 0$, with $\phi = \alpha$ (see fig. 2.8, right panel). Remarkably, the first pattern of the hierarchy is recalled without errors, besides the presence of cross-talk.

2.4.2 Strong interference for $\delta > \gamma$

As the number of patterns whose dilution is insufficient is increased, the parallel retrieval capabilities of the system may be naively expected to get drastically compromised. In this section, we will actually show that the patterns recall degrades only very slowly, highlighting the regions where symmetric retrieval is stable. Here we consider $P = \alpha N^\delta$ patterns, with $\delta > \gamma$, and we assume the number of retrieved patterns to be $n = \psi N^\delta$.

As in sec. 2.4.1, we analyse the stability of symmetric solutions of the form $\mathbf{m} = m(1, \dots, 1, 0, \dots, 0)$. The amplitude of symmetric mixtures follows from (2.4.115) as

$$m = \sum_z \tilde{P}_\psi(z) \tanh\left(\hat{\beta} m (1+z)\right) \ , \tag{2.4.123}$$

where $\tilde{P}_\psi(z) = \left\langle \delta_{z, \sum_{\nu=1}^{\psi N^\delta} \xi^\nu} \right\rangle_{\boldsymbol{\xi}}$ is the discrete noise distribution. Using the integral representation of the Kronecker-δ and averaging over $\boldsymbol{\xi}$, the latter can be written as

$$\tilde{P}_\psi(z) = \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega z + c\psi N^{\delta-\gamma}(\cos\omega - 1)} \ . \tag{2.4.124}$$

As for large N the integral above has most of its mass concentrated around $\omega = 0$,

we can use a small-ω expansion in the exponent, yielding

$$\tilde{P}_{\psi}(z) \;=\; \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{-\mathrm{i}\omega z - \frac{c\psi N^{\delta-\gamma}}{2}\omega^2} \;. \tag{2.4.125}$$

Setting $\omega \to \omega/\sqrt{c\psi N^{\delta-\gamma}}$ and extending the rescaled boundaries to infinity, due to the fast decay of the exponential at large values of ω, by Gaussian integration we obtain

$$\tilde{P}_{\psi}(z) = \frac{\mathrm{e}^{-z^2/2c\psi N^{\delta-\gamma}}}{\sqrt{2\pi c\psi N^{\delta-\gamma}}} \;. \tag{2.4.126}$$

Hence, in this regime the cross-talk between a large number of insufficiently diluted condensed patterns appears as a Gaussian noise. The variance of the Gaussian noise is given by $\hat{\psi} = c\psi N^{\delta-\gamma}$ and quantifies the interference level between patterns. The effect of interference on the symmetric mixtures is stronger compared to the $\delta = \gamma$ regime, analysed in sec. 2.4.1, leading to smaller amplitudes of the symmetric recall.

Analysing the statics of this system in the very same regime, an ansatz was put forward [26] that symmetric mixtures were stable in all regions of the phase space $(\hat{T}, \hat{\psi})$. Our analysis based on a dynamical approach highlights the limitations of the ansatz and its regime of validity. In particular, we anticipate that in the limit of large interference, $\hat{\psi} \gg 1$, symmetric mixtures become unstable and a hierarchical recall is, instead, accomplished.

The linear stability analysis of the symmetric mixtures $\mathbf{m} = m(1, \dots, 1, 0 \dots, 0)$ with $n = \psi N^{\delta}$ can be performed as in Sec 2.4.1, by identifying $\phi = \psi N^{\delta-\gamma}$. In particular, we have

$$\lambda_1 \simeq \hat{\beta}(1 - r_1) - 1 \;, \quad \text{with} \quad deg(\lambda_1) = \psi N^{\delta} \;, \tag{2.4.127}$$

$$\lambda_2 \simeq \hat{\beta}(1 - r) - 1 \;, \quad \text{with} \quad deg(\lambda_2) = N^{\delta}(\alpha - \psi) \;, \tag{2.4.128}$$

with

$$r_1 = \sum_z \tilde{P}_\psi(z) \tanh^2\Big(\hat{\beta} m(1+z)\Big) \ , \tag{2.4.129}$$

$$r = \sum_z \tilde{P}_\psi(z) \tanh^2\Big(\hat{\beta} m z\Big) \ . \tag{2.4.130}$$

For $T \simeq T_c$, we get from equations (2.4.129) and (2.4.130),

$$r_1 \simeq \frac{3\epsilon c\psi N^{\delta-\gamma}}{1+3c\psi N^{\delta-\gamma}} \simeq \epsilon\Big(1+\frac{2}{3c\psi N^{\delta-\gamma}}\Big) \ , \tag{2.4.131}$$

$$r \simeq \frac{3\epsilon c\psi N^{\delta-\gamma}}{1+3\epsilon c\psi N^{\delta-\gamma}} \simeq \epsilon\Big(1-\frac{1}{3c\psi N^{\delta-\gamma}}\Big) \ , \tag{2.4.132}$$

where we performed a series expansion for large N. From eq. (2.4.127), (2.4.128) and the expansion above, we find

$$\lambda_1 \simeq \frac{-2\epsilon}{3c\psi N^{\delta-\gamma}} \ , \quad deg(\lambda_1) = \psi N^\delta \ , \tag{2.4.133}$$

$$\lambda_2 \simeq \frac{\epsilon}{3c\psi N^{\delta-\gamma}} \ , \quad deg(\lambda_2) = N^\delta(\alpha-\psi) \ . \tag{2.4.134}$$

For any large but finite N, symmetric mixtures near T_c are stable only if *all* components are recalled, as $\lambda_1 < 0$ and $\lambda_2 > 0$. In this case we would have $\alpha = \psi$, i.e. $P = n$, and the positive eigenvalue accounting for the stability of non-retrieved patterns would disappear. In the thermodynamic limit, eq. (2.4.133)(2.4.134) vanish, and stability cannot be assessed to linear order in ϵ. In the low temperature limit, $T \to 0$, we have, similarly to (2.4.113, 2.4.114),

$$r_1 \to 1 - \tilde{P}_\psi(-1) \ , \qquad \lambda_1 \to \hat{\beta}\tilde{P}_\psi(1) - 1 \ , \tag{2.4.135}$$

$$r \to 1 - \tilde{P}_\psi(0) \ , \qquad \lambda_2 \to \hat{\beta}\tilde{P}_\psi(0) - 1 \ . \tag{2.4.136}$$

Note that we used the symmetry $\tilde{P}_\psi(z) = \tilde{P}_\psi(-z)$. For large N, $\tilde{P}_\psi(1) \simeq \tilde{P}_\psi(0) \simeq N^{(\gamma-\delta)/2}/\sqrt{2\pi c\psi}$ and we get

$$\lambda_1 \simeq \lambda_2 \to \frac{\hat{\beta} N^{(\gamma-\delta)/2}}{\sqrt{2\pi c\psi}} \ , \tag{2.4.137}$$

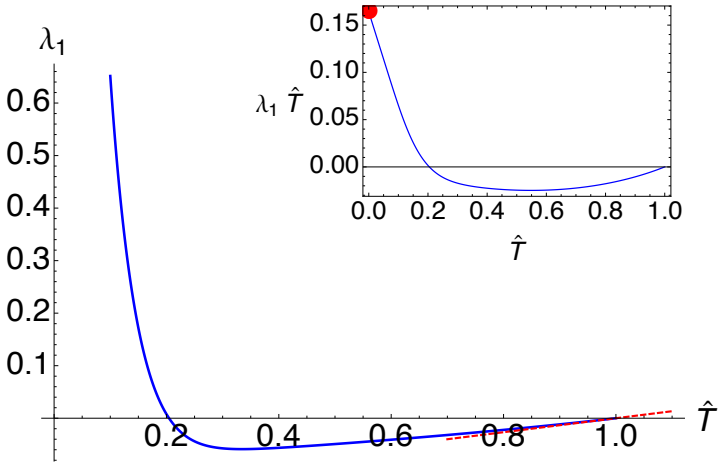

Figure 2.9: Eigenvalue λ_1 as a function of the temperature $\hat{T}$, for a fixed $\hat{\psi} = c\psi N^{\delta-\gamma} = 5$. The red dashed line gives the theoretical prediction near the critical temperature (2.4.133). The figure in the inset shows agreement with (2.4.137), which gives $\lambda_1(\hat{T}) \to \frac{N^{(\gamma-\delta)/2}}{\sqrt{2\pi c\psi}}$ as $\hat{T} \to 0$ (red marker).

hence for any large but finite N, both eigenvalues are positive and symmetric mixtures are unstable, indicating that a hierarchical retrieval must take place. We can compute $\lambda_{1,2}$ numerically from (2.4.127, 2.4.128) by using (2.4.129, 2.4.130) where $\tilde{P}_\phi(z)$ is replaced with (2.4.125) and the amplitude m is determined from (2.4.123). In fig. 2.9 we plot the eigenvalue λ_1 as a function of $\hat{T}$ at a fixed value of $\hat{\psi}$, as well as the predicted limits for $\hat{T} \to 0$ and $\hat{T} \to 1$ given in (2.4.133) and (2.4.137) respectively. We check numerically that λ_2 stays positive for any temperature below criticality. A contour plot of the critical line where $\lambda_1 = 0$ in the $(\hat{T}, \hat{\psi})$ plane is shown in fig. 2.10. In particular, we highlight the region (**S**) where symmetric mixtures are stable and the (**H**) region where they destabilise.

In general, we expect that in presence of strong interference or at low temperature the symmetry among patterns should be broken. However, in this regime, it is more difficult to predict what the shape of the bifurcating solutions is, even at low temperature, due to the presence of strong interference between patterns. In particular, it is not *a priori* clear whether a hierarchical retrieval would involve all patterns or just a fraction of them. Nonetheless, we can inspect the behaviour

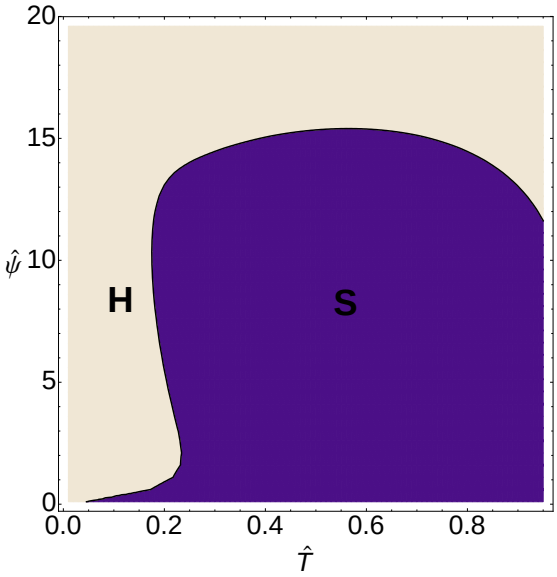

Figure 2.10: Phase diagram in the parameter space $(\hat{T}, \hat{\psi})$ for the case $\delta > \gamma$. The (**S**) area represents the region where the symmetric mixtures are stable, while the (**H**) region is characterised by hierarchical states. The (**S**) region is obtained as the contour plot of the equation $\lambda_1 = 0$ solved numerically together with equation (2.4.123).

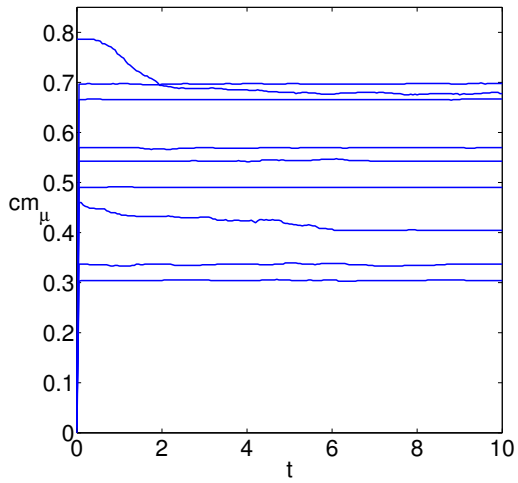

Figure 2.11: Monte Carlo simulations with $N = 10^4$ spins. We plot the overlaps m_μ with different patterns ξ_μ, as a function of time t, for $\hat{T} = 0.015$, $c = 0.8$, $\delta = 0.25$, $\gamma = 0.2$ and $\alpha = 1$.

of the system by looking at the activation dynamics, described by

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \left\langle \tanh\left(\hat{\beta}\left(m_\mu + \sum_{\nu\neq\mu}^{P}\xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} - m_\mu \ . \tag{2.4.138}$$

Introducing the noise density $P(z|\{m_\nu\}) = \left\langle \delta\left(z - \sum_{\nu\neq\mu}^{P}\xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}}$ we have (as in sec. 2.2.3)

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \int_{-\infty}^{\infty}\mathrm{d}z P(z|\{m_\nu\})\tanh\left(\hat{\beta}\left(m_\mu + z\right)\right) - m_\mu \ . \tag{2.4.139}$$

The distribution can be written, using the standard manipulations introduced so far, as

$$P(z|\{m_\nu\}) = \int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{2\pi}\mathrm{e}^{\mathrm{i}z\omega + \frac{c}{N^\gamma}\sum_{\nu\neq\mu}^{P}(\cos(\omega m_\nu)-1)} \ . \tag{2.4.140}$$

At zero temperature, one may expect that the system shows an overlap $\mathcal{O}(1)$ with a non-zero number of patterns. Assuming that the number of these condensed patterns is $\mathcal{O}(N^\gamma)$ and the remaining (non-condensed) $\mathcal{O}(N^\delta)$ patterns have an overlap with the system configuration $\mathcal{O}(N^{\frac{\gamma-1}{2}})$, we can write the noise distribution by splitting the contribution from the condensed (labelled by λ) and the non-condensed (labelled by ρ) patterns

$$\begin{aligned} P(z|\{m_\nu\}) &= \int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{2\pi}\mathrm{e}^{\mathrm{i}z\omega + \frac{c}{N^\gamma}\sum_{\lambda=1}^{\psi N^\gamma}(\cos(\omega m_\lambda)-1)}\mathrm{e}^{\frac{c}{N^\gamma}\sum_{\rho=1}^{\alpha N^\delta}(\cos(\omega m_\rho)-1)} \\ &= \int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{2\pi}\mathrm{e}^{\mathrm{i}z\omega + \frac{c}{N^\gamma}\sum_{\lambda=1}^{\psi N^\gamma}(\cos(\omega m_\lambda)-1)}\mathrm{e}^{-\frac{c}{2N^\gamma}\sum_{\rho=1}^{\alpha N^\delta} m_\rho^2\omega^2} \ . \end{aligned} \tag{2.4.141}$$

In the last step, we also Taylor-expanded the m_ρ-dependent terms for m_ρ small. For large N we obtain

$$P(z|\{m_\nu\}) \simeq P(z|\{m_\lambda\}) + \mathcal{O}(N^{\delta-1}) \ . \tag{2.4.142}$$

Hence, the evolution of the non-condensed patterns is given, to leading order, by

$$\frac{\mathrm{d}m_\rho}{\mathrm{d}t} \simeq \int_{-\infty}^{+\infty} \mathrm{d}z P(z|\{m_\lambda\}) \tanh\left(\hat{\beta}(m_\rho + z)\right) - m_\rho + \mathcal{O}(N^{\delta-1}) \ , \quad (2.4.143)$$

which depends on the noise distribution of condensed patterns $\{m_\lambda\}$ only. Taylor-expanding for small m_ρ we obtain

$$\frac{\mathrm{d}m_\rho}{\mathrm{d}t} \simeq m_\rho \left[\hat{\beta} \int_{-\infty}^{+\infty} \mathrm{d}z P(z|\{m_\lambda\}) \left(1 - \tanh^2\left(\hat{\beta}z\right)\right) - 1\right] + \mathcal{O}(N^{\delta-1}) \quad (2.4.144)$$

and for $T \to 0$ we get

$$\frac{\mathrm{d}m_\rho}{\mathrm{d}t} \simeq m_\rho \left[\hat{\beta} P(0|\{m_\lambda\}) - 1\right] . \quad (2.4.145)$$

As explained in 2.4.1.1, $P(0|\{m_\lambda\}) \geq \mathrm{e}^{-\psi c}$. This suggests that the non-condensed solutions are unstable and the system will attempt to recall hierarchically all patterns, as supported by Monte Carlo simulations shown in fig. 2.11. However, the retrieval degrades with respect to that described in (2.4.121), valid at the onset of pattern cross-talk, where the first pattern in the hierarchy was retrieved without errors.

2.5 Summary

In this chapter, we analysed the dynamics of associative memories with diluted patterns, in different regimes of pattern dilution and storage load (away from saturation). Our aim was to understand the behaviour of the system in terms of signals exchange between clones and cytokine patterns recall.

In particular, we have identified three regimes as highlighted in fig. 2.12: (I) $\delta < \gamma$, (II) $\delta = \gamma$ and (III) $\delta > \gamma$. In (I) there is no interference between cytokine patterns, as these are relatively low in number and sufficiently diluted, so the network behaves as a set of P independent ferromagnets each evolving according to the Curie-Weiss equation. As a result, the parallel retrieval accomplished by

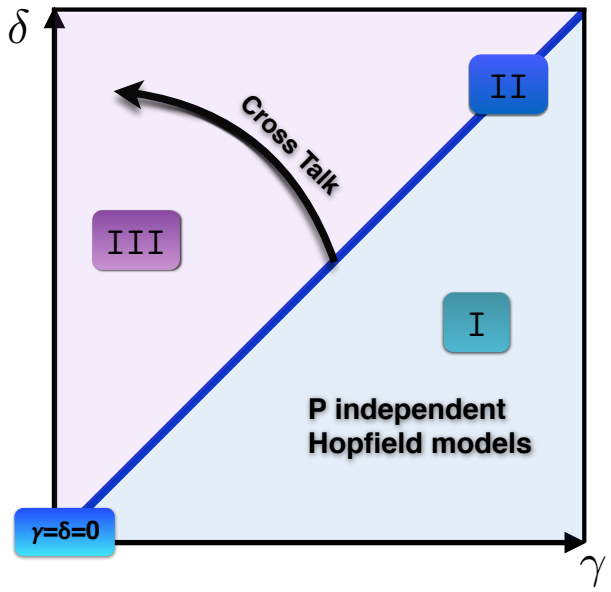

Figure 2.12: Schematic phase space (γ, δ). In region I ($\delta < \gamma$), we have a set of P independent ferromagnets and parallel retrieval is accomplished in a symmetric fashion. In region II cross-talk between patterns appears; the intensity of symmetric recall is decreased due to the presence of a noise, whose distribution is identical to the one of a lazy, unbiased random walker, and the region of stability of the symmetric retrieval is provided in the phase diagram (fig. 2.7). A similar behaviour is found for $\delta = \gamma = 0$, but the region of stability of the symmetric region shrinks (fig. 2.3), due to the noise distribution retaining a pattern dependence (sec. 2.2.3). Finally, in region III, cross-talk effects are strong and increase the larger δ and the smaller γ. These decrease the strength of symmetric retrieval via a Gaussian interference noise, and are seen to degrade gradually the hierarchical retrieval of the network, which thus retains its parallel processing capabilities.

the system in this regime is symmetric for any temperature below criticality.

In (II) a pattern cross-talk appears and symmetric mixtures are no longer stable for all temperatures below criticality. We derived the phase diagram showing their region of stability in the parameter space. The cross-talk between patterns materialises on symmetric mixtures as a noise term affecting the distribution of a lazy random walker, taking discrete steps at discrete time (for $\gamma = 0$) and in continuous time (for $\gamma > 0$). The laziness of the random walker is due to pattern dilution and removes the periodicity in the return times, which is responsible for the different role played by even and odd mixtures, in the traditional (i.e. undiluted) Hopfield model, at zero temperature [29]. The region where symmetric mixtures are stable is broad for $\gamma > 0$, where the noise distribution is pattern-independent, whereas the retrieval is mostly hierarchical for $\gamma = 0$, where the symmetry of patterns is mostly a close-to-criticality effect.

Finally, in (III), the effect of cross-talk on symmetric mixtures turns out to be Gaussian noise, whose variance is directly related to the level of pattern interference, which decreases the quality of parallel retrieval. We derived the region of stability of symmetric mixtures and showed that for large interference or low temperature, symmetric mixtures are unstable. In particular, at zero temperature the system is able to retrieve all patterns in a hierarchical fashion, despite the presence of strong interference may suggest that no retrieval is accomplished in this regime [26].

There is a growing community of theorists who want, as it were, more out of life. [..] We want to reconcile the physicists desire for concise, unifying theoretical principles with the obvious complexity and diversity of life.

W. Bialek, Perspectives on theory at the interface of physics and biology

3

Towards a more realistic immune system model

The immune system is, indeed, a beautiful and powerful machinery, able to perform multiple tasks simultaneously by coordinating the sub-parts that it is composed of. In the previous chapters we described, for example, a simple model that captures the ability of T clones to orchestrate the same-time activation of B clones in the adaptive immune system. Yet, there is more to this description that Nature has been able to create: from ingenious mechanisms to modulate the immune response to smart and efficient ways to discern self-cells from external pathogens and prevent unwanted activations. To this end, we propose here an extended model of the adaptive immune system to incorporate some important biological features of real immune systems. In particular, we will include the features listed below.

(a) *Receptor promiscuity*: we introduce the effect of having B clones with a variable number of receptors expressed on their surfaces, one of the most

important mechanisms preventing autoimmune reactions and diseases. A *receptor editing process* is, indeed, very commonly observed during the B clones maturation, where self-reactive cells, which may be responsible for the onset of autoimmune responses, are suppressed at an early stage of the development by reducing their number of receptors [47].

(b) *Idiotypic interactions*: we consider interactions between B clones, which are not only able to recognise antigens, but also other B cells with complementary receptors. In healthy situations, B-B interactions generate a suppressive "force" that must be overcome in order to induce an immune response and may prevent unwanted activations [48]. According to recent experimental studies [49–52] this so-called *idiotypic network* of interactions seems also to play a central role in autoimmune diseases. In this case, it supports a cascade of unwanted events such as autoantibodies production, which recognise each other and modulate the immune response.

(c) *Antigens*: we incorporate the effect of an external antigenic field acting on the B-T system. In particular, we investigate the *immunological memory* [2], i.e. the ability of the immune system to produce a more effective and faster response at a second encounter with an antigen [1]. We also discuss how the presence of infections may affect the immune system *basal activity* [53], i.e. B clones activation due to the interactions with T clones in the absence of external antigens, and surveillance.

In terms of modelling, we will use the techniques and the setting introduced in chapter 2. In the next sections, we will discuss how we actually integrate the new ingredients in the previous framework and the main results obtained. We will analyse the system's behaviour in different regions of the parameters space, i.e. varying the number of clones and triggered receptors, noise level or B-B interaction strength, through linear stability analysis and Monte Carlo simulations.

3.1 Adding new ingredients

The players involved in this model are, as before, T clones $\sigma_i = \pm 1$, $i = 1, \dots, N$ and B clones $b_\mu \in \mathbb{R}$, $\mu = 1, \dots, P$ with the addition of antigens with concentrations $\psi_a \in \mathbb{R}, a = 1, \dots, N_A$ (with respect to a reference level). Both the number of B clones and the number of antigens are sub-linear in the system size N, with $P = N^\delta$, $\delta \in [0, 1)$ and $N_A = N^\lambda$, $\lambda \in [0, 1)$. In the absence of antigens and interactions with T clones, the B clones' log-concentrations b_μ can be regarded as Gaussian variables ($\mathcal{N}(0, T)$ in absence of idiotypic interactions as shown in chapter 2). The interactions between the i-th T clone and the μ-th B clone, mediated via the cytokines, are represented by the variables $\xi_i^\mu \in \{+1, -1, 0\}$, respectively corresponding to receptor promiscuityexcitatory, inhibitory or absent signals. We will regard those as random variables drawn from $\mathbb{P}(\{\boldsymbol{\xi}\}|\mathbf{q})$, which will depend on the set of parameter $\mathbf{q} = (q_1, \dots, q_P)$ tuning the receptor promiscuity per B clone. Note that by setting $\mathbf{q} = c(1, \dots, 1)$ we would recover the case with fixed receptor promiscuity analysed in chapter 2.

The combined interacting system of B, T clones and antigens can be phenomenologically described by the following Hamiltonian:

$$\mathcal{H}(\boldsymbol{\sigma}, \mathbf{b}|\boldsymbol{\xi}) = -\sum_{\mu=1}^{P} b_\mu \left(\sum_{a=1}^{N_A} \psi_a \eta_a^\mu + N^{1-\gamma} \sum_{i=1}^{N} \xi_i^\mu \sigma_i \right) + \frac{1}{2} \sum_{\mu,\nu=1}^{P} b_\mu A_{\mu\nu} b_\nu \ . \quad (3.1.1)$$

The first term takes into account the interactions between antigens and B clones via the matrix η_a^μ, the second term is related to B - T interactions mediated by cytokines, while the third one accounts for the effect of the idiotypic network, i.e. B-B interactions via the matrix $A_{\mu\nu}$. These interactions are schematically summarised in fig. 3.1, 3.2. Assuming that the B clones log-concentrations evolve

according to a gradient descent on the Hamiltonian (3.1.1), we have

$$\frac{\mathrm{d}b_\mu}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial b_\mu} + \chi_\mu(t) = \sum_{a=1}^{N_A} \psi_a \eta_a^\mu + N^{\gamma-1} \sum_{i=1}^{N} \xi_i^\mu \sigma_i - \frac{1}{2} \sum_{\nu=1}^{P} A_{\mu\nu} b_\nu + \chi_\mu(t) \ , \tag{3.1.2}$$

where $\chi_\mu(t)$ is a Gaussian white noise with $\langle \chi_\mu(t) \rangle = 0$, $\langle \chi_\mu(t) \chi_\nu(t') \rangle = 2T\delta_{\mu,\nu}\delta(t-t')$ and strength $T = 1/\beta$. Hence, the log-concentration of the μ-th B-clone is directly affected by this new interaction mechanisms, i.e. B clone-antigen and B-B, decreasing or increasing depending on the processes and type of interactions (excitatory or inhibitory).

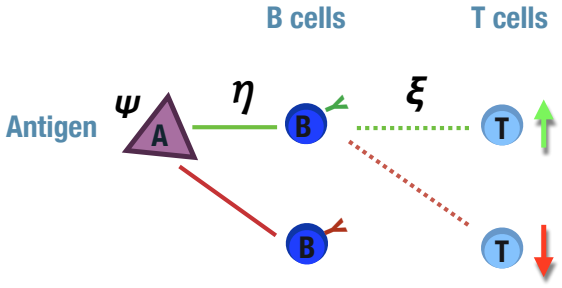

Figure 3.1: Schematic interactions between B, T clones and the antigen A. In the presence of an antigen with concentration ψ, the complementary B clone will detect it (B-A interactions mediated via the matrix $\boldsymbol{\eta}$) and will receive a confirmatory signal from the active T clones (represented by up arrows) via the cytokines ξ_i^μ .

Assuming as in chapter 2 that B clones are faster than T clones and that T clones minimise the same energy function (3.1.1), we can introduce a Glauber sequential dynamics with transition rates

$$w_i(\boldsymbol{\sigma}) = \frac{1}{2}\left[1 - \sigma_i \tanh\left(\beta h_i^{\mathrm{eff}}(\boldsymbol{\sigma})\right)\right] \tag{3.1.3}$$

and effective fields

$$h_i^{\mathrm{eff}}(\boldsymbol{\sigma}) = N^{\gamma-1}\left(\mathbf{M}^T \mathbf{A}^{-1} \boldsymbol{\xi}_i + \boldsymbol{\psi}^T \boldsymbol{\eta} \mathbf{A}^{-1} \boldsymbol{\xi}_i\right) \ . \tag{3.1.4}$$

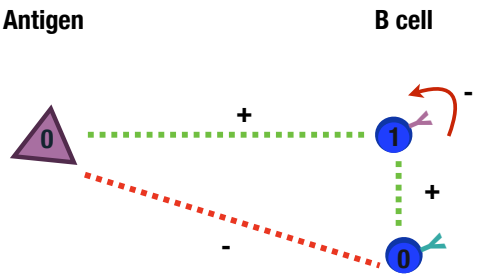

Figure 3.2: Scheme of B-B and B-A interactions. Antigens and B clones are denoted by variables $0, 1$ representing the shape of their receptors. Different B clones excite each other and each of them represses itself, while the antigen will excite complementary B repressing the identical one (0-0).

The dependence on $\boldsymbol{\sigma}$ is through the vector of magnetisations $\mathbf{M}^T = (M_1, \dots, M_P)$, where

$$M_\mu(\boldsymbol{\sigma}) = \frac{1}{N^{1-\gamma}} \sum_{i=1}^{N} \sigma_i \xi_i^\mu \ . \tag{3.1.5}$$

Here $\boldsymbol{\psi}^T = (\psi_1, \dots, \psi_a)$ represents the set of antigenic fields and we denote $\mathbf{M}^T$ and $\boldsymbol{\psi}^T$ the transpose of $\mathbf{M}$ and $\boldsymbol{\psi}$ respectively. The matrix $\boldsymbol{\eta}$ encodes the interaction between antigens and B clones, while $\mathbf{A}$ refers to the B-B interactions. As shown in sec. 2.2.1.1, from the microscopic probabilities $p_t(\boldsymbol{\sigma})$ of finding the system in state $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N)$ we can write the dynamics for the B clones activation (magnetisations), described by the probability of finding the system in a macroscopic state $\mathbf{M}$[1] at time t, namely

$$P_t(\mathbf{M}) = \sum_{\boldsymbol{\sigma}} p_t(\boldsymbol{\sigma}) \delta(\mathbf{M} - \mathbf{M}(\boldsymbol{\sigma})) \ . \tag{3.1.6}$$

Finally, via a Kramers-Moyal expansion for large system size and away from saturation, it is possible to show that $P_t(\mathbf{M})$ evolves according to a Liouville equation, as shown in ch. 2 for the corresponding $P_t(\mathbf{m})$. Hence, it follows that

[1]The attentive reader will have noticed the different normalisation of the overlaps in (3.1.5), compared to ch. 2. For convenience we work here with dilution-dependent overlaps and we will normalise them later, depending on the choice of dilution.

$\mathbf{M}$ evolve deterministically with a dynamics described by

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = \left\langle \boldsymbol{\xi}^T \tanh\left[\beta\left(\mathbf{M}^T\mathbf{A}^{-1}\boldsymbol{\xi} + \boldsymbol{\psi}^T\mathbf{C}\boldsymbol{\xi}\right)\right] \right\rangle_{\boldsymbol{\xi}} - \mathbf{M}\ , \tag{3.1.7}$$

where $\mathbf{C} = \boldsymbol{\eta}\mathbf{A}^{-1}$ and $\langle\cdot\rangle_{\boldsymbol{\xi}}$ denotes the average over the distribution $\mathbb{P}(\boldsymbol{\xi}|\mathbf{q})$. The choice of the distribution will be specified in the following.

In the next sections, we will consider the following cases:

- in sec. 3.2 we study only the effect of receptor promiscuity, in the absence of antigens ($\boldsymbol{\psi} = \mathbf{0}$) and idiotypic interactions ($\mathbf{A} = \mathbf{1}$);
- in sec. 3.3 we consider the effects of idiotypic interactions in the absence of antigens ($\boldsymbol{\psi} = \mathbf{0}$) and with homogeneous promiscuities, i.e. $\mathbf{q} = c(1, \dots, 1)$;
- in sec. 3.4 we analyse the effects of antigens ($\boldsymbol{\psi} \neq \mathbf{0}$) in the absence of idiotypic interactions ($\mathbf{A} = \mathbf{1}$) and with homogeneous promiscuities, i.e. $\mathbf{q} = c(1, \dots, 1)$.

In each of the sections mentioned above, we will consider the low storage regime and finite dilution, i.e. $P = 2$ patterns and $\gamma = 0$, and the medium load with extreme dilution, i.e. $P = \alpha N^\delta$ patterns and $\delta, \gamma > 0$.

3.2 Effects of receptor promiscuity

We first consider the case where there is no antigen $\psi_a = 0\ \forall\ a$, and no B-B interactions i.e. $\mathbf{A} = \mathbf{1}$, and we focus on the effect of having a variable number of triggered receptors on different B clones, i.e. heterogeneous q_μ. The receptor promiscuity is introduced in the model via the cytokines or patterns distribution.

We will regard the cytokines as P random variables drawn from

$$\mathbb{P}(\boldsymbol{\xi}|\mathbf{q}) = \prod_{\mu} \left[\frac{q_\mu}{2N^\gamma} \left(\delta_{\xi^\mu,1} + \delta_{\xi^\mu,-1} \right) + \left(1 - \frac{q_\mu}{N^\gamma} \right) \delta_{\xi^\mu,0} \right], \quad q_\mu = \mathcal{O}(N^0) \quad \forall \mu \ , \tag{3.2.8}$$

where the q_μ's are drawn from a given distribution. The expected number of triggered receptors of the μ-th B-clone is given by $n_r^\mu = \frac{q_\mu}{N^\gamma} N$. All clones have a sub-extensive number of receptors $\mathcal{O}(N^{1-\gamma})$, while q_μ finely tunes the interactions promiscuity. Experiments [54] suggest that the number of receptors in T cells have to be at least ~ 8000 in order to be activated, whereas $N = \mathcal{O}(10^8)$. The fraction of non-zero B-T links determines the degree of dilution of the system: for $\gamma = 0$ the system is finitely diluted, whereas for $\gamma > 0$ the system is extremely diluted.

3.2.1 Dynamical equations for B clones activation

In order to compare the activation of B clones with different numbers of receptors, it is convenient to look at the activation per receptor, given by the normalised order parameters $m_\mu = M_\mu / q_\mu$, $\mu = 1, \dots, P$, which take values in the range $[-1, 1]$ for all clones μ. The dynamical equations (3.1.7) then read

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \frac{N^\gamma}{q_\mu} \left\langle \xi^\mu \tanh \left(\beta \sum_{\nu=1}^{P} q_\nu \xi^\nu m_\nu \right) \right\rangle_{\boldsymbol{\xi}} - m_\mu \ , \tag{3.2.9}$$

where the average is taken over 3.2.8. At the critical temperature $T_c = q_{\max}$, where $q_{\max} = \max_\mu [q_\mu]$, the system undergoes a phase transition, with the equilibrium phase at $T > T_c$ characterised by $\mathbf{m} = \mathbf{0}$, and B clones activation ($\mathbf{m} \neq \mathbf{0}$) occurring for $T < T_c$. At the steady state ($\mathrm{d}\mathbf{m}/\mathrm{d}t = \mathbf{0}$), we have

$$m_\mu = \frac{N^\gamma}{q_\mu} \left\langle \xi^\mu \tanh \left(\beta \sum_{\nu=1}^{P} q_\nu \xi^\nu m_\nu \right) \right\rangle_{\boldsymbol{\xi}} \ . \tag{3.2.10}$$

Taking the scalar product with $\mathbf{m}$ and using the inequality $|\tanh x| \leq |x|$, yields

$$\mathbf{m}^2 \leq \frac{N^\gamma}{q_\mu}\beta \sum_{\mu=1}^{P} m_\mu \sum_{\nu=1}^{P} q_\nu m_\nu \langle \xi^\nu \xi^\mu \rangle_{\boldsymbol{\xi}} = \beta \sum_{\mu=1}^{P} q_\mu m_\mu^2 \leq \beta q_{\max} \mathbf{m}^2 \ . \tag{3.2.11}$$

This implies $\mathbf{m}^2 = \mathbf{0}$, hence $\mathbf{m} = \mathbf{0}$, for $\beta q_{\max} < 1$, meaning that none of the B clones get activated for noise levels above the critical value $T > q_{\max}$. Although $\mathbf{m} = \mathbf{0}$ is a steady state solution of (3.2.9) for any value of T, a linear stability analysis shows that it becomes unstable for $T < q_{\max}$. To this end, we compute the Jacobian of the linearised dynamics about the steady state $\mathbf{m}^\star = \mathbf{0}$

$$\mathcal{J}_{\mu\nu} = \left.\frac{\partial F_\mu^{(1)}(\mathbf{m})}{\partial m_\nu}\right|_{\mathbf{m}=\mathbf{m}^\star}, \qquad F_\mu^{(1)}(\mathbf{m}) = \frac{N^\gamma}{q_\mu}\left\langle \xi^\mu \tanh\left(\beta \sum_{\nu=1}^{P} \xi^\nu q_\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} - m_\mu \ , \tag{3.2.12}$$

which gives

$$\mathcal{J}_{\mu\nu} = \frac{N^\gamma \beta q_\nu}{q_\mu}\left\langle \xi^\nu \xi^\mu \left[1 - \tanh^2\left(\beta \sum_{\nu=1}^{P} \xi^\nu q_\nu m_\nu^\star\right)\right]\right\rangle_{\boldsymbol{\xi}} - \delta_{\mu\nu} \ . \tag{3.2.13}$$

Substituting $\mathbf{m}^\star = \mathbf{0}$ we get $\mathcal{J}_{\mu\nu} = (\beta q_\mu - 1)\delta_{\mu\nu}$. The largest eigenvalue, which gives the stability of $\mathbf{m}^\star = \mathbf{0}$, is $\lambda_{\max} = \beta q_{\max} - 1$. This becomes positive for $\beta q_{\max} > 1$, showing that non-zero solutions $\mathbf{m} \neq \mathbf{0}$ will bifurcate away from $\mathbf{m} = \mathbf{0}$ at $T = q_{\max}$.

In the next sections, we will inspect the structure and the stability of the bifurcating solutions first for the case with two B clones (sec. 3.2.2) and then for the general one with P B clones (sec. 3.2.3). Results in sec. 3.2 will show that cells with few receptors may be transiently activated but fail to sustain the signal even if strongly excitatory (fig. 3.3). The receptor promiscuity affects both the immune response strength and the critical temperature at which they get activated: both decrease with the number of triggered receptors. In addition, competition emerges between B clones to be activated. As the number of activated B clones increases,

signalling pathways to inactive clones get more noisy due to the interference of active clones. Therefore, the critical temperature for the activation of inactive clones is a decreasing function of the fraction of active clones (see eq.(3.2.67)).

3.2.2 A toy model with two B clones

Here we study a toy model with $P = 2$ B clones, and assume $q_1 > q_2$. For $\gamma > 0$, this model reduces to two independent Curie-Weiss ferromagnets, with critical temperatures q_1 and q_2 respectively (see discussion in sec. 3.2.3). Hence, the most interesting case is obtained for $\gamma = 0$. The state $\mathbf{m}^\star = (0, 0)$ is the only steady state of the system for $T > q_1$, but it destabilises for $T < q_1$. To investigate the system's behaviour below criticality, we numerically solve its dynamical equations

$$\begin{aligned}\frac{\mathrm{d}m_1}{\mathrm{d}t} &= (1-q_2)\tanh(\beta q_1 m_1) + \frac{q_2}{2}[\tanh(\beta(q_1 m_1 + q_2 m_2)) + \\ &\quad + \tanh(\beta(q_1 m_1 - q_2 m_2))] - m_1 \ , \qquad (3.2.14)\\ \frac{\mathrm{d}m_2}{\mathrm{d}t} &= (1-q_1)\tanh(\beta q_2 m_2) + \frac{q_1}{2}[\tanh(\beta(q_1 m_1 + q_2 m_2)) + \\ &\quad - \tanh(\beta(q_1 m_1 - q_2 m_2))] - m_2 \ . \qquad (3.2.15)\end{aligned}$$

In fig. 3.3 we show the flow diagram and the stable fixed points of the dynamical system at different temperatures: first, we notice that B clones with a higher promiscuity produce a higher immune response (m_1), whereas a lower promiscuity q_μ results in a lower or null activation (m_2), depending on the temperature. Hence, the number of receptors on B clones surface affects their responsiveness. In particular, if T is high, only clones with the highest number of receptors are activated and the system's fixed point corresponds to the pure state $m_1 \neq 0, m_2 = 0$. Lowering T induces the activation of clones with fewer receptors but with a lower intensity ($m_1 > m_2$).

Theoretical results are consistent with Monte Carlo simulations, shown in fig. 3.4. In addition, Monte Carlo simulations are in qualitative agreement with experimental results [54] showing that clones with very few receptors are triggered

transiently but fail to remain active in the long run. Fig. 3.5 shows that our model can qualitatively reproduce this effect: clones with few receptors (green) have a lower activation even if triggered by a strong signal (initial condition) and tend to be switched off after a short transient, conversely cells with a higher number of receptors (blue) produce a strong immune response, even if triggered by a weak signal (initial condition).

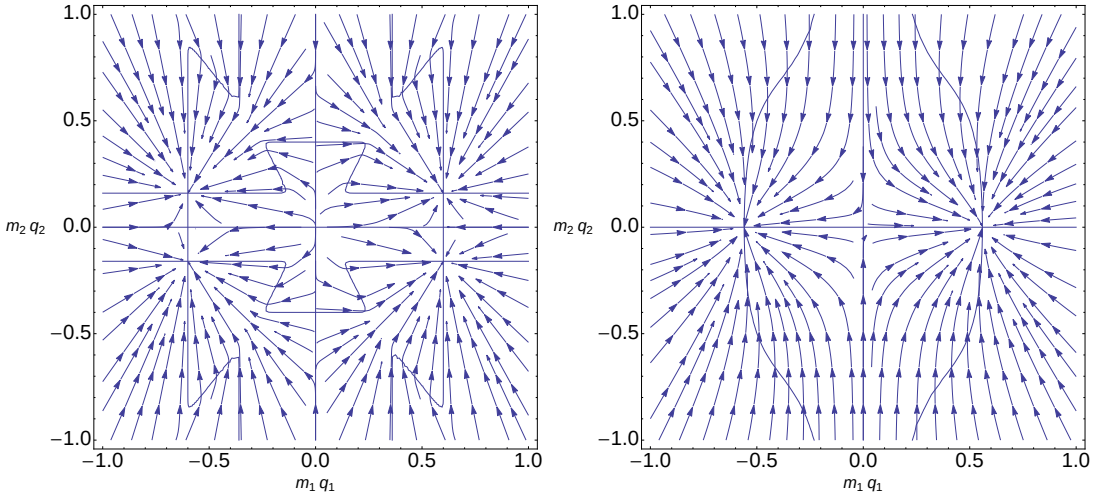

Figure 3.3: Phase portrait of the dynamical system (3.2.14, 3.2.15) for $q_1 = 0.6$, $q_2 = 0.4$ at low temperature $T = 0.01$ (left) and high temperature $T = 0.2$ (right). Blue lines represent null-clines and stationary states are at the intersection of null-clines.

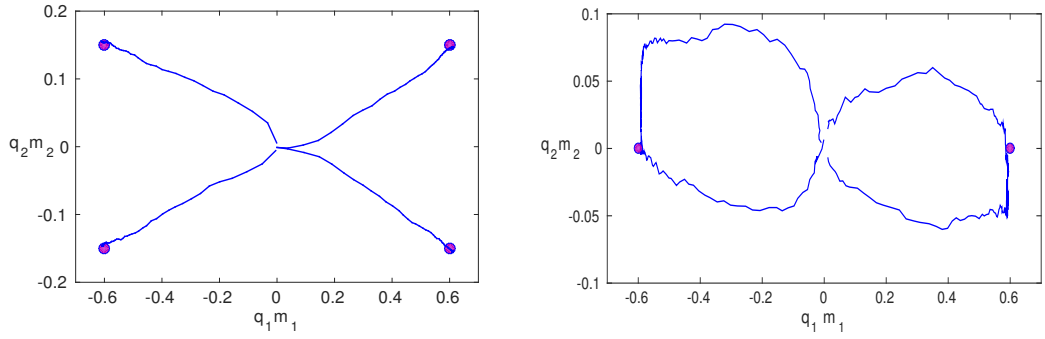

Figure 3.4: Monte Carlo simulations with 10^4 spins for $q_1 = 0.6$, $q_2 = 0.4$ at low temperature $T = 0.01$ (left) and at high temperature $T = 0.2$ (right). Markers represent the numerical solutions of (3.2.14, 3.2.15) .

Analytically, we can investigate the structure of the first states to bifurcate below T_c by expanding the steady state equations, obtained setting $\mathrm{d}m_{1,2}/\mathrm{d}t = 0$

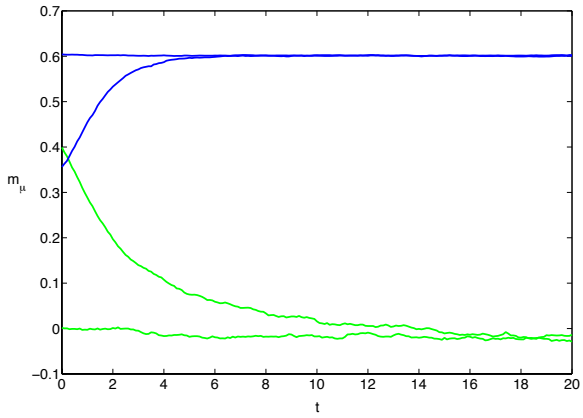

Figure 3.5: Monte Carlo simulations with $N = 10^4$ spins, with $q_1 = 0.6$, $q_2 = 0.4$ at $T = 0.2$. Clone activations m_1, m_2 as a function of time for different initial conditions. Clones with few receptors [q_2 (green)] fail to get activated even if triggered by a strong signal.

in (3.2.14, 3.2.15), for small m_1, m_2, close to criticality, i.e. at $\beta q_1 = 1 + \epsilon$. We get as possible solutions

$$m_1^2 = 3\epsilon - 3\frac{q_2^3}{q_1^2} m_2^2 \,, \tag{3.2.16}$$

$$m_2^2 = 3\frac{q_1^2}{q_2^2}(1+\epsilon) - 3\frac{q_1^3}{q_2^3} - 3\frac{q_1^3}{q_2^2} m_1^2 \,, \tag{3.2.17}$$

as well as $m_1 = 0, m_2 = 0$. Solution (3.2.17) is not physical as it remains $\mathcal{O}(1)$ at $\epsilon = 0$ for $q_1 \neq q_2$, hence $m_2 = 0$. Inserting $m_2 = 0$ in (3.2.16), we get $m_1^2 = 3\epsilon$. Hence, the first state bifurcating away from $\mathbf{m} = (0, 0)$ is in the form $\mathbf{m} = (\sqrt{3\epsilon}, 0)$. At high temperature (below criticality) only the clone with the highest promiscuity is switched on. For $q_1 = q_2$ we clearly retrieve the results presented in chapter 2 and both B clones are activated with the same intensity below criticality.

It is, then, interesting to understand in which region of the phase space clones with lower promiscuities become responsive. Naively one might expect that m_2 becomes active at $T \simeq q_2$. In reality, heterogeneities in the clone promiscuities deeply affect cell responsiveness and cells with fewer receptors will remain quies-

cent even at very low T. By Taylor-expanding (3.2.9) for small m_2 in powers of $\epsilon = q_2\beta - 1$, we obtain

$$m_1 = \tanh(\beta q_1 m_1) \simeq \tanh\left(\frac{q_1}{q_2} m_1\right) , \tag{3.2.18}$$

$$m_2^2 = 3\epsilon - 3\frac{q_1^3}{q_2^2} m_1^2 . \tag{3.2.19}$$

Since m_1 is $\mathcal{O}(1)$, the non-zero solution for m_2 is impossible close to $T = q_2$. The pure state $\mathbf{m} = (m_1, 0)$ will then have a wider stability region, which can be found by analysing the eigenvalues of the Jacobian (3.2.13) at $\mathbf{m}^\star = (m_1, 0)$,

$$\lambda_1 = \beta q_1 - \beta q_1 \tanh^2(\beta q_1 m_1) - 1 , \tag{3.2.20}$$

$$\lambda_2 = \beta q_2 - \beta q_1 q_2 \tanh^2(\beta q_1 m_1) - 1 . \tag{3.2.21}$$

Analytically we can calculate $\lambda_{1,2}$ near $T \simeq 0$

$$\lambda_1 \simeq -1 , \tag{3.2.22}$$

$$\lambda_2 \simeq \beta q_2 - \beta q_1 q_2 - 1 , \tag{3.2.23}$$

and near T_c, i.e. at $\beta q_1 = 1 + \epsilon$

$$\lambda_1 \simeq -2\epsilon , \tag{3.2.24}$$

$$\lambda_2 \simeq \frac{q_2\epsilon + q_2 - 3\epsilon q_1 q_2 - q_1}{q_1} + \mathcal{O}(\epsilon^2) . \tag{3.2.25}$$

For intermediate T we compute $\lambda_{1,2}$ numerically. Plots of the eigenvalues as a function of the temperature are shown in fig. 3.6, where the theoretical predictions for $T \simeq 0$ (3.2.23) and $T \simeq T_c$ (3.2.25) are highlighted. We note that $\lambda_1 < 0, \forall\, T$, hence the stability of the pure state is determined by the sign of λ_2. In fig. 3.7 we show a contour plot of $\lambda_2 = 0$ in the $T - q_2$ plane fixing q_1. The linear behaviour can be understood as follows. In the pure state region we have, using the steady

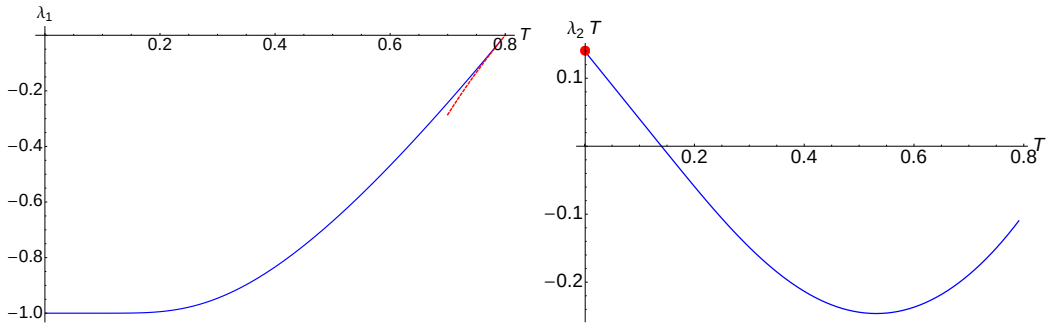

Figure 3.6: Eigenvalues (3.2.20),(3.2.21) as a function of T for $q_1 = 0.8$ and $q_2 = 0.7$. Left: λ_1, the red dashed line represents the behaviour near $T \simeq T_c$ (3.2.25). Right: $\lambda_2 T$, the red marker represents the limit at $T \simeq 0$ (3.2.23).

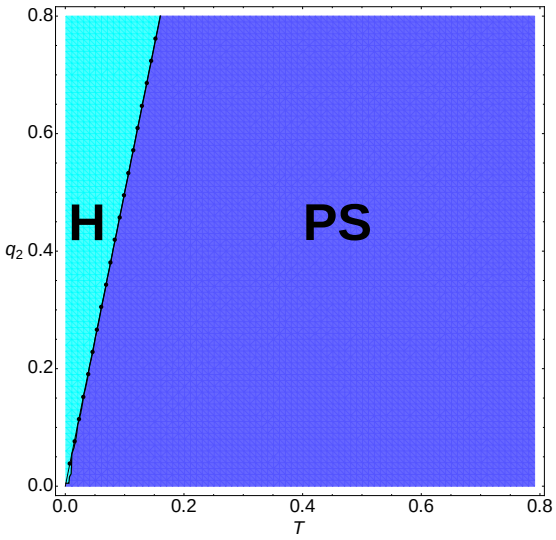

Figure 3.7: Phase diagram in the space (T, q_2) fixing $q_1 = 0.8$ obtained from the condition $\lambda_2 < 0$ (3.2.21). The dotted line represents the theoretical critical temperature (3.2.28). In the (**PS**) region the pure state is stable ($m_1 \neq 0, m_2 = 0$). At low T, B clones are hierarchically activated (**H**).

state equation $m_1 = \tanh(\beta q_1 m_1)$,

$$\lambda_1 = \beta q_1 (1 - m_1^2) - 1 \ , \tag{3.2.26}$$

$$\lambda_2 = \beta q_2 (1 - q_1 m_1^2) - 1 \ . \tag{3.2.27}$$

As T decreases (below q_2) m_1 increases, so that the eigenvalues stay negative and the stability is ensured, until m_1 reaches its maximum value $m_1 = 1$. At this point, a further decrease of the temperature will make λ_2 positive, destabilising

the pure state, and a new state (m_1, m_2) will take over. The temperature at which bifurcations from the pure state are expected is thus found from the condition $\lambda_2(T, m_1 = 1) = 0$, as

$$T = q_2(1 - q_1) \ , \tag{3.2.28}$$

which is in agreement with the critical temperature computed numerically (fig. 3.7). Deviations from the linear behaviour are expected in the regime $q_1 \simeq q_2$, where a symmetric activation occurs for $q_1 \leq 1/3$, as shown in ch. 2. In fig. 3.8 we plot m_1, m_2 as a function of q_2 in the (**PS**) region (left) and crossing the critical line where m_2 becomes non-zero (right). In the latter region, i.e. for $q_2 > T/(1 - q_1)$, the stable state is $\mathbf{m} = (1, m_2)$ where m_2 is the T-dependent stationary solution of (3.2.15) at $m_1 = 1$. In particular, for $T = 0$ the stable state is

$$m_1 = 1 \ , \tag{3.2.29}$$

$$m_2 = 1 - q_1 \ , \tag{3.2.30}$$

in agreement with simulations and flow diagrams (fig. 3.3, 3.4). In conclusion, the system can activate clones with different numbers of receptors simultaneously for $T < q_2(1 - q_1)$. The activation is hierarchical (**H**), with clones with higher promiscuity being prioritised with respect to the others. In particular, clones with the highest number of receptors are activated with the strongest possible signal in a wide region of the phase diagram.

3.2.3 The case of P B clones with a variable promiscuity

In this section, we study the case where the number of B clones is $P = N^\delta$ with $\delta \in [0, 1)$, and N is the number of T-clones. The dynamical equations are

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \frac{N^\gamma}{q_\mu} \left\langle \xi^\mu \tanh\left(\beta \sum_{\nu=1}^{P} q_\nu \xi^\nu m_\nu\right) \right\rangle_{\boldsymbol{\xi}} - m_\mu \ , \tag{3.2.31}$$

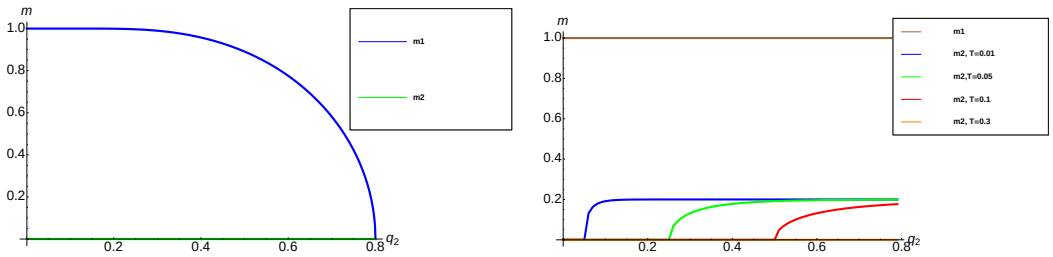

Figure 3.8: Plot of m_1, m_2 as a function of q_2. Left : We consider $q_1 = 0.8$ at $T = q_2$, following the magnetisations in the (**PS**) region with $\mathbf{m} = (m_1, 0)$. Right: For $q_1 = 0.8$, and $T = 0.01, 0.05, 0.1, 0.3$ we show the bifurcation of m_2, while m_1 has already reached its maximum $m_1 = 1$.

which can be rewritten, averaging over non-zero values of ξ^μ, as

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \left\langle \tanh\left(\beta\left(q_\mu m_\mu + \sum_{\nu\neq\mu}^{P} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} - m_\mu \ . \tag{3.2.32}$$

Introducing, as done in sec. 2.2.3, the noise distribution $P_\mu(z|\{m_\nu, q_\nu\}) = \left\langle \delta\left(z - \sum_{\nu\neq\mu}^{P} \xi^\nu q_\nu m_\nu\right)\right\rangle$ on clone μ, we have

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \int_{-\infty}^{+\infty} \mathrm{d}z P_\mu(z|\{m_\nu, q_\nu\}) \tanh\left(\beta(q_\mu m_\mu + z)\right) - m_\mu \ , \tag{3.2.33}$$

where $P_\mu(z|\{m_\nu, q_\nu\})$ can be written, using the Fourier representation of the Dirac delta and carrying out the average over $\boldsymbol{\xi}$, as

$$\begin{aligned} P_\mu(z|\{m_\nu, q_\nu\}) &= \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{\mathrm{i}z\omega} \left\langle \prod_{\nu\neq\mu}^{P} \mathrm{e}^{-\mathrm{i}\omega\xi^\nu q_\nu m_\nu} \right\rangle_{\boldsymbol{\xi}} \\ &= \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{\mathrm{i}z\omega} \mathrm{e}^{\sum_{\nu\neq\mu}^{P} \frac{q_\nu}{N^\gamma}[\cos(\omega q_\nu m_\nu) - 1]} \ . \end{aligned} \tag{3.2.34}$$

If $\gamma > 0$, extending the sum in the exponent to all patterns will add a negligible contribution $\mathcal{O}(N^{-\gamma})$ in the thermodynamic limit, hence, as $N \to \infty$ all clones will have the same noise distribution

$$P_\mu(z|\{m_\nu, q_\nu\}) \to P_P(z|\mathbf{m}, \mathbf{q}) = \left\langle \delta\left(z - \sum_{\nu=1}^{P} q_\nu \xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} \ . \tag{3.2.35}$$

We note that the sum on the rhs of (3.2.35) is at most $\sim \mathcal{O}(N^{\delta-\gamma})$, hence for $\delta < \gamma$ it is negligible in the thermodynamic limit. This yields $P_P(z|\mathbf{m},\mathbf{q}) \to \delta(z)$ as $N \to \infty$, so that equations (3.2.33) decouple and the system reduces, for $\delta < \gamma$, to a set of indipendent Curie-Weiss ferromagnets, each evolving according to

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \tanh(\beta q_\mu m_\mu) - m_\mu \ . \tag{3.2.36}$$

At the steady state, each B clone μ becomes active at its own critical temperature $T_c = q_\mu$, independently of the other clones (fig. 3.9, left). In contrast, for $\delta \geq \gamma$ the noise distribution $P_P(z|\mathbf{m},\mathbf{q})$ has a finite width, due to clonal interference, and the equations for the evolution of clonal activations are coupled. In this regime, clones compete to be activated and the ones with fewer triggered receptors will fail to get switched on (fig. 3.9, right).

In the following section, we will analyse the effect of receptor promiscuity in the regime of competing clones. In particular, we will show that clonal interference affects both critical temperature and intensity of B clones activations.

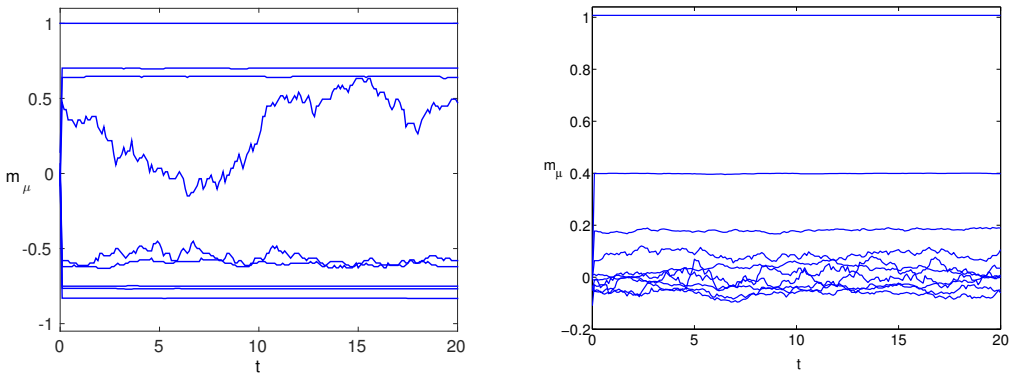

Figure 3.9: Monte Carlo simulations with $N = 10^4$ spins, $q_\mu = (0.7)^\mu q$. $T = 0.01$, $q = 6$. Left: $\delta = 0.2, \gamma = 0.25$. In the non-competing regime ($\delta < \gamma$) *all* clones are activated. Right: $\delta = \gamma = 0.25$. In this case only few clones with the highest promiscuity are active.

3.2.3.1 Bifurcations near the critical temperature and stability region in the regime of competing clones.

In this section we study the bifurcations away from $\mathbf{m} = (0, \dots, 0)$ below the critical temperature $T_c = q_{\max}$. Without loss of generality, we assume $q_{\max} = q_1$. We Taylor expand the steady state equations obtained by setting $\mathrm{d}m_\mu/\mathrm{d}t = 0$ in (3.2.31), for small m_μ at $\beta q_1 = 1 + \epsilon$

$$\begin{aligned} m_\mu &\simeq \frac{N^\gamma}{q_\mu}\beta \sum_{\nu=1}^{P} q_\nu \langle \xi^\mu \xi^\nu \rangle_{\boldsymbol{\xi}} m_\nu - \frac{N^\gamma}{q_\mu}\frac{\beta^3}{3} \sum_{\nu,\rho,\lambda=1}^{P} q_\nu q_\rho q_\lambda \langle \xi^\mu \xi^\nu \xi^\rho \xi^\lambda \rangle_{\boldsymbol{\xi}} m_\nu m_\rho m_\lambda \\ &= \beta q_\mu m_\mu - \frac{\beta^3 q_\mu^3}{3} m_\mu^3 - \beta^3 \frac{q_\mu}{N^\gamma} m_\mu \sum_{\rho \neq \mu}^{P} m_\rho^2 q_\rho^3 \ . \end{aligned} \tag{3.2.37}$$

For any μ, $m_\mu = 0$ is always a solution. Non-zero solutions are given, for $\mu = 1$, by

$$\beta q_1 - 1 - \frac{\beta^3 q_1^3}{3} m_1^2 - \frac{\beta^3}{N^\gamma} q_1 \sum_{\rho>1}^{P} m_\rho^2 q_\rho^3 = 0 \ , \tag{3.2.38}$$

which at $\beta q_1 = 1 + \epsilon$ and for large N yields

$$m_1^2 = 3\epsilon - \frac{3}{q_1^2 N^\gamma} \sum_{\rho>1}^{P} m_\rho^2 q_\rho^3 \ . \tag{3.2.39}$$

For $\mu \neq 1$, we have

$$\beta q_\mu - 1 - \frac{\beta^3 q_\mu^3}{3} m_\mu^2 + \frac{\beta q_\mu^4}{N^\gamma} m_\mu^2 - \frac{\beta^3}{N^\gamma} q_\mu \sum_{\rho>1}^{P} m_\rho^2 q_\rho^3 = 0 \ , \tag{3.2.40}$$

which gives, for $N \to \infty$ and to $\mathcal{O}(\epsilon^0)$,

$$m_\mu^2 = \frac{3}{q_\mu^3}(q_1^2 q_\mu - q_1^3) - \frac{3}{N^\gamma q_\mu^2} \sum_{\rho>1}^{P} m_\rho^2 q_\rho^3 + \mathcal{O}(N^{-\gamma}) \ . \tag{3.2.41}$$

Summing over $\mu > 1$ we get

$$\sum_{\mu>1}^{P} m_\mu^2 q_\mu^3 \left(1 + \frac{3}{N^\gamma} \sum_{\rho>1}^{P} q_\rho\right) = 3\left(\sum_{\mu>1}^{P} (q_\mu - q_1)\right) . \tag{3.2.42}$$

Since $q_\mu < q_1$, $\forall \mu > 1$, this equality can never be satisfied, showing that $m_\mu = 0, \forall \mu > 1$. Substituting this result into (3.2.39), we find $m_1^2 = 3\epsilon$, hence the first state to bifurcate away from $\mathbf{m} = (0, \ldots, 0)$ is $\mathbf{m} = (\sqrt{3\epsilon}, 0, \ldots, 0)$.

The stability region of the pure state $\mathbf{m} = (m_1, 0, \ldots, 0)$ can be found by inspecting the sign of the eigenvalues of the Jacobian of the linearised equations of motion about the steady state. For a steady state with the general structure $\mathbf{m} = (m_1, m_2, \ldots, m_n, 0, \ldots, 0)$, where n is the fraction of activated clones, the Jacobian (3.2.13) has a block structure, where diagonal terms are, for $\mu \leq n$

$$\mathcal{J}_{\mu\mu} = \beta q_\mu \left(1 - \left\langle \tanh^2 \left(\beta \left(q_\mu m_\mu + \sum_{\nu\neq\mu}^{n} \xi^\nu q_\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}}\right) - 1 \tag{3.2.43}$$

and for $\mu > n$

$$\mathcal{J}_{\mu\mu} = \beta q_\mu \left(1 - \left\langle \tanh^2 \left(\beta \sum_{\nu=1}^{n} \xi^\nu q_\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}}\right) - 1 . \tag{3.2.44}$$

Off-diagonal elements are, for $\mu, \nu \leq n$

$$\mathcal{J}_{\mu\nu} = -\frac{\beta q_\mu}{N^\gamma} \left\langle \tanh^2 \left(\beta \sum_{\nu=1}^{n} \xi^\nu q_\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} - 1 \tag{3.2.45}$$

and $\mathcal{J}_{\mu\nu} = 0$ otherwise. For $N \to \infty$ the matrix becomes diagonal with eigenvalues $\lambda_\mu = \mathcal{J}_{\mu\mu}$ given by (3.2.43) for $\mu \leq n$ and (3.2.44) for $\mu > n$. In the pure state $\mathbf{m}^\star = (m_1, 0, \ldots, 0)$, where only one clone is activated, we have

$$\lambda_1 = \beta q_1 \left(1 - \tanh^2(\beta q_1 m_1)\right) - 1 , \tag{3.2.46}$$

$$\lambda_\mu = \beta q_\mu \left(1 - \langle \tanh^2(\beta \xi^1 q_1 m_1)\rangle_{\boldsymbol{\xi}}\right) - 1 , \qquad \mu > 1 . \tag{3.2.47}$$

Near the critical temperature $T_c = q_1$, setting $\beta q_1 = 1 + \epsilon$ and using $m_1^2 \simeq 3\epsilon$ gives

$$\lambda_1 = \beta q_1 - 1 - \beta^3 m_1^2 q_1^3 = -2\epsilon < 0 \ , \tag{3.2.48}$$

$$\lambda_\mu = \beta q_\mu - 1 - \beta^3 q_1^2 q_\mu \langle (\xi^1)^2 m_1^2 \rangle_{\boldsymbol{\xi}} = (1+\epsilon)\frac{q_\mu}{q_1} - 1 \ , \quad \mu > 1 \ , \tag{3.2.49}$$

showing that for $q_\mu < q_1$ the pure state is stable near criticality, as opposed to the case $q_\mu = q_1 = c, \forall \mu$ studied in ch. 2, where all clones are activated with the same intensity below T_c. In the opposite limit $T \to 0$, we get from (3.2.46, 3.2.47)

$$\lambda_1 \simeq -1 \ , \tag{3.2.50}$$

$$\lambda_\mu \simeq \beta q_\mu \left(1 - \frac{q_1}{N^\gamma}\right) - 1 \simeq \beta q_\mu - 1 \ , \quad \mu > 1 \ , \tag{3.2.51}$$

showing that the pure state is unstable at low temperature.

Indeed, decreasing the temperature below q_1, we expect clones with fewer receptors to get active. In particular, at $T = 0$ we expect all clones to be activated, in a hierarchical fashion. Similarly to the case with no promiscuity analysed in sec. 2.4.1.1, the system sends the strongest possible signal to the clone with maximum promiscuity, while the clone with the second highest promiscuity receives signals from the remaining $N - q_1 N^{1-\gamma}$ spare T clones and so on. Considering the overlaps $M_\mu = q_\mu N^{1-\gamma} m_\mu$ we have that

$$\begin{aligned}
M_1 &= q_1 N^{1-\gamma} \\
M_2 &= q_2 N^{1-\gamma}\left(1 - \frac{q_1}{N^\gamma}\right) \\
M_3 &= q_3 N^{1-\gamma}\left(1 - \frac{q_1}{N^\gamma}\right)\left(1 - \frac{q_2}{N^\gamma}\right) \\
&\dots \\
M_n &= q_n N^{1-\gamma}\left(1 - \frac{q_1}{N^\gamma}\right)\left(1 - \frac{q_2}{N^\gamma}\right)\dots\left(1 - \frac{q_{n-1}}{N^\gamma}\right) \ .
\end{aligned} \tag{3.2.52}$$

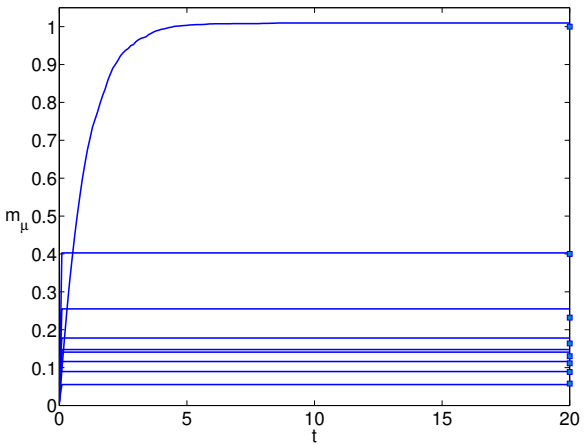

Figure 3.10: Monte Carlo simulations with $N = 10^4$ spins, $q_\mu = (0.7)^\mu q$. $T = 0.001$, $q = 6$, $\delta = \gamma = 0.25$. The markers represent the theoretical predictions at $T = 0$ in eq.(3.2.53).

This leads to the following heuristic rule for noiseless clone activations

$$m_\mu = \prod_{j=1}^{\mu-1} \left(1 - \frac{q_j}{N^\gamma}\right), \qquad \mu = 1, \dots, P, \tag{3.2.53}$$

in agreement, up to finite size effects, with Monte Carlo simulations shown in fig. 3.10.

3.2.3.2 Sequential B clones activation: critical temperature and interference effects.

In this section, we calculate the critical temperature at which clones $\mu \geq 2$ with fewer receptors get activated. We focus on the regime of clonal interference $\delta \geq \gamma$, as for $\delta < \gamma$ each clone μ gets active at its own critical temperature $T_\mu = q_\mu$. Without loss of generality we can set $P = \alpha N^\gamma$ for $\delta \geq \gamma$, where $\alpha = 1$ for $\delta = \gamma$ and $\alpha \to \infty$ for $\delta > \gamma$. In the following we will consider $q_1 > q_2 > \cdots > q_P$. Assuming that $\mathbf{m}$ bifurcates continuously from the pure state, we can Taylor expand (3.2.31) at the steady state for small m_ν, with $\nu \neq 1$, while $m_1 = \mathcal{O}(1)$.

For $\mu \neq 1$ we have

$$m_\mu = \frac{N^\gamma}{q_\mu}\left\langle \xi^\mu \tanh\left(\beta(\xi^1 q_1 m_1 + \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu)\right)\right\rangle_{\boldsymbol{\xi}} \tag{3.2.54}$$

$$= \left\langle \tanh\left(\beta\left(\xi^1 q_1 m_1 + \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}}, \tag{3.2.55}$$

where in (3.2.55) we averaged over $\xi^\mu \neq 0$. Taking the average over $\xi^1 \neq 0$ we have

$$\begin{aligned} m_\mu = &\left(1 - \frac{q_1}{N^\gamma}\right)\left\langle \tanh\left(\beta \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} + \\ &+ \frac{q_1}{2N^\gamma}\left(\left\langle \tanh\left(\beta\left(q_1 m_1 + \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} + \right. \\ &\left. + \left\langle \tanh\left(\beta\left(-q_1 m_1 + \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}}\right). \end{aligned} \tag{3.2.56}$$

Indeed, Taylor-expanding in (3.2.56) for small $\{m_\nu\}$ we have

$$\begin{aligned} &\frac{q_1}{2N^\gamma}\left(\left\langle \xi^\mu \tanh\left(\beta\left(q_1 m_1 + \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} + \right. \\ &\left. + \left\langle \xi^\mu \tanh\left(\beta\left(-q_1 m_1 + \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}}\right) \simeq \\ &\simeq \frac{q_1}{N^\gamma}(\tanh(\beta q_1 m_1) + \beta(1 - \tanh^2(\beta q_1 m_1))\left\langle \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right\rangle_{\boldsymbol{\xi}} + \\ &+ \beta^2(\tanh^3(\beta q_1 m_1) - \tanh(\beta q_1 m_1))\left\langle \sum_{\nu,\rho=2}^{P} q_\nu q_\rho \xi^\nu \xi^\rho m_\nu m_\rho\right\rangle_{\boldsymbol{\xi}}, \end{aligned} \tag{3.2.57}$$

with $\left\langle \sum_{\nu,\rho=2}^{P} q_\nu q_\rho \xi^\nu \xi^\rho m_\nu m_\rho \right\rangle_{\boldsymbol{\xi}} = \mathcal{O}(N^{-\gamma})$ and $\left\langle \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu \right\rangle_{\boldsymbol{\xi}} = 0$. Hence, the leading term is given by

$$m_\mu \simeq \left\langle \tanh\left(\beta \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} + \mathcal{O}(N^{-\gamma}). \tag{3.2.58}$$

Expanding for small $\{m_\nu\}$, the leading term one obtains

$$\begin{aligned} m_\mu &\simeq \beta \sum_{\nu=2}^{P} q_\nu \langle \xi^\mu \xi^\nu \rangle_{\boldsymbol{\xi}} m_\nu - \frac{\beta^3}{3} \sum_{\rho,\nu,\lambda}^{P} \langle \xi^\mu \xi^\nu \xi^\rho \xi^\lambda \rangle_{\boldsymbol{\xi}} m_\nu m_\rho m_\lambda q_\nu q_\rho q_\lambda \\ &\simeq \beta q_\mu m_\mu - \frac{\beta^3 q_\mu^3}{3} m_\mu^3 - \frac{\beta^3 q_\mu m_\mu}{N^\gamma} \sum_{\rho \neq \mu}^{P} q_\rho^3 m_\rho^2 \ . \end{aligned} \tag{3.2.59}$$

The solutions are $m_\mu = 0$, or

$$1 = \beta q_\mu - \frac{\beta^3 q_\mu^3}{3} m_\mu^2 - \frac{\beta^3 q_\mu}{N^\gamma} \sum_{\rho \neq \mu}^{P} q_\rho^3 m_\rho^2 \ . \tag{3.2.60}$$

Hence, for $1 - \beta q_\mu > 0$, $m_\mu = 0$ while for $1 - \beta q_\mu < 0$ the μ-th clone may be activated. Hence, the first state to bifurcate away from the pure state is $\mathbf{m} = (m_1, m_2, 0, \ldots, 0)$ at $T = q_2$. Its amplitude at $\beta q_2 = 1 - \epsilon$ is

$$m_2^2 \simeq 3\epsilon - \frac{3 q_1^3}{q_2^2 N^\gamma} m_1^2 \ , \tag{3.2.61}$$

i.e. $m_2^2 = 3\epsilon + \mathcal{O}(N^{-\gamma})$. Lowering T below q_2 we expect that the clones will activate sequentially one after another, each at its own temperature. In particular, assuming $m_{1,2} = \mathcal{O}(1)$ and repeating the reasoning shown above, we have for $\mu > 2$

$$m_\mu = \left\langle \tanh\left(\beta \sum_{\nu>2}^{P} q_\nu \xi^\nu m_\nu \right) \right\rangle_{\boldsymbol{\xi}} + \mathcal{O}(N^{-\gamma}) \ , \tag{3.2.62}$$

and expanding for $m_{\nu>2}$ small at $T < q_2$ shows that m_3 becomes non-zero at $T = q_3 + \mathcal{O}(N^{-\gamma})$.

Generalizing to $n \ll N^\gamma$ activated clones $m_1, \ldots, m_n = \mathcal{O}(1)$, we have for $\mu > n$

$$m_\mu = \left\langle \tanh\left(\beta \sum_{\nu>n}^{P} q_\nu \xi^\nu m_\nu \right) \right\rangle_{\boldsymbol{\xi}} + \mathcal{O}(nN^{-\gamma}) \ , \tag{3.2.63}$$

giving as bifurcation temperature $T_n = q_n + \mathcal{O}(nN^{-\gamma})$. As the number of activated clones increases, their cumulative effect on the activation temperature of the remaining clones increases and can no longer be neglected for $n = \mathcal{O}(N^{\gamma})$. The activation temperature of pattern $n+1$, when $n \sim N^{\gamma}$ clones have been activated, can be computed from

$$m_{n+1} = \left\langle \tanh\left(\beta\left(q_{n+1}m_{n+1} + \sum_{\nu=1}^{n} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} . \tag{3.2.64}$$

Inserting $\int_{-\infty}^{\infty} \mathrm{d}z\delta\left(z - \sum_{\nu=1}^{n} q_\nu \xi^\nu m_\nu\right) = 1$, we obtain

$$m_{n+1} = \int \mathrm{d}z P_n(z|\mathbf{m},\mathbf{q}) \tanh(\beta(q_{n+1}m_{n+1} + z)) , \tag{3.2.65}$$

with $P_n(z|\mathbf{m},\mathbf{q})$ defined in (3.2.35). Taylor expanding for small m_{n+1} we have, to leading order

$$m_{n+1} = \int \mathrm{d}z P_n(z|\mathbf{m},\mathbf{q}) \left[(1 - \tanh^2(\beta z))\beta q_{n+1}m_{n+1} + \mathcal{O}(m_{n+1}^2)\right] , \tag{3.2.66}$$

where we have used $P_n(z|\mathbf{m},\mathbf{q}) = P_n(-z|\mathbf{m},\mathbf{q})$. One solution is $m_{n+1} = 0$ and a non-zero solution is possible when

$$\beta q_{n+1} = \frac{1}{1 - \int \mathrm{d}z P_n(z|\mathbf{m},\mathbf{q}) \tanh^2(\beta z)} . \tag{3.2.67}$$

For $n \ll N^{\gamma}$, $P_n(z|\mathbf{m},\mathbf{q}) = \delta(z)$, and we retrieve $\beta q_{n+1} = 1$ for the temperature at which m_{n+1} becomes non-zero. For $P_n(z|\mathbf{m},\mathbf{q})$ having a small but finite width we can use $\tanh(\beta z) \simeq \beta z$

$$\beta q_{n+1} \simeq 1 + \frac{\langle z^2 \rangle}{q_{n+1}^2} \simeq 1 + \frac{\sum_{\mu=1}^{n} q_\mu^3 m_\mu^2}{q_{n+1}^2 N^{\gamma}} , \tag{3.2.68}$$

showing that $T_{n+1} < q_{n+1}$ and deviations from q_{n+1} depend on the promiscuity distribution of the activated clones. Equation (3.2.67) shows that as more clones are activated, these create an interference, encoded in $P_n(z|\mathbf{m},\mathbf{q})$, which decreases

the activation temperature of the inactive ones. Furthermore, it suggests that the number n of clones that the system can activate (i.e. the number of $\mathcal{O}(1)$ order parameters m_μ) at small but finite temperature is $\mathcal{O}(N^\gamma)$.

3.2.3.3 Numerical examples.

In this section we test (3.2.67) and look at the effect of receptor promiscuity on the intensity of B clones activation, for three simple cases that can be treated analytically, with $q_1 > q_2$ and $P = \alpha N^\gamma$:

(i) $\mathbf{q} = (q_1, q_2, \dots, q_2)$: only one clone has a higher promiscuity;

(ii) $\mathbf{q} = (q_1, \dots, q_1, q_2, \dots, q_2)$: half of the clones have promiscuity q_1, and half have promiscuity q_2;

(iii) $\mathbf{q} = (q_1, \dots, q_1, q_2)$: only one clone has a smaller promiscuity.

Our goal is to analyse the increasing interference effect due to the activated clones with more receptors on the quiescent ones with less receptors. According to (3.2.67) active clones should play the role of interference terms that lower the critical temperature of the quiescent clones.

(i) Near T_c clones with the same promiscuity will be activated with the same intensity, as discussed in chapter 2. In this case the activation vector bifurcating away from the pure state will have the form $\mathbf{m} = (m_1, m_2, \dots, m_2)$, where m_1, m_2 are the amplitudes associated to clones with promiscuity q_1 and q_2 respectively. The amplitudes m_1 and m_2 , computed from the steady

state in (3.2.10), are

$$\begin{aligned} m_1 &= \frac{N^\gamma}{q_1}\left\langle \xi^1 \tanh\left(\beta\left(\xi^1 q_1 m_1 + q_2 \sum_{\nu=2}^{P} \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} \\ &= \left\langle \tanh\left(\beta\left(q_1 m_1 + q_2 m_2 \sum_{\nu=2}^{P} \xi^\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} \\ &= \sum_z \tilde{P}_{P-1}(z|\mathbf{q}) \tanh(\beta(q_1 m_1 + q_2 m_2 z)) \ . \end{aligned} \tag{3.2.69}$$

For m_2 we obtain

$$\begin{aligned} m_2 &= \sum_z \tilde{P}_{P-2}(z|\mathbf{q}) \langle \tanh(\beta(q_1 \xi^1 m_1 + q_2 m_2(1+z)))\rangle_{\xi^1} \\ &= \sum_z \tilde{P}_{P-2}(z|\mathbf{q}) \tanh(\beta q_2 m_2(1+z)) + \mathcal{O}(N^{-\gamma}) \ . \end{aligned} \tag{3.2.70}$$

Analogously to sec. 2.3.2.1 (eq. (2.3.77)), we have introduced here a discrete noise distribution of the form

$$\tilde{P}_P(z|\mathbf{q}) = \left\langle \delta_{z,\sum_{\nu=1}^{P} \xi^\nu} \right\rangle_{\boldsymbol{\xi}} . \tag{3.2.71}$$

Using (3.2.34) and $P = \alpha N^\gamma$, we can write $\tilde{P}_P(z|\mathbf{q}) = \mathrm{e}^{-q_2\alpha} I_z(q_2\alpha)$ where $I_z(x)$ is a modified Bessel function of the first kind [45] and $\tilde{P}_{P-1}(z|\mathbf{q}) \simeq \tilde{P}_{P-2}(z|\mathbf{q}) \simeq \tilde{P}_P(z|\mathbf{q})$. The activation temperature of the clones with smaller promiscuity follows from Taylor expansion of (3.2.70) for small m_2 giving

$$\begin{aligned} m_2 &= \sum_z \tilde{P}_{P-2}(z|\mathbf{q}) \left[\beta q_2 m_2(1+z) - \frac{1}{3}(\beta q_2 m_2)^3 (1+z)^3\right] + \mathcal{O}(N^{-\gamma}) \\ &= \beta q_2 m_2 - \frac{1}{3}(\beta q_2 m_2)^3 (1 + 3\langle z^2\rangle) + \mathcal{O}(N^{-\gamma}) \\ &= \beta q_2 m_2 - \frac{1}{3}(\beta q_2 m_2)^3 (1 + 3\alpha q_2) + \mathcal{O}(N^{-\gamma}) \ , \end{aligned} \tag{3.2.72}$$

where we used the parity of $\tilde{P}_P(z|\mathbf{q}) = \tilde{P}_P(-z|\mathbf{q})\ \forall\, n$ and $\langle z^2\rangle \equiv \sum_z \tilde{P}_P(z|\mathbf{q}) z^2 = \alpha q_2$. For $P = 1$ we retrieve $\beta q_2 = 1 + \mathcal{O}(N^{-\gamma})$ for the activation temperature of m_2, consistently with (3.2.67). The activation intensity at $\beta q_2 = 1 + \epsilon$

follows from (3.2.72) as

$$m_2 = \frac{3\epsilon}{1 + 3q_2\alpha} \ , \tag{3.2.73}$$

as one expects for clones with homogeneous promiscuity q_2 as derived in chapter 3 in eq. (2.4.101). In fig. 3.11 we plot the amplitudes m_1, m_2 resulting from (3.2.69, 3.2.70), as a function of the temperature (right) and those resulting from Monte Carlo simulations, as a function of time (left). The simulations are in agreement with the theoretically predicted steady state in (3.2.69), (3.2.70). A contour plot of $m_2 = 0$ in the $T - q_2$ plane,

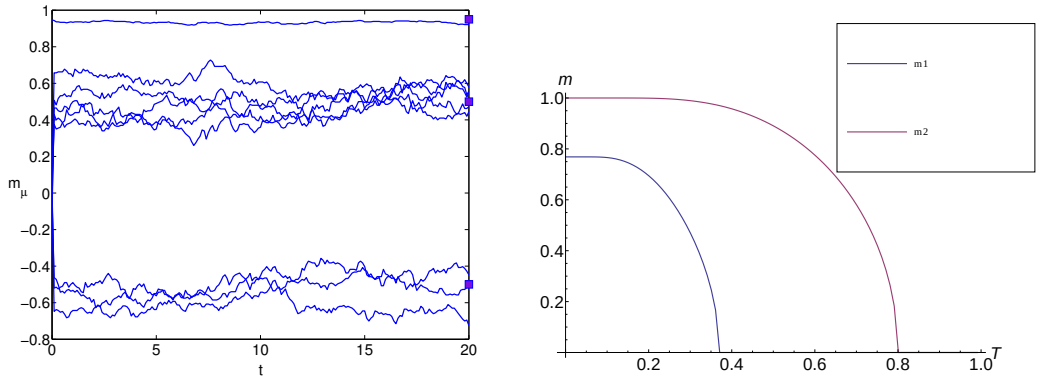

Figure 3.11: Left: Monte Carlo simulations with $N = 10^4$ spins, $\delta = \gamma = 0.25$, $T = 0.3$, $q_1 = 0.8$, $q_2 = 0.4$. We plot the magnetisations as a function of time at fixed temperature; the markers are the theoretically predicted symmetric steady state activations for m_1, m_2 from the dynamical system (3.2.69, 3.2.70). Right: Full temperature dependence of the steady state solutions m_1, m_2 of the dynamical system (3.2.69, 3.2.70) for $q_1 = 0.8, q_2 = 0.4$.

computed numerically from (3.2.70), is shown in fig. 3.12. Deviations from the line $T = q_2$ are consistent with finite size effects $N^{-\gamma}$.

In conclusion, the presence of one clone with a higher number of receptors, does not affect, in the thermodynamic limit, the activation temperature nor the activation intensity of clones with fewer receptors.

(ii) In this case we expect a transition from the state $\mathbf{m} = (m_1, \ldots, m_1, 0, \ldots, 0)$

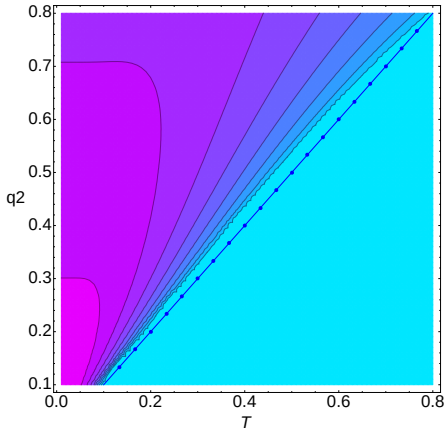

Figure 3.12: Contours of constant m_2 with $\mathbf{q} = (q_1, q_2, \ldots, q_2)$, $q_1 = 0.8$. The contour indicating the onset of a non-zero value of m_2 is in agreement with the theoretically predicted blue dotted line $T = q_2$.

to $\mathbf{m} = (m_1, \ldots, m_1, m_2, \ldots, m_2)$, with

$$m_1 = \sum_{z_1,z_2} \tilde{P}_{P/2}(z_1|\mathbf{q})\tilde{P}_{P/2}(z_2|\mathbf{q}) \tanh(\beta(q_1 m_1(1+z_1) + q_2 m_2 z_2)) \ ,$$
$$m_2 = \sum_{z_1,z_2} \tilde{P}_{P/2}(z_1|\mathbf{q})\tilde{P}_{P/2}(z_2|\mathbf{q}) \tanh(\beta(q_2 m_2(1+z_2) + q_1 m_1 z_1)) \ . \quad (3.2.74)$$

The temperature at which the onset of non-zero m_2 occurs can be found by Taylor expanding (3.2.74) for small m_2 with $m_1 = \mathcal{O}(1)$

$$\begin{aligned} m_2 = \sum_{z_1,z_2} \tilde{P}_{P/2}(z_1|\mathbf{q})\tilde{P}_{P/2}(z_2|\mathbf{q}) \Big[& \tanh(\beta q_1 m_1 z_1) + \\ & + (1 - \tanh^2(\beta q_1 m_1 z_1))\beta q_2 m_2(1+z_2) + \mathcal{O}(m_2^2) \Big] = \\ = \beta q_2 m_2 \sum_{z_1} & \tilde{P}_{P/2}(z_1|\mathbf{q})(1 - \tanh^2(\beta q_1 m_1 z_1)) \end{aligned} \quad (3.2.75)$$

giving $m_2 \neq 0$ for

$$\beta q_2 = \frac{1}{1 - \sum_z \tilde{P}_{P/2}(z|\mathbf{q}) \tanh^2(\beta q_1 m_1 z)} \ . \quad (3.2.76)$$

The theoretically predicted critical line (3.2.76) is in good agreement with the contour plot of $m_2 \neq 0$ in the $T - q_2$ plane, shown in fig. 3.13 (left),

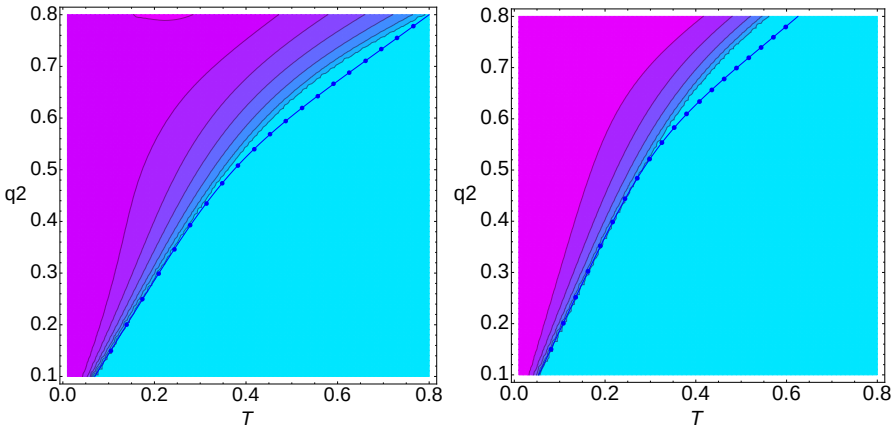

Figure 3.13: Contour plot of m_2 in the (T, q_2) plane for $\mathbf{q} = (q_1, \ldots, q_1, q_2, \ldots, q_2)$ (left) and $\mathbf{q} = (q_1, \ldots, q_1, q_2)$ (right) with $q_1 = 0.8$. The blue dotted line represents the theoretical critical temperature line computed using respectively the self-consistent equations (3.2.74) (left) and (3.2.78)(right) together with the theoretical predictions for the critical temperature (3.2.76)(left) and (3.2.81)(right). Deviations from the numerical results when $q_2 \simeq q_1$ are due to the fact that (3.2.67) is obtained assuming $m_2 \ll m_1$, condition which is not satisfied when $q_2 \simeq q_1$.

computed numerically from (3.2.74). The plot shows that in the presence of $\mathcal{O}(N^\gamma)$ clones with higher numbers of receptors, the activation temperature of those with smaller promiscuity q_2 will deviate from the line $T = q_2$.

(iii) This is the case where deviations from the line $T = q_2$ are expected to be the largest. The steady state equations for the magnetisation read

$$\begin{aligned} m_1 &= \frac{N^\gamma}{q_1}\left\langle \xi^1 \tanh\left(\beta\left(\xi^1 q_1 m_1 + \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} \\ &= \left\langle \tanh\left(\beta\left(q_1 m_1 + \sum_{\nu=2}^{P} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}}, \end{aligned} \tag{3.2.77}$$

where in the last line we are averaging over $\xi^1 \neq 0$. Separating the contribution of the only clone with promiscuity q_2 and introducing the usual

discrete noise distribution, yields

$$\begin{aligned} m_1 &= \sum_z \tilde{P}_{P-2}(z|\mathbf{q})\langle \tanh(\beta(q_1 m_1(1+z) + q_2 m_2 \xi^2))\rangle_{\xi^2} \\ &= \sum_z \tilde{P}_{P-2}(z|\mathbf{q}) \tanh(\beta q_1 m_1(1+z)) + \mathcal{O}(N^{-\gamma}) \ . \end{aligned} \tag{3.2.78}$$

Equivalently for m_2 we have

$$\begin{aligned} m_2 &= \left\langle \tanh\left(\beta\left(q_2 m_2 + \sum_{\nu\neq 2}^{P} q_\nu \xi^\nu m_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} \\ &= \sum_z \tilde{P}_{P-1}(z|\mathbf{q}) \tanh\left(\beta(q_2 m_2 + q_1 m_1 z)\right) \ . \end{aligned} \tag{3.2.79}$$

When clones with lower promiscuity become active, we expect $m_1 = \mathcal{O}(1)$ and m_2 small. Taylor expanding for m_2 small and using the parity of $\tilde{P}_P(z|\mathbf{q}) = \tilde{P}_P(-z|\mathbf{q}) \ \forall \ n$

$$\begin{aligned} m_2 &= \sum_z \tilde{P}_{P-1}(z|\mathbf{q})[\tanh(\beta q_1 m_1 z) + (1 - \tanh^2(\beta q_1 m_1 z)\beta q_2 m_2)] + \mathcal{O}(m_2^2) \\ &= \beta q_2 m_2 \left[1 - \sum_z \tilde{P}_{P-1}(z|\mathbf{q}) \tanh^2(\beta q_1 m_1 z)\right] + \mathcal{O}(m_2^2) \ , \end{aligned} \tag{3.2.80}$$

we obtain for the activation temperature of the only clone with promiscuity q_2

$$\beta q_2 = \frac{1}{1 - \sum_z \tilde{P}_{P-1}(z|\mathbf{q}) \tanh^2(\beta q_1 m_1 z)} \ . \tag{3.2.81}$$

This result is in agreement with the zero contour plot of m_2 in the $T - q_2$, obtained by numerically solving (3.2.80), shown in fig. 3.13 (right). Deviations are compatible with finite size effects and are more evident for $q_2 \simeq q_1$ where the assumption $m_2 \ll m_1$ is no longer valid. Monte Carlo simulations with 10 B clones, one of which has a lower promiscuity $q_2 = 0.6$, are shown in fig. 3.14: at $T = q_2$ the pattern with the lowest promiscuity is inactive and is activated only at a much lower temperature $T \simeq 0.2$.

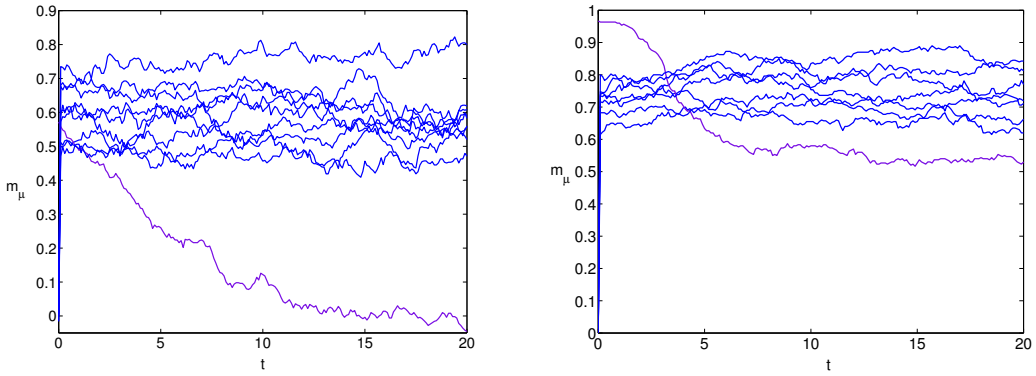

Figure 3.14: Time evolution of B clones activation in Monte Carlo simulations with 10^4 spins and $\mathbf{q} = (q_1, \ldots, q_1, q_2)$ (case (iii)) with $q_1 = 0.8$ (blue) and $q_2 = 0.6$ (violet). The activation for the clone with fewer receptors (violet) decays to 0 for $T = 0.5$ (left) and stays non-zero for $T = 0.2$ (right), i.e. at temperatures considerably lower than in the absence of clonal interference $T = q_2$.

Finally in fig. 3.15 we plot m_2 as a function of the temperature in the three different cases with promiscuity (i) $\mathbf{q} = (q_1, q_2, \ldots, q_2)$, (ii) $\mathbf{q} = (q_1, \ldots, q_1, q_2, \ldots, q_2)$, (iii) $\mathbf{q} = (q_1, \ldots, q_1, q_2)$. The presence of clones with higher promiscuity q_1 and activation intensity m_1, does not only affect the activation temperature of clones with lower promiscuity q_2, but also the intensity m_2 of their immune responses. In conclusion, clones with fewer triggered receptors are activated at a lower temperature and will produce a weaker response than clones with a higher number of active receptors.

3.3 Idiotypic interactions

Cells and molecules of the adaptive immune system can recognise, react to and regulate each other forming an interacting network. In particular, B cells can bind, recognise and suppress or excite each other forming the so-called *idiotypic network* of interactions. Different B-B interaction mechanisms have been proposed, e. g. via cell receptors or antibodies, but in this model we are not concerned with the microscopic details of such interactions and we will rather focus on their mere existence. The first formulation of an immune network theory by

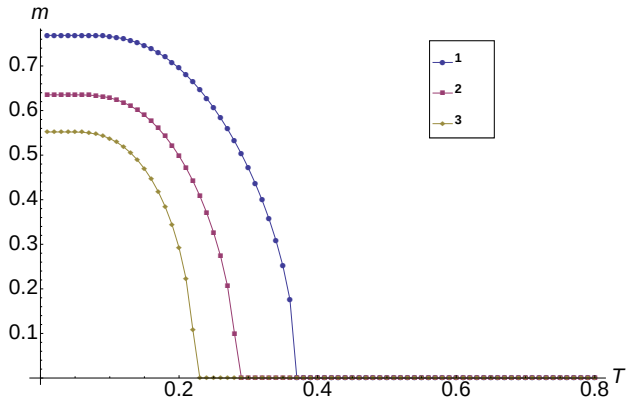

Figure 3.15: Plot of m_2 as a function of T in the different cases: (1) $\mathbf{q} = (q_1, q_2, \ldots, q_2)$, (2) $\mathbf{q} = (q_1, \ldots, q_1, q_2, \ldots, q_2)$, (3) $\mathbf{q} = (q_1, \ldots, q_1, q_2)$. Increasing the interference due to clones activated at a higher temperature, the m_2 intensity decreases.

N. K. Jerne [4] dates back to the early 1970s and was soon after followed by preliminary experimental confirmations [55–57] and various mathematical models to capture this interaction mechanism [14]. In particular, linear idiotypic network models [58], where antibodies belonging to the i-th group could interact with antibodies of the $i-1$ and $i+1$-th group, or a cyclic version of it [59] have been proposed, together with models with random connections between antibodies [18]. In these models, the dynamics of the B clones concentration is modelled via large systems of coupled ODEs, mostly solved via numerical simulations. More recently, experiments also showed a new connection between the existence of an "idiotypic network" and the onset of autoimmune diseases, spotlighting the importance of such interactions and the need for a more careful analysis of their effects [49–52].

To investigate B-B interactions, we choose the simplest configuration of the receptor space which retains the key biological features. We represent B clones receptors as binary strings and assume that complementary strings, like e.g. $010\ldots$ and $101\ldots$, excite each other. Noting that one can always label each of $P = 2^d$ binary strings of length d by an index μ, so that the string complementary to μ is labeled by $\mu + P/2$, we order the strings on a ring in such a way that each string

sits close to similar strings and opposite to complementary ones. The reason for this choice is that B-B interactions are believed to be very specific [4, 60], with each B clone interacting with a number $\mathcal{O}(1)$ of other clones. For simplicity, we assume that each B clone interacts only with another clone, but this assumption can be easily relaxed; for example, one may assume that each B clone does not only interact with the cell strictly complementary to it (with Hamming distance d), but also with the d clones which are nearly complementary (with a Hamming distance $d-1$). We expect that as long as the connectivity of B clones is $\sim d \ll N$ the effect should be qualitatively very similar. We suppose that the μ-th B clone expansion is triggered by the $(\mu + P/2)$-th B clone [mod P], which is precisely complementary to that B clone. Each B clone will die at a certain rate, while complementary B clones will excite each other. Therefore, for the B-B interactions we use the following matrix

$$A_{\mu\nu} = \delta_{\mu\nu} - k\delta_{\mu,(\nu+P/2) \bmod P}, \tag{3.3.82}$$

with $k \in [0, 1)$ representing the strength of idiotypic interactions. The non-zero diagonal terms $A_{\mu\mu}$ account for a decay term in the population dynamics (taken as a gradient descent on $\mathcal{H}(\boldsymbol{\sigma}, \mathbf{b})$ [61]) of each clone μ, proportional to the population size b_μ itself. For simplicity we assumed a unit death rate for all clones. Intra-clonal competition can be introduced by allowing non-zero off-diagonal entries in $\mathbf{A}$. These elements would need to scale as $1/P$ to ensure that the decay term is overall $\mathcal{O}(1)$. Assuming that off-diagonal terms are all equal i.e. $A_{\mu\nu} = 1/P \ \forall \ \mu \neq \nu$ and the overall log-concentration of clones $\sum_{\mu=1}^{P} b_\mu = B$ is constant, the loss term in the population dynamics gains a constant term $-B/P$, which can be absorbed in the definition of b_μ. This choice of the matrix also allows for a straightforward matrix inversion, useful for the analytical treatment, while retaining the two key biological ingredients of (i) suppression effects preventing abnormal clonal expansion and (ii) excitatory signalling between complementary

clones. Its inverse is

$$(\mathbf{A}^{-1})_{\mu\nu} = \frac{1}{1-k^2}\delta_{\mu\nu} + \frac{k}{1-k^2}\delta_{\mu,(\nu+P/2) \bmod P}. \tag{3.3.83}$$

We note that $\mathbf{A}$ is positive definite and symmetric and the following relations hold: $A_{\mu\nu} = \mathbf{A}(\mu - \nu)$ with $\mathbf{A}(n+P) = \mathbf{A}(n)$ and $\mathbf{A}(n) = \mathbf{A}(-n)$. The eigenvalue spectrum $\rho(\mu)$ of $\mathbf{A}$ has a finite limit as $N \to \infty$, ensuring the correct scaling of the Hamiltonian.

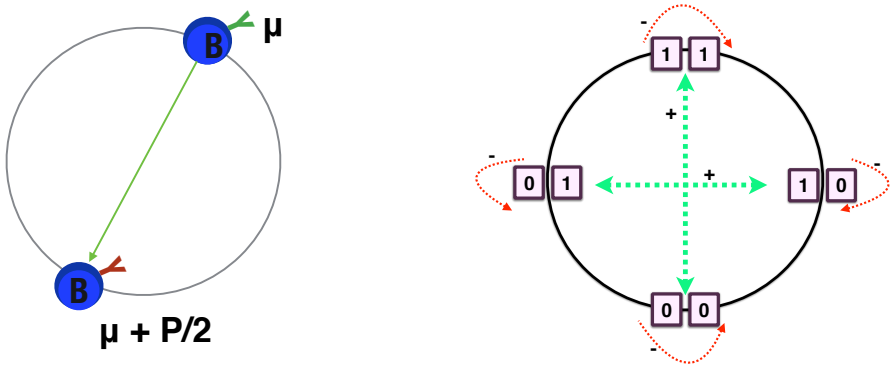

Figure 3.16: Left: B-B interaction via $A_{\mu\nu}$. The expansion of the μ-th clone is triggered by its complementary clone, $\mu + P/2$, where P is the number of B clones. Right: We represent the epitopes as binary strings, organising them on a ring and assuming that the complementary strings interact.

In the following, we will show that the idiotypic interactions contribute to the overall stability of the network, sustaining the propagation of a suppressive/excitatory signal for the immune response through the system. In healthy situations, this may prevent unwanted activations and increase the region where all clones are equally activated and ready to start an immune response upon arrival of new infections. In particular, including B-B interactions in the model affects the critical temperature, in this case widening the region where a symmetric immune response is stable, as the interactions strength k is increased (fig. 3.19). This increased network stability may not always have positive effects on the immune system performances: in presence of autoimmune diseases, for example, this may encourage the transmission of a wrong signal, e.g. producing autoantibodies,

which would attack self-cells [49].

3.3.1 Dynamical equations

In this section, we study the effect of the idiotypic interactions on the ability of the system to activate multiple clones in parallel. For simplicity, here we will assume homogeneous receptor promiscuities i.e. $q_\mu = c$, $\forall \mu$, so that the dynamical equations (3.1.7) read

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \frac{N^\gamma}{c}\left\langle \xi^\mu \tanh\left(\beta c \sum_{\rho,\nu=1}^{P} \xi^\nu (\mathbf{A}^{-1})_{\rho\nu} m_\rho\right)\right\rangle_{\boldsymbol{\xi}} - m_\mu \ . \tag{3.3.84}$$

The matrix $\mathbf{A}$ is positive definite for $k \in (0,1)$ and symmetric, hence the free energy of the system is a Lyapunov function for the dynamics [62] and the system will converge to a steady state. Next, we compute the critical temperature at which clonal activation emerges. Starting from the steady state

$$m_\mu = \frac{N^\gamma}{c}\left\langle \xi^\mu \tanh\left(\beta c \sum_{\rho,\nu=1}^{P} \xi^\nu (\mathbf{A}^{-1})_{\rho\nu} m_\rho\right)\right\rangle_{\boldsymbol{\xi}} , \tag{3.3.85}$$

and summing over $\sum_{\mu,\lambda=1}^{P} m_\lambda (\mathbf{A}^{-1})_{\lambda\mu}$, yields

$$\sum_{\mu,\lambda=1}^{P} m_\lambda (\mathbf{A}^{-1})_{\lambda\mu} m_\mu = \frac{N^\gamma}{c}\left\langle \sum_{\mu,\lambda=1}^{P} m_\lambda (\mathbf{A}^{-1})_{\lambda\mu} \xi^\mu \tanh\left(\beta c \sum_{\rho,\nu=1}^{P} \xi^\nu (\mathbf{A}^{-1})_{\rho\nu} m_\rho\right)\right\rangle_{\boldsymbol{\xi}} . \tag{3.3.86}$$

Using the inequality $|\tanh(x)| \leq |x|$ and averaging over the disorder we obtain

$$\sum_{\mu,\lambda=1}^{P} m_\lambda (\mathbf{A}^{-1})_{\lambda\mu} m_\mu \leq \beta c \sum_{\rho,\lambda=1}^{P} m_\lambda (\mathbf{A}^{-2})_{\lambda\rho} m_\rho \ . \tag{3.3.87}$$

Next we diagonalise the matrix $\mathbf{A}$ by means of the similarity transformation $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, where $\mathbf{D}$ is the diagonal matrix constructed from the eigenvalues $\{\mu_\rho\}_{\rho=1}^{P}$ of $\mathbf{A}$ and $\mathbf{P}$ is the orthogonal matrix of eigenvectors, i.e. $\mathbf{P}^{-1} = \mathbf{P}^T$.

Hence, we can rewrite the equation above for the transformed vector $\mathbf{v} = \mathbf{P}^{-1}\mathbf{m}$ as

$$\mathbf{v}^T\mathbf{D}^{-1}\mathbf{v} - \beta c\mathbf{v}^T\mathbf{D}^{-2}\mathbf{v} \leq 0 \ , \tag{3.3.88}$$

which gives

$$\sum_{\nu=1}^{P}\left(\frac{v_\nu}{\mu_\nu}\right)^2(\beta c - \mu_\nu) \geq 0 \ . \tag{3.3.89}$$

This yields $\mathbf{v} = \mathbf{0}$ for $\beta c < \mu_{\min}$, where $\mu_{\min} = \min_\nu \mu_\nu$. The eigenvalues of $\mathbf{A}$ are

$$\mu_1 = 1 - k, \qquad deg(\mu_1) = P/2 \ , \tag{3.3.90}$$

$$\mu_2 = 1 + k, \qquad deg(\mu_2) = P/2 \ , \tag{3.3.91}$$

hence, above the critical temperature $T_c = c(\mu_{\min})^{-1} = c/(1-k)$ we have $\mathbf{m} = \mathbf{0}$, whereas below criticality $\mathbf{m} \neq \mathbf{0}$ is possible (and expected). Remarkably, the critical temperature increases with k, meaning that idiotypic interactions enhance the immune system's activation. Next, we investigate the structure of the states bifurcating away from $\mathbf{m} = \mathbf{0}$ below T_c and their stability.

3.3.2 Dynamical equations for two B clones

As before, it is useful to consider first the toy model with $P = 2$ clones. For $k \neq 0$, the dynamics of different clones is always coupled, for all $\gamma \geq 0$. We first

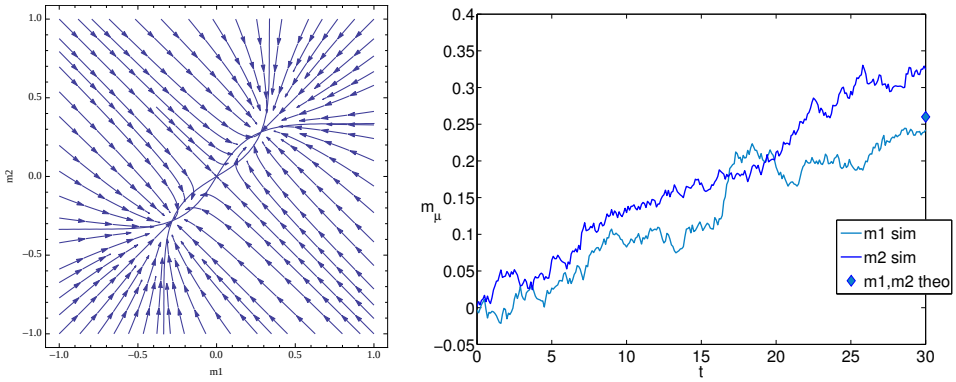

Figure 3.17: Flow diagram in the plane (m_1, m_2) (left) and Monte Carlo simulations with 10^4 spins (right) for the dynamical system (3.3.92) with $T = 1.7, c = 0.55, k = 0.7$ ($T_c = 1.83$).

look at the case $\gamma = 0$, where the system evolves according to the equations

$$\begin{aligned}
\frac{\mathrm{d}m_1}{\mathrm{d}t} &= (1-c)\tanh\left(\frac{\beta c}{1-k^2}(m_1+km_2)\right)+\frac{c}{2}\Bigg[\tanh\left(\frac{\beta c}{1-k}(m_1+m_2)\right)\\
&\quad+\tanh\left(\frac{\beta c}{1-k}(m_1-m_2)\right)\Bigg]-m_1\ ,\\
\frac{\mathrm{d}m_2}{\mathrm{d}t} &= (1-c)\tanh\left(\frac{\beta c}{1-k^2}(km_1+m_2)\right)+\frac{c}{2}\Bigg[\tanh\left(\frac{\beta c}{1-k}(m_1+m_2)\right)\\
&\quad-\tanh\left(\frac{\beta c}{1-k}(m_1-m_2)\right)\Bigg]-m_2\ .
\end{aligned} \tag{3.3.92}$$

The state $\mathbf{m} = (0,0)$ is a fixed point of the dynamics at all temperatures, but we expect it to become unstable below $T_c = c/(1-k)$. In order to inspect the structure of the bifurcating state, it is convenient to analyse the steady state equations in terms of the variables $x = m_1 + m_2$ and $y = m_1 - m_2$

$$\begin{aligned}
x &= c\tanh\left(\frac{\beta c x}{1-k}\right)+(1-c)\Bigg[\tanh\left(\frac{\beta c}{1-k^2}(x+(k-1)m_2)\right)\\
&\quad+\tanh\left(\frac{\beta c}{1-k^2}(x+(k-1)m_1)\right)\Bigg]\ ,
\end{aligned} \tag{3.3.93}$$

$$\begin{aligned}
y &= c\tanh\left(\frac{\beta c y}{1-k}\right)+(1-c)\Bigg[\tanh\left(\frac{\beta c}{1-k^2}(x+(k-1)m_2)\right)\\
&\quad-\tanh\left(\frac{\beta c}{1-k^2}(x+(k-1)m_1)\right)\Bigg]\ .
\end{aligned} \tag{3.3.94}$$

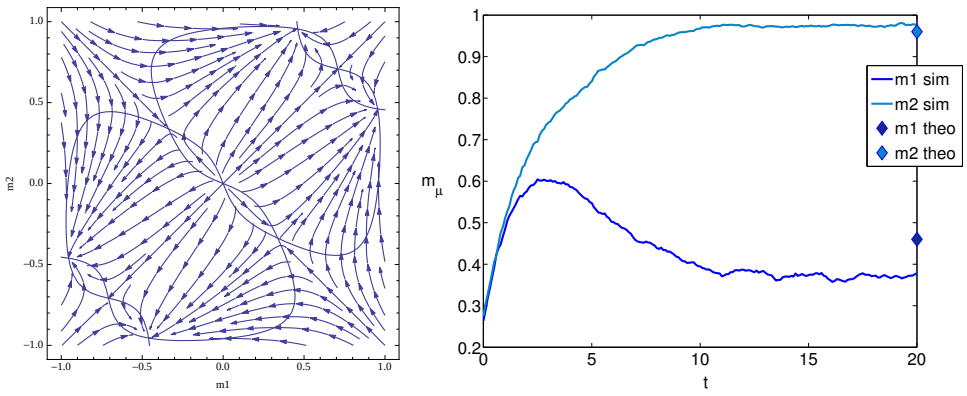

Figure 3.18: Flow diagram in the plane (m_1, m_2) (left) and Monte Carlo simulations with 10^4 spins (right) for the dynamical system (3.3.92), for $T = 0.65, c = 0.55, k = 0.7$ ($T_c = 1.83$).

Assuming continuous bifurcations, we Taylor expand for small $\mathbf{m}$ near $\beta c = 1 - k$ i.e. at $\tilde{\beta} \equiv \frac{\beta c}{1-k} = 1 + \epsilon$, obtaining, to leading orders

$$\begin{aligned} x &= c\tilde{\beta}x - \frac{1}{3}c\tilde{\beta}^3x^3 + (1-c)\tilde{\beta}x - \frac{1-c}{3(1+k)^2}\tilde{\beta}^3 x\Big[(x+(k-1)m_2)^2\ , \\ &\quad +(x+(k-1)m_1)^2 - (x+(k-1)m_1)(x+(k-1)m_2)\Big]\ , \qquad (3.3.95) \\ y &= c\tilde{\beta}y - \frac{1}{3}c\tilde{\beta}^3y^3 + \frac{(1-c)(1-k)}{(1+k)}\tilde{\beta}y - \frac{(1-c)(1-k)}{3(1+k)^3}\tilde{\beta}^3 y\Big[(x+(k-1)m_2)^2 \\ &\quad +(x+(k-1)m_1)^2 - (x+(k-1)m_1)(x+(k-1)m_2)\Big]\ . \qquad (3.3.96) \end{aligned}$$

We note that $x = y = 0$ is always a solution (corresponding to $\mathbf{m} = (0,0)$). A solution $y \neq 0$ is not possible as in (3.3.96) terms $\mathcal{O}(\epsilon^0)$ do not simplify. Hence, $y = 0$ is the only solution, implying $m_1 = m_2$. In contrast, in (3.3.95) first order terms simplify, hence we can have $x \neq 0$. We can conclude that mixtures bifurcate from $\mathbf{m} = (0,0)$ in a symmetric fashion $\mathbf{m} = m(1,1)$. We can compute the amplitude m, Taylor expanding (3.3.92) at the steady state near T_c, for small m,

$$m = \tilde{\beta}m - \frac{1}{3}\tilde{\beta}^3m^3 - c\tilde{\beta}^3m^3\ , \qquad (3.3.97)$$

yielding, at $\tilde{\beta} = 1 + \epsilon$,

$$m^2 = \frac{3\epsilon}{1+3c} \ , \tag{3.3.98}$$

or the trivial solution $m = 0$. Flow diagrams and Monte Carlo simulations are in agreement with theoretical predictions, showing that close to criticality both B clones are activated with the same intensity (fig. 3.17). The stability of symmetric solutions can be analysed by computing the eigenvalues of the Jacobian

$$\mathcal{J}_{\mu\nu} = \beta c(\mathbf{A}^{-1})_{\mu\nu} \left(1 - \left\langle \tanh^2 \left(\beta c \sum_{\rho,\lambda=1}^{P} \xi^{\lambda} (\mathbf{A}^{-1})_{\rho\lambda} m_{\rho}^{\star} \right) \right\rangle_{\boldsymbol{\xi}} \right) - \delta_{\mu\nu} + \mathcal{O}(N^{-\gamma}) \tag{3.3.99}$$

and substituting $\mathbf{m}^* = m(1,1)$ we find for the eigenvalues

$$\lambda_1 = \frac{\beta c}{1-k} - 1 - \frac{\beta c(1-c)}{1-k} \tanh^2 \left(\frac{m\beta c}{1-k} \right) - \beta \frac{c^2}{1-k} \tanh^2 \left(\frac{2m\beta c}{1-k} \right) \ , \tag{3.3.100}$$

$$\lambda_2 = \frac{\beta c(1-c)}{1+k} \left(1 - \tanh^2 \left(\frac{m\beta c}{1-k} \right) \right) + \frac{\beta c^2}{1-k} - 1 \ . \tag{3.3.101}$$

These can be calculated analytically near criticality i.e. at $\beta c/1-k = 1+\epsilon$ where

$$\lambda_1 \simeq -2\epsilon \ , \tag{3.3.102}$$

$$\lambda_2 \simeq \frac{-2\epsilon + 6\epsilon c - 2k + 2k\epsilon - 4kc - 4kc\epsilon}{(1+3c)(1+k)} \ , \tag{3.3.103}$$

and for $T \to 0$, where

$$\lambda_1 \simeq -1 \ , \tag{3.3.104}$$

$$\lambda_2 \simeq \frac{\beta c^2}{1-k} - 1 \ . \tag{3.3.105}$$

We deduce that λ_1 is always negative, hence the stability of symmetric solutions is determined by λ_2. A plot of λ_2 (rescaled by T) as a function of T is shown

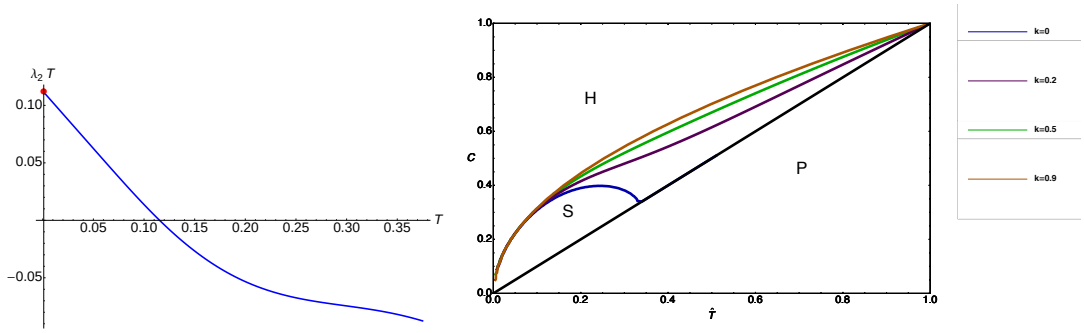

Figure 3.19: Left: $\lambda_2 T$ as a function of T for $c = 0.3$ and $k = 0.2$. The value at $T = 0$ matches the analytical prediction $\lambda_2 T = c^2/1 - k$ (red marker). Right: Contour plot of $\lambda_2 = 0$ with different interaction strengths $k = 0.8, k = 0.5, k = 0.2$ in the parameter space (T, c). The region where clones are activated in parallel with the same intensity (**S**) becomes wider as k increases.

in fig. 3.19 (left) for fixed dilution and B-B interaction strength k. The critical line where $\lambda_2 = 0$ in the $T - c$ plane for different values of k is shown in fig. 3.19 (right). Remarkably, the region (**S**) where the activation of parallel immune responses is accomplished with the same intensity becomes wider as k increases. For $\gamma > 0$, the system evolves according to the equations

$$\begin{aligned} \frac{\mathrm{d}m_1}{\mathrm{d}t} &\simeq \tanh\left(\frac{\beta c}{1-k^2}(m_1 + k m_2)\right) - m_1 + \mathcal{O}(N^{-\gamma}) \ , \\ \frac{\mathrm{d}m_2}{\mathrm{d}t} &\simeq \tanh\left(\frac{\beta c}{1-k^2}(k m_1 + m_2)\right) - m_2 + \mathcal{O}(N^{-\gamma}) \ . \end{aligned} \tag{3.3.106}$$

One can show again that $\mathbf{m} = (0, 0)$ is the only fixed point above $T_c = c/(1-k)$, and symmetric mixtures $\mathbf{m} = m(1, 1)$ bifurcate away from $\mathbf{m} = (0, 0)$ at T_c. Now, however, symmetric mixtures remain stable for any $T < T_c$ (see sec. 3.3.3.1), with the intensity m of the symmetric activation found from

$$m = \tanh\left(\frac{\beta c m}{1-k}\right) \ . \tag{3.3.107}$$

Hence for $\gamma > 0$ the system reduces to independent Curie-Weiss ferromagnets, even in the presence of idiotypic interactions, with the critical temperature $T_c = c/(1-k)$ increasing with the strength of interactions.

3.3.3 Generalisation to P clones

In this section we consider the case of $P = N^\delta$, $0 < \delta < 1$ B clones. At $T_c = c/(1-k)$, we expect clones to be activated in a symmetric fashion. For $\delta < \gamma$ we expect symmetric solutions to be stable for all $T < T_c$, whereas for $\delta \geq \gamma$ we expect them to destabilise at low temperature. Without loss of generality, we can set $P = \alpha N^\gamma$ with $\alpha = 1$ for $\delta = \gamma$ and the cases $\delta < \gamma$ and $\delta > \gamma$ retrieved in the limits $\alpha \to 0$ and $\alpha \to \infty$ respectively.

To inspect the behaviour of the system near criticality, it is convenient to write the steady state equations in terms of the rescaled matrix $\hat{\mathbf{A}} = \mathbf{A}/(1-k)$, with eigenvalues

$$\mu_1 = 1 \ , \qquad deg(\mu_1) = P/2 \ , \tag{3.3.108}$$

$$\mu_2 = \frac{1+k}{1-k} \ , \qquad deg(\mu_2) = P/2 \ . \tag{3.3.109}$$

Taylor-expanding the steady state equations (3.3.85) for small m near T_c, introducing for convenience $\tilde{\beta} = \frac{\beta c}{1-k}$, we have

$$m_\mu \simeq \frac{N^\gamma}{c}\tilde{\beta}\sum_{\rho,\nu=1}^{P}(\hat{\mathbf{A}}^{-1})_{\rho\nu}m_\rho\langle\xi^\mu\xi^\nu\rangle_{\boldsymbol{\xi}} + \\ - \frac{\tilde{\beta}^3}{3}\frac{N^\gamma}{c}\Bigg[\sum_{\rho_1,\nu_1,\rho_2,\nu_2,\rho_3,\nu_3=1}^{P}\langle\xi^\mu\xi^{\nu_1}\xi^{\nu_2}\xi^{\nu_3}\rangle_{\boldsymbol{\xi}}m_{\rho_1}(\hat{\mathbf{A}}^{-1})_{\rho_1\nu_1}m_{\rho_2}(\hat{\mathbf{A}}^{-1})_{\rho_2\nu_2}m_{\rho_3}(\hat{\mathbf{A}}^{-1})_{\rho_3\nu_3}\Bigg] \tag{3.3.110}$$

and averaging over the disorder gives

$$m_\mu = \tilde{\beta}\sum_{\rho=1}^{P}(\hat{\mathbf{A}}^{-1})_{\rho\mu}m_\rho - \frac{\tilde{\beta}^3}{3}\sum_{\rho}(\hat{\mathbf{A}}^{-3})_{\rho\mu}m_\rho^3 - \tilde{\beta}^3\frac{c}{N^\gamma}\sum_{\rho=1}^{P}(\hat{\mathbf{A}}^{-1})_{\rho\mu}m_\rho\sum_{\tau,\pi=1}^{P}(\hat{\mathbf{A}}^{-2})_{\tau\pi}m_\pi^2 \ . \tag{3.3.111}$$

To make progress, we write $\mathbf{m} = \sum_{j=1}^{P} a_j\mathbf{v}^{(j)}$ as a linear combination of the

eigenvectors of $\hat{\mathbf{A}}^{-1}$, defined by $\hat{\mathbf{A}}^{-1}\mathbf{v}^{(j)} = \mu_j \mathbf{v}^{(j)}$, hence we get

$$\sum_{j=1}^{P} a_j v_\mu^{(j)} = \tilde{\beta} \sum_{j=1}^{P} a_j \mu_j v_\mu^{(j)} - \frac{\tilde{\beta}^3}{3} \sum_{i,j,k=1}^{P} a_i a_j a_k \mu_i \mu_j \mu_k v_\mu^{(i)} v_\mu^{(j)} v_\mu^{(k)} \\ -\tilde{\beta}^3 \frac{c}{N^\gamma} \sum_{i=1}^{P} a_i \mu_i v_\mu^{(i)} \sum_{\rho \neq \mu}^{P} \sum_{j,k=1}^{P} a_j a_k \mu_j \mu_k v_\rho^{(j)} v_\rho^{(k)} \ . \qquad (3.3.112)$$

We then split the sum into $j = 1, \ldots, P/2$ and $j = P/2+1, \ldots, P$, substitute the eigenvalues (3.3.108, 3.3.109), and equating linear terms we get

$$\sum_{j=1}^{P/2} a_j v_\mu^{(j)} (1 - \tilde{\beta}) + \sum_{j=P/2+1}^{P} a_j v_\mu^{(j)} \left(1 - \tilde{\beta} \frac{1-k}{1+k}\right) = 0, \quad \forall\ \mu \ . \qquad (3.3.113)$$

At $\tilde{\beta} = 1$ linear terms cancel only if $\sum_{j=P/2+1}^{P} a_j v_\mu^{(j)} = 0\ \forall\ \mu$. Since the eigenvectors of $\hat{\mathbf{A}}^{-1}$ are in the form

$$v_\mu^{(j)} = \frac{1}{\sqrt{2}} \left(\delta_{\mu,j} + \delta_{\mu,j+P/2}\right) \ , \qquad j = 1, .., P/2 \ , \qquad (3.3.114)$$

$$v_\mu^{(j)} = \frac{1}{\sqrt{2}} \left(\delta_{\mu,j} - \delta_{\mu,j+P/2}\right) \ , \qquad j = P/2+1, .., P \ , \qquad (3.3.115)$$

this implies $a_j = 0\ \forall\ j = \frac{P}{2} + 1, \ldots, P$. Substituting in (3.3.112) and using $m_\mu = \sum_{j=1}^{P} a_j v_\mu^{(j)} = \sum_{j=1}^{P/2} a_j v_\mu^{(j)}$, we obtain

$$m_\mu(-1 + \tilde{\beta}) - \frac{\tilde{\beta}^3}{3} m_\mu^3 - \tilde{\beta}^3 \frac{c}{N^\gamma} m_\mu \sum_{\rho \neq \mu}^{P} m_\rho^2 = 0 \ . \qquad (3.3.116)$$

This gives at $\tilde{\beta} = 1 + \epsilon$,

$$m_\mu \epsilon - \frac{1}{3} m_\mu^3 - \frac{c}{N^\gamma} m_\mu |\mathbf{m}|^2 = 0 \ , \qquad (3.3.117)$$

where $|\mathbf{m}|^2 = \sum_{\rho=1}^{P} m_\rho^2$, yielding $m_\mu = 0$ or

$$m_\mu^2 = 3\epsilon - \frac{3c}{N^\gamma} |\mathbf{m}|^2. \qquad (3.3.118)$$

Each clone activation m_μ depends on the whole vector $\mathbf{m}$, hence all clones will have the same activation strength $m_\mu \in (-m, +m, 0)$. Assuming the non-zero components are a fraction $\phi \leq \alpha$ of the total number of components P, we have $|\mathbf{m}|^2 = \alpha\phi N^\gamma m^2$ yielding

$$m^2 = \frac{3\epsilon}{1 + 3c\alpha\phi} \ . \tag{3.3.119}$$

We will see in the next section that while for $k = 0$ stability of symmetric mixtures is only ensured for $\phi = 1$, in the presence of idiotypic interactions, i.e. for $k \neq 0$, values of $\phi \neq 1$ are possible.

3.3.3.1 Linear stability analysis and phase diagram

In this section we study the stability of symmetric clonal activation, $\mathbf{m}^\star = m(1, \ldots, 1, 0, \ldots 0)$ with $n = \alpha\phi N^\gamma$ activated clones below criticality. To this end, we study the eigenvalues of the Jacobian of the dynamical system (3.3.84), which has a block structure with diagonal elements given for $\mu < n$ by

$$\mathcal{J}_{\mu\mu} = \beta c(\mathbf{A}^{-1})_{\mu\mu} \left[1 - \left\langle \tanh^2 \left(\beta c \left(\sum_{\rho=1}^{P} (\mathbf{A}^{-1})_{\rho\mu} m_\rho + \sum_{\rho,\lambda\neq\mu}^{P} \xi^\lambda (\mathbf{A}^{-1})_{\rho\lambda} m_\rho \right) \right) \right\rangle_{\boldsymbol{\xi}} \right] - 1, \tag{3.3.120}$$

and for $\mu > n$ by

$$\mathcal{J}_{\mu\mu} = \beta c(\mathbf{A}^{-1})_{\mu\mu} \left(1 - \left\langle \tanh^2 \left(\beta c \sum_{\rho,\lambda\neq\mu}^{P} \xi^\lambda (\mathbf{A}^{-1})_{\rho\lambda} m_\rho \right) \right\rangle_{\boldsymbol{\xi}} \right) - 1. \tag{3.3.121}$$

Off-diagonal elements for $\mu, \nu \leq n$ are

$$\mathcal{J}_{\mu\nu} = -\frac{\beta c(\mathbf{A}^{-1})_{\mu\nu}}{N^\gamma} \left\langle \tanh^2 \left(\beta c^2 \sum_{\rho,\lambda\neq\mu}^{P} \xi^\lambda (\mathbf{A}^{-1})_{\rho\lambda} m_\rho \right) \right\rangle_{\boldsymbol{\xi}} - 1 \tag{3.3.122}$$

and zero otherwise. Note that in the limit $N \to \infty$ the matrix becomes diagonal. At the symmetric fixed point, using $\sum_\rho (\mathbf{A}^{-1})_{\rho\lambda} = 1/(1-k)\ \forall\ \lambda$, we get the eigenvalues

$$\lambda_1 = \frac{\hat{\beta}}{1+k}\left(1 - \sum_z \tilde{P}_n(z)\tanh^2\left(\hat{\beta}m(1+z)\right)\right) - 1\ , \quad deg(\lambda_1) = n\ , \tag{3.3.123}$$

$$\lambda_2 = \frac{\hat{\beta}}{1+k}\left(1 - \sum_z \tilde{P}_n(z)\tanh^2\left(\hat{\beta}mz\right)\right) - 1\ , \quad deg(\lambda_2) = P - n\ , \tag{3.3.124}$$

where $\hat{\beta} = \beta c$ and the discrete noise distribution $\tilde{P}_n(z) = \left\langle \delta_{z,\sum_{\nu=1}^n \xi^\nu} \right\rangle_{\boldsymbol{\xi}}$. $\tilde{P}_n(z)$ can be written as a modified Bessel function of the first kind $I_z(x)$

$$\tilde{P}_n(z) = \int_{-\pi}^{+\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{\mathrm{i}\omega z} \prod_{\nu=1}^{n} \left\langle \mathrm{e}^{-\mathrm{i}\omega\xi^\nu} \right\rangle_{\boldsymbol{\xi}} \simeq \int_{-\pi}^{+\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{\mathrm{i}\omega z} \mathrm{e}^{-\alpha\phi c(1-\cos(\omega))} = \mathrm{e}^{-\alpha\phi c} I_z(\alpha\phi c)\ , \tag{3.3.125}$$

where we used $n = \phi P = \phi\alpha N^\gamma$ fraction of the total number of B clones. Close to criticality, by Taylor-expanding in powers of $\epsilon = \hat{\beta} - 1$ we obtain

$$\lambda_1 = \frac{-2\epsilon - k(1+3c\alpha\phi)}{(1+k)(1+3c\alpha\phi)}\ , \tag{3.3.126}$$

$$\lambda_2 = \frac{\epsilon - k(1+3c\alpha\phi)}{(1+k)(1+3c\alpha\phi)}\ . \tag{3.3.127}$$

In contrast to the case $k = 0$ analysed in sec. 2.4, where near T_c $\lambda_2 > 0$ and the symmetric activation was stable only when involving *all* clones, in the presence of idiotypic interactions, i.e. for $k \neq 0$, both eigenvalues are negative, showing the possibility of a partial clonal activation.

In the opposite limit, i.e. $T \to 0$ we get

$$\lambda_1 = \frac{\hat{\beta}}{1+k}\mathrm{e}^{-\alpha\phi c} I_1(\alpha\phi c) - 1\ , \tag{3.3.128}$$

$$\lambda_2 = \frac{\hat{\beta}}{1+k}\mathrm{e}^{-\alpha\phi c} I_0(\alpha\phi c) - 1\ . \tag{3.3.129}$$

For $\delta < \gamma$, $\alpha = 0$, hence using the properties of Bessel functions $I_1(0) = 0$ and $I_0(0) = 1$, we get $\lambda_1 < 0$ and $\lambda_2 > 0$. Since λ_2 has degeneracy $P - n$, symmetric mixtures $\mathbf{m} = m(1, \dots, 1)$ i.e. with $n = P$ will be stable for all $T < T_c$. In contrast, for $\delta = \gamma$ i.e. $\alpha \neq 0$, $\lambda_1 > 0$ as $I_1(x) > 0$ for any $x > 0$, meaning that symmetric mixtures are unstable at low temperature for any n. Finally, for $\delta > \gamma$ one has $\alpha \to \infty$ in the thermodynamic limit, so $\mathrm{e}^{-\alpha\phi c} I_z(\alpha\phi c) \simeq (2\pi\alpha\phi c)^{-1/2}$ $\forall$ z, and symmetric mixtures will gain stability at low temperature as $N \to \infty$.

Let us now compute the critical line in the phase diagram where symmetric mixtures become unstable. We note that $\lambda_2 > \lambda_1$ for $T \to 0$ and for $T \to T_c$, and deduce that $\lambda_2 > \lambda_1$ for all $T < T_c$. Hence, stability of n-mixtures $\mathbf{m} = m(1, \dots, 1, 0, \dots, 0)$ is given by the region where $\lambda_2 < 0$. In fig. 3.20 (left) we show the critical lines where λ_2 becomes zero in the space of scaled parameters $\hat{T} = T/c, \hat{\phi} = \phi c$ for different values of k and $\alpha = 1$. As k increases, the region where clones are activated with the same intensity widens. When λ_2 destabilises, symmetric mixtures can only be stable for $n = P$, i.e. for $\phi = 1$, in the region where $\lambda_1 < 0$. A contour plot of $\lambda_1 = 0$ for $\phi = 1$ in the $T - k$ plane is shown in fig. 3.20 (right).

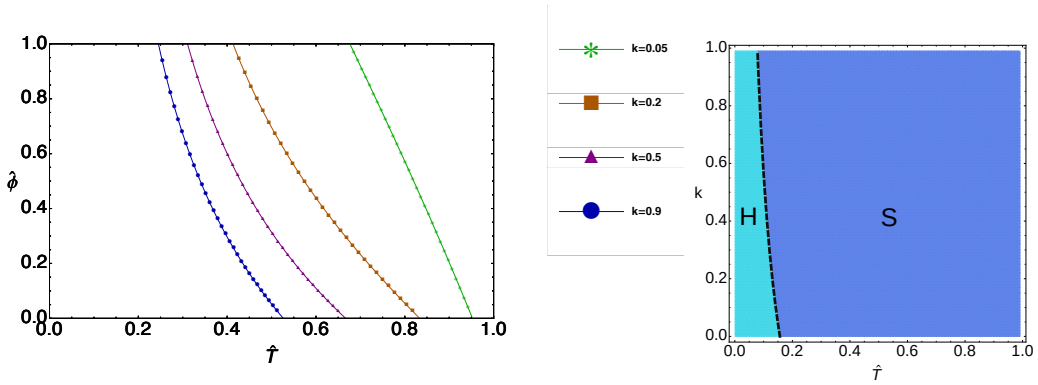

Figure 3.20: Left: Contour plot of $\lambda_2 = 0$ (obtained from (3.3.124)) for $\alpha = 1$, as a function of the scaled parameters $(\hat{T} = T/c, \hat{\phi} = \phi c)$, for different value of B-B interaction strength k. Increasing k the region where symmetric mixtures are stable (to the right of the critical line) becomes wider. Right: Contour plot of $\lambda_1 = 0$ (obtained from (3.3.123)) for $\alpha = \phi = 1$, as a function of the scaled temperature $\hat{T} = T/c$ and strength k of the idiotypic interactions.

3.4 Antigen effect

In this section, we investigate the effect of the antigens on the basal activity of the immune system. We will mainly focus on how the response of non-infected B clones decreases with the fraction of infected [2] ones, due to the interference of strongly activated B clones; non-infected clones will also be activated at a lower noise level, making the whole system unresponsive to new incoming pathogens (see fig. 3.23). In our model, antigens involved in B clones activation are modelled as external fields acting on coupled ferromagnetic systems. The introduction of antigens also induces hysteresis phenomena [63], which could explain short-term memory effect in the immune response [2, 64], even in the absence of memory cells.

We can further investigate the B-Antigen interactions analysing the stochastic process in (3.1.2)

$$\frac{\mathrm{d}b_\mu}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial b_\mu} + \chi_\mu(t) = \sum_{a=1}^{N_A} \psi_a \eta_a^\mu + N^{\gamma-1} \sum_{i=1}^{N} \xi_i^\mu \sigma_i - \frac{1}{2} \sum_{\nu=1}^{P} A_{\mu\nu} b_\nu + \chi_\mu(t) \ , \tag{3.4.130}$$

which governs the dynamics of B clones concentration. Denoting by ψ_μ the concentration of the antigen complementary to the b_μ clone, we need the b_μ concentration to increase when $\psi_\mu \neq 0$. Also, we assume that clone $b_{\mu+P/2}$, complementary to b_μ and thus carrying the same epitope as antigen ψ_μ, is inhibited by the presence of antigen ψ_μ, and we denote by k_1 the strength of this inhibition. This leads to a matrix for the B-Antigen interaction of the form

$$\eta_a^\mu = \delta_{\mu a} - k_1 \delta_{\mu,(a+P/2) \bmod P} \ . \tag{3.4.131}$$

According to (3.4.130), the concentration of the μ-th clone also increases in the

[2] By *infected clones* we mean B clones μ with complementary receptors to antigen A.

presence of excitatory signals received by T clones (second term), while the third suppressive term represents a B-B interactions acting as a threshold to be overcome to start the immune response. Assuming that the total B clones concentration is conserved on average leads to a relation between k and k_1

$$\frac{\mathrm{d}}{\mathrm{dt}}\sum_{\mu}\langle b_\mu\rangle = 0 \rightarrow \frac{1-k_1}{1-k} = \frac{1}{2}\frac{\sum_\nu\langle b_\nu\rangle}{\sum_a\langle\psi_a\rangle}\ , \tag{3.4.132}$$

which depends on the steady state concentrations of B cells and antigens. For simplicity we will set $k = k_1$, which leads to $\boldsymbol{\eta} = \mathbf{A}$. Different choices of the matrix

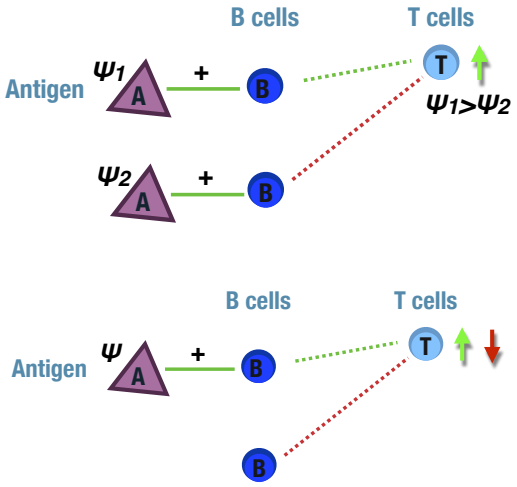

Figure 3.21: Top: effect of having $\mathbf{C} = \boldsymbol{\eta}\mathbf{A}^{-1} = \mathbf{1}$, in the presence of two antigens whose complementary B clones are signalled by different cytokines from the same T clone: the field acting on the T clone is $\psi_1 - \psi_2$ and T gets activated only if $\psi_1 > \psi_2$. Bottom: an illustration of the effect of having $\mathbf{C} = \mathbf{A}^{-1}$, which may result in an inhibitory signal on T clones from inactive B clones.

$\boldsymbol{\eta}$ or of the constants k, k_1 will lead to different forms of the matrix $\mathbf{C} = \boldsymbol{\eta}\mathbf{A}^{-1}$, however the choice $\mathbf{C} = \mathbf{1}$ seems to ensure the strongest excitatory signal to antigen-activated B clones. As an illustration, let us consider the simple case where we have just one antigen i.e. ψ_1: for the choice $\mathbf{C} = \mathbf{1}$ the field acting on the i-th T cell exciting the complementary clone b_1 is ψ_1, hence i receives an activation signal only from the clone B activated by the antigen. With two antigens present, whose complementary B cells are signalled by different cytokines

from the same T cell, the field acting on the latter is $\psi_1 - \psi_2$ and the T cell will get activated only if $\psi_1 > \psi_2$, as biologically desired (fig. 3.21, top). For $\mathbf{C} \neq \mathbf{1}$ there might be a negative interference on the desired signal. For example, for the alternative choice $\boldsymbol{\eta} = \mathbf{1}$, leading to $\mathbf{C} = \mathbf{A}^{-1}$, the field acting on T in the presence of a single virus $\psi_1 \neq 0$ would be $\frac{\psi_1}{1-k^2}[\xi_i^1 + k\xi^{1+\frac{P}{2}}]$, meaning that for $\xi^1 = +1$, $\xi^{1+\frac{P}{2}} = -1$ the T clone would receive, in addition to the excitatory signal from the B clone activated by the antigen, an inhibitory signal from a non-activated B clone (fig. 3.21, bottom). As a result the overall field and immune response will decrease. Similar arguments can be given for $\boldsymbol{\eta}$ defined in (3.4.131) with $k_1 \neq k$.

Therefore, for $\boldsymbol{\eta}$ defined in (3.4.131) with $k = k_1$, we have $\mathbf{C} = \boldsymbol{\eta}\mathbf{A}^{-1} = \mathbf{1}$, and each order parameter $m_\mu = M_\mu/c$ evolves according to

$$\frac{\mathrm{d}m_\mu}{\mathrm{d}t} = \frac{N^\gamma}{c}\left\langle \xi^\mu \tanh\left(\hat{\beta}\sum_{\nu=1}^{P}\xi^\nu\left(m_\nu + \psi_\nu\right)\right)\right\rangle_{\boldsymbol{\xi}} - m_\mu \ , \tag{3.4.133}$$

where we considered homogeneous promiscuities, $\mathbf{q} = c(1,\dots,1)$ and absent idiotypic interactions , i.e. $A_{\mu\nu} = \delta_{\mu\nu}$. Considering the steady state of (3.4.133), we can derive two sets of equations, for the activation m_μ, $\mu = 1,\dots,N_A$ of B clones complementary to the incoming viruses and those for the activation m_ν, $\nu = N_A+1,\dots,P$, for non-infected clones performing basal activity:

$$m_\mu = \left\langle \tanh\left(\hat{\beta}\left(m_\mu + \psi_\mu + \sum_{\rho\neq\mu}^{N_A}\psi_\rho\xi^\rho + \sum_{\lambda\neq\mu}^{P} m_\lambda\xi^\lambda\right)\right)\right\rangle_{\boldsymbol{\xi}} \ , \quad \mu = 1,\dots,N_A \ , \tag{3.4.134}$$

$$m_\nu = \left\langle \tanh\left(\hat{\beta}\left(m_\nu + \sum_{\rho\neq\nu}^{N_A}\psi_\rho\xi^\rho + \sum_{\lambda\neq\nu}^{P} m_\lambda\xi^\lambda\right)\right)\right\rangle_{\boldsymbol{\xi}} \ , \quad \nu = N_A+1,\dots,P \ . \tag{3.4.135}$$

3.4.1 No interference case, $A \ll N^\gamma$

In this section, we will consider the case with a number $N_A \ll N^\gamma$ of antigens. This means that we have few infected B clones in the system. We will show that the infected clones, which will be strongly activated, will not interfere with non-infected clones nor compromise the basal activity of the immune system in this regime. In this case, the equations (3.4.134, 3.4.135) can be written as

$$m_\mu = \int \mathrm{d}z P(z|\mathbf{m}) \tanh\left(\hat{\beta}\left(m_\mu + \psi_\mu + z\right)\right) , \qquad \mu = 1, \dots, N_A , \tag{3.4.136}$$

$$m_\nu = \int \mathrm{d}z P(z|\mathbf{m}) \tanh\left(\hat{\beta}\left(m_\nu + z\right)\right) , \qquad \nu = N_A + 1, \dots, P . \tag{3.4.137}$$

Here we introduced the noise distribution $P(z|\mathbf{m})$, which is the large N limit of $P_\mu(z|\{m_\rho, c\})$ defined in (3.2.34), and only depends on the vector $\mathbf{m}$ of basal activation (non-infected clones) [3]. The large N limit depends on the scaling of P with the system size N, as we will discuss in the following.

For $P \ll N^\gamma$, we have $P(z|\mathbf{m}) \equiv \left\langle \delta\left(z - \sum_{\nu=1}^{P} \xi^\nu m_\nu\right)\right\rangle_{\boldsymbol{\xi}} \simeq \delta(z)$ and the equations decouple, hence there is no interference between infected and non-infected clones

$$m_\mu = \tanh\left(\hat{\beta}\left(m_\mu + \psi_\mu\right)\right) , \qquad \mu = 1, \dots, N_A , \tag{3.4.138}$$

$$m_\nu = \tanh\left(\hat{\beta} m_\nu\right) , \qquad \nu = N_A + 1, \dots, P . \tag{3.4.139}$$

Introducing a field induces a discontinuous transition from zero to non-zero magnetisations and a hysteresis effect in the clonal activation of the infected clones [63]. In particular, considering the simple case with one antigenic field, i.e. $\boldsymbol{\psi} = \psi(1, 0, \dots, 0)$ we can compute the amplitude of the coercive field ψ^* [65].

[3] In principle, we should have introduced another noise distribution of the form $Q_\mu(z_\psi|\{\psi_\nu\}) = \left\langle \delta\left(z_\psi - \sum_{\rho\neq\mu}^{N_A}(m_\rho + \psi_\rho)\xi^\rho\right)\right\rangle_{\boldsymbol{\xi}} = \int \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{\mathrm{i}\omega z_\psi} \mathrm{e}^{\sum_{\rho\neq\mu}^{N_A} \frac{c}{N^\gamma} \cos(\omega(m_\rho + \psi_\rho))}$, but as $N_A \ll N^\gamma$ in the large N-limit it converges to $\delta(z_\psi)$.

For $\psi < \psi^*$ the systems shows bi-stability with two possible solutions for the magnetisation $\pm m^*$. This means that it could be trapped in a metastable state, e. g. with a positive field and a negative magnetisation. For $\psi > \psi^*$, the system has only one possible solution with the magnetisation concordant with the field.

This may explain immunological memory effects [2, 64], without the requirement of dedicated memory cells: after an infection, the responsive B clones retain a non-zero activation as the antigenic field decreases and the antigens' concentration goes to zero, and on a successive encounter with the same antigen they will provide a stronger and faster response. In addition, the fact that for antigenic fields smaller than ψ^* the system is not properly activated as discussed, may be interpreted as a *low-dose tolerance effect* [1, 66]: if the antigen is presented with a concentration below a certain threshold, the immune system does not produce a response. This mechanism is also exploited to prevent immune activations to self-cells.

The threshold ψ^* for the case with no interference and one antigenic field ψ, is easily found by combining the equation for the activation of infected clones

$$m = \tanh\left(\hat{\beta}(m+\psi)\right) \ , \tag{3.4.140}$$

with the condition on the slope of the curve $F(m) = \tanh\left(\hat{\beta}(m+\psi)\right)$ in $(\pm m^*, \pm\psi^*)$

$$1 = \hat{\beta}\left[1 - \tanh^2\left(\hat{\beta}(m+\psi)\right)\right] \ . \tag{3.4.141}$$

Combining (3.4.140) and (3.4.141) we obtain

$$m^* = \pm\sqrt{1 - 1/\hat{\beta}} \ , \tag{3.4.142}$$

$$\psi^* = \frac{1}{\hat{\beta}}\,\mathrm{atanh}(m^*) - m^* \ , \tag{3.4.143}$$

with $\hat{\beta} = \beta c$, similarly to what is obtained for the standard Curie-Weiss [65] (note the dependence on the dilution c in $\hat{T}$). The bifurcation diagram for ψ^* as

a function of $\hat{T}$ is shown fig. 3.22 : below $\hat{\beta}_c = 1$ the system is in the low-dose tolerance phase, while above it the B clone is correctly activated.

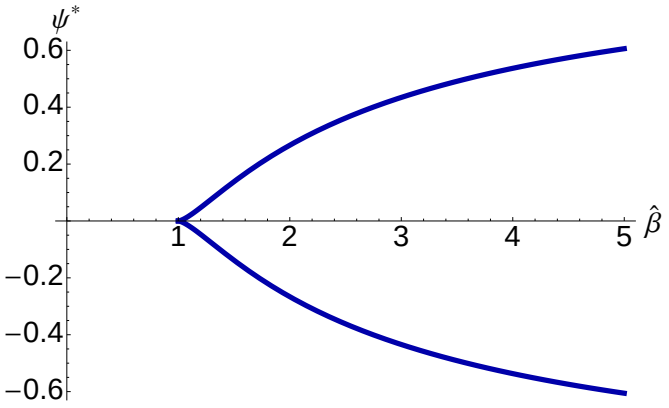

Figure 3.22: Plot of the coercive field ψ^* as a function of $\hat{\beta}$. The coercive field materialises above the value $\hat{\beta} = \beta c = 1$ of the inverse temperature.

For $P = \mathcal{O}(N^\gamma)$, $P(z|\mathbf{m}) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{\mathrm{i}z\omega} \mathrm{e}^{\sum_{\nu=1}^{P} \frac{c}{N^\gamma}[\cos(\omega c m_\nu)-1]}$ in (3.4.136, 3.4.137), has a finite width, and both the antigen-induced and basal activities reduce, due to clonal interference. Also, the hysteresis cycles become smaller. However, the presence of antigens does not affect the basal activity of non-infected clones $\nu = N_A + 1, \ldots, P$, as long as $N_A \ll N^\gamma$.

3.4.2 Increasing the number of infections

To investigate the interplay between infected and non-infected clones increasing the number of antigens, it is convenient to consider small fields ψ that allow us to work with continuous transitions. In particular, we are interested in understanding whether the system is able or not to fight $N_A = \mathcal{O}(N^\gamma)$ antigens in parallel, and how this compromises the basal activity. For simplicity, we set $P = N^\gamma$ and $N_A = \phi_1 N^\gamma$, with ϕ_1 denoting the fraction of infected clones, and we assume that all viruses have the same concentrations i.e. $\boldsymbol{\psi} = \psi(1, \ldots, 1, 0, \ldots, 0)$. This

leads to the steady state equations

$$m_\mu = \left\langle \tanh\left(\hat\beta\left(m_\mu + \psi + \sum_{\rho\neq\mu}^{N_A}(m_\rho+\psi)\xi^\rho + \sum_{\lambda=N_A+1}^{P} m_\lambda\xi^\lambda\right)\right)\right\rangle_{\boldsymbol{\xi}} \quad , \quad \mu = 1,\ldots,N_A \ ,$$
$$m_\nu = \left\langle \tanh\left(\hat\beta\left(m_\nu + \sum_{\rho=1}^{N_A}(m_\rho+\psi)\xi^\rho + \sum_{\lambda\neq\nu}^{P} m_\lambda\xi^\lambda\right)\right)\right\rangle_{\boldsymbol{\xi}} \quad , \quad \nu = N_A+1,\ldots,P \ . \tag{3.4.144}$$

In the small field limit, we can Taylor-expand (3.4.144) near $\hat\beta = 1+\epsilon$ and small (non-infected) m_ν, obtaining

$$m_\mu \simeq (1+\epsilon)m_\mu + \psi(1+\epsilon) - \frac{1}{3}\left((m_\mu+\psi)^3 + \frac{3c}{N^\gamma}(m_\mu+\psi)\sum_{\rho\neq\mu}(m_\rho+\psi)\right) . \tag{3.4.145}$$

For $\mu < N_A$ (infected clones) we have for $\psi \ll \epsilon$

$$m_\mu \simeq \frac{-\psi(1+\epsilon)}{\epsilon} \equiv m_1 \ , \qquad \mu = 1,\ldots,N_A \ . \tag{3.4.146}$$

Note that for fields $\psi \gg \epsilon$, $m_\mu = \mathcal{O}(1)$ and the transition is discontinuous. For $\nu > N_A$ ($\psi_\nu = 0$) we have

$$m_\nu \simeq (1+\epsilon)m_\nu - \frac{1}{3}\left(m_\nu^3 + \frac{3c}{N^\gamma}m_\nu\sum_{\rho\neq\nu}(m_\rho+\psi)^2\right) . \tag{3.4.147}$$

Hence, $m_\nu = 0$ is always a solution (non-infected clones may not be activated) together with

$$\begin{aligned} m_\nu^2 &\simeq 3\epsilon - \frac{3c}{N^\gamma}\sum_{\rho=1}^{P}(m_\rho+\psi)^2 = \\ &= 3\epsilon - \frac{3c}{N^\gamma}\left(\sum_{\rho=1}^{P} m_\rho^2 + \sum_{\rho=1}^{N_A}\left(2m_1\psi+\psi^2\right)\right) \quad , \quad \nu = N_A+1,\ldots,P \ . \end{aligned} \tag{3.4.148}$$

This shows that non-infected clones are symmetrically activated, each with intensity

$$m_2^2 = \frac{3\epsilon - 3c\phi_1\psi^2(2(1+\epsilon)/\epsilon + 1)}{1 + 3c\phi_2} = \frac{3\epsilon\left(1 - 2c\phi_1\left(\frac{\psi}{\epsilon}\right)^2\right)}{1 + 3c\phi_2}\,, \qquad (3.4.149)$$

where $\phi_2 \leq 1 - \phi_1$ is the fraction of active non-infected clones and m_1 is computed in (3.4.146).

For $\psi \ll \epsilon$, clonal activation close to criticality will have the form $\mathbf{m}^* = (m_1, \ldots, m_1, m_2, \ldots, m_2, 0, \ldots, 0)$. However, upon increasing the fraction ϕ_1 of infected clones or the antigenic field ψ, (3.4.149) shows that non-zero values of m_2 may become impossible and non-infected clones may get activated at a lower temperature (similarly to clones with smaller receptor promiscuity we dealt with in sec. 3.2). This is confirmed by numerical results in fig. 3.23, where the response

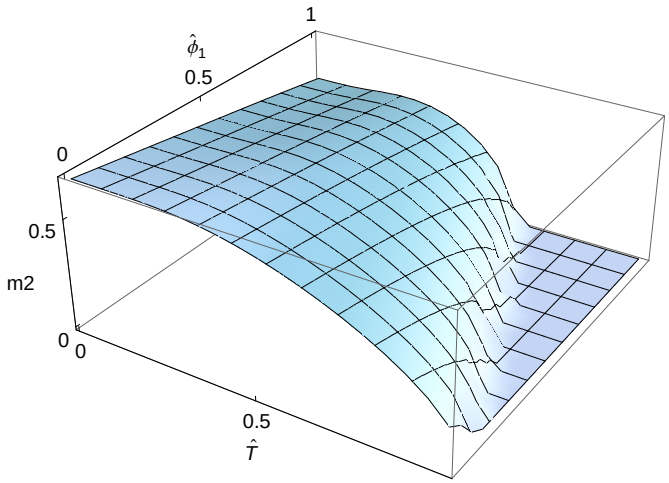

Figure 3.23: 3D plot of the activation m_2 of non-infected clones versus the fraction of infected clones $\hat{\phi}_1 = \phi_1 c$ and the scaled temperature $\hat{T} = T/c$, for fixed $\psi = 0.1$. Increasing $\hat{\phi}_1$ both the intensity and the critical activation temperature decrease, due to antigenic interference.

of non-infected B clones and their activation temperature decrease for increasing fractions of infected clones. This results in a reduced basal activity of the immune system, which is important to accomplish the homeostatic control of cell numbers [27, 53]. In fig. 3.24 we study the impact of antigen concentration on the critical

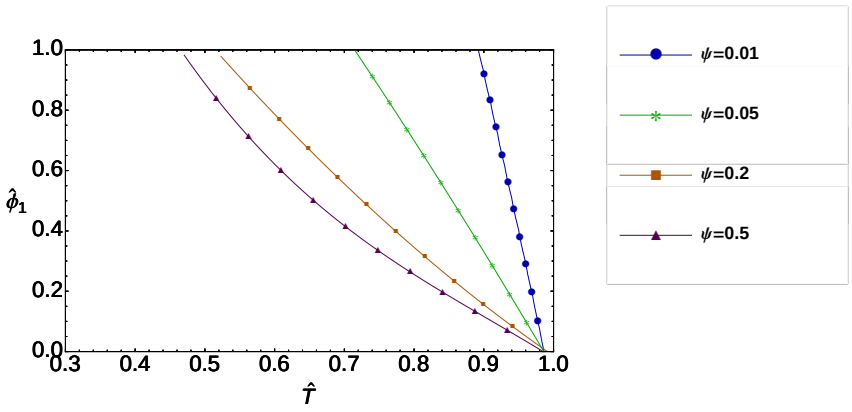

Figure 3.24: Critical line for the activation of non-infected clones in the space of scaled parameters $\hat{T} = T/c$, $\hat{\phi}_1 = \phi_1 c$, obtained from the condition $m_2 \neq 0$ (3.4.144), for different values of the antigenic field ψ. Increasing ψ, the region where non-infected clones receive signals shrinks.

temperature at which non-infected B clones become responsive, by plotting the critical temperature against the fraction of infected clones, for different values of antigen concentration. This shows that as the fraction of infected clones and the field increase, the basal activity is more and more compromised.

Subsequently, we inspect the stability region of the steady state $\mathbf{m}^* = (m_1, \dots, m_1, m_2, \dots, m_2, 0, \dots, 0)$ by looking at the eigenvalues of the Jacobian of the dynamical system (3.4.133). In the thermodynamic limit off-diagonal elements become negligible. We define by $n_1 = \phi_1 N^\gamma$ and $n_2 = \phi_2 N^\gamma$ the number of infected and non-infected clones respectively. Diagonal terms for $\mu < n_1 = \phi_1 N^\gamma$ are

$$\mathcal{J}_{\mu\mu} = \hat{\beta}\left(1 - \left\langle \tanh^2\left(\hat{\beta}\left(m_\mu + \psi + \sum_{\rho=n_1+1}^{n_2} m_\rho \xi^\rho + \sum_{\lambda=1}^{n_1}(m_\lambda + \psi)\xi^\lambda\right)\right)\right\rangle_{\boldsymbol{\xi}}\right) . \tag{3.4.150}$$

For $n_1 < \mu < n_2 = \phi_2 N^\gamma$ we have

$$\mathcal{J}_{\mu\mu} = \hat{\beta}\left(1 - \left\langle \tanh^2\left(\hat{\beta}\left(m_\mu + \sum_{\rho=n_1+1}^{n_2} m_\rho \xi^\rho + \sum_{\lambda=1}^{n_1}(m_\lambda + \psi)\xi^\lambda\right)\right)\right\rangle_{\boldsymbol{\xi}}\right) , \tag{3.4.151}$$

while for $\mu > n_2$

$$\mathcal{J}_{\mu\mu} = \hat{\beta}\left(1 - \left\langle \tanh^2\left(\hat{\beta}\left(\sum_{\rho=n_1+1}^{n_2} m_\rho \xi^\rho + \sum_{\lambda=1}^{n_1}(m_\lambda + \psi)\xi^\lambda\right)\right)\right\rangle_{\boldsymbol{\xi}}\right) . \tag{3.4.152}$$

Evaluating the Jacobian at the symmetric fixed point $\mathbf{m}^\star$ and introducing the discrete noise distribution $\tilde{P}_n(z) = \left\langle \delta_{z,\sum_{\nu=1}^n \xi^\nu}\right\rangle_{\boldsymbol{\xi}}$, supported over $z \in (-n, \ldots, n)$ gives

$$\lambda_1 = \hat{\beta}\left(1 - \sum_{z_1,z_2} \tilde{P}_{n_1}(z_1)\tilde{P}_{n_2}(z_2)\tanh^2\left(\hat{\beta}((\psi + m_1)(1+z_1) + m_2 z_2)\right)\right) - 1 , \quad deg(\lambda_1) = n_1 , \tag{3.4.153}$$

$$\lambda_2 = \hat{\beta}\left(1 - \sum_{z_1,z_2} \tilde{P}_{n_1}(z_1)\tilde{P}_{n_2}(z_2)\tanh^2\left(\hat{\beta}((\psi + m_1)z_1 + m_2(1+z_2))\right)\right) - 1 , \quad deg(\lambda_2) = n_2 , \tag{3.4.154}$$

$$\lambda_3 = \hat{\beta}\left(1 - \sum_{z_1,z_2} \tilde{P}_{n_1}(z_1)\tilde{P}_{n_2}(z_2)\tanh^2\left(\hat{\beta}((\psi + m_1)z_1 + m_2 z_2)\right)\right) - 1 , \quad deg(\lambda_3) = P - n_1 - n_2 . \tag{3.4.155}$$

The amplitudes m_1, m_2 follow from (3.4.146), (3.4.147)

$$m_1 = \sum_{z_1,z_2} \tilde{P}_{n_1}(z_1)\tilde{P}_{n_2}(z_2)\tanh\left(\hat{\beta}((\psi + m_1)(1+z_1) + m_2 z_2)\right) , \tag{3.4.156}$$

$$m_2 = \sum_{z_1,z_2} \tilde{P}_{n_1}(z_1)\tilde{P}_{n_2}(z_2)\tanh\left(\hat{\beta}((\psi + m_1)z_1 + m_2(1+z_2))\right) . \tag{3.4.157}$$

In fig. 3.25 (left) we show the critical lines $\lambda_1 = 0$, $\lambda_2 = 0$ and $\lambda_3 = 0$ in the space of scaled parameters $\hat{T} = T/c, \hat{\phi}_2 = \phi_2 c$. Lowering the temperature, λ_3

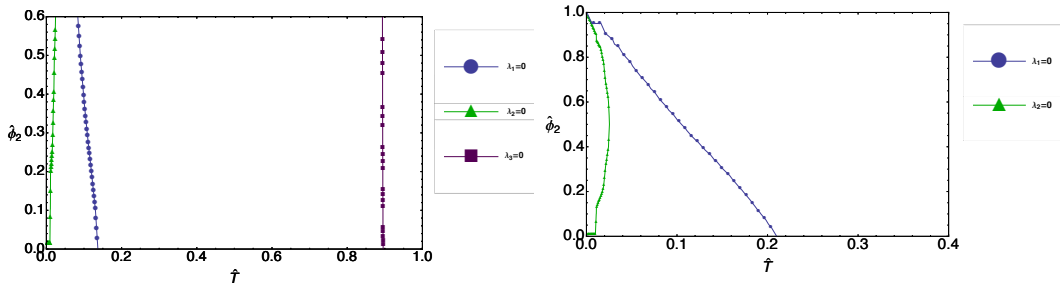

Figure 3.25: Left: Phase diagram in the space of scaled parameters $\hat{T} = T/c, \hat{\phi}_2 = \phi_2 c$ with $\phi_1 = 0.4$. Lines represent contours of $\lambda_1 = 0$ (circles), $\lambda_2 = 0$ (triangles) and $\lambda_3 = 0$ (squares). To the right of the line $\lambda_3 = 0$, solutions where non-infected clones are partially activated are stable. Lowering the temperature and crossing the line $\lambda_1 = 0$, the m_1 symmetric mixtures destabilise, implying that infected clones are hierarchically activated. Crossing the line $\lambda_2 = 0$, the m_2 symmetric mixtures destabilise, and non-infected clones are hierarchically activated. Right: Plots of $\lambda_1 = 0$ (circles), $\lambda_2 = 0$ (triangles) in the space $\hat{T} = T/c, \hat{\phi}_2 = \phi_2 c$ for $\phi_2 = 1 - \phi_1$.

is the first eigenvalue to destabilise, meaning that clonal activation will get in the form $\mathbf{m} = (m_1, \dots, m_1, m_2, \dots, m_2)$. Decreasing the temperature further, the system will first prioritise activation of infected clones, while keeping activation of non-infected clones symmetric, and later, at low temperature, will activate non-infected clones in a hierarchical fashion (see fig. 3.25, right). Near the critical temperature, symmetric mixtures with $\phi_2 \leq 1 - \phi_1$ are stable. To investigate the optimal value of ϕ_2, we calculate the free-energy as a function of ϕ_2 for fixed T, ϕ_1

$$F(\phi_1, \phi_2) = -\frac{1}{\beta} \sum_{z_1, z_2} \tilde{P}_{n_1}(z_1) \tilde{P}_{n_2}(z_2) \log(2 \cosh(\beta c(m_1 + \psi) z_1 + m_2 z_2)) + \\ + \frac{c}{2}(\phi_1 m_1^2 + \phi_2 m_2^2) \ , \tag{3.4.158}$$

which is minimal (fig. 3.26) at $\phi_2 = 1 - \phi_1$, implying that all non-infected clones will keep receiving signals.

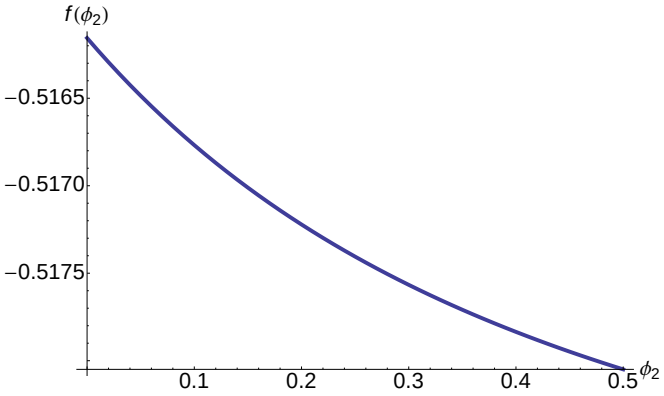

Figure 3.26: Free energy $f(\phi_2) = F(\phi_1 = 0.5, \phi_2)$ as a function of the fraction of active non-infected clones (ϕ_2) for $\phi_1 = 0.5$, $T = 0.7, c = 0.2, \psi = 0.05$.

3.5 Summary

In this chapter, we introduced new interesting mechanisms in the model, which the adaptive immune system exploits to control, sustain and manage immune responses. In table 3.1 we summarise them, highlighting their main effects on the system's behaviour. Our model is able to capture important collective features of the real immune system, such as the ability of simultaneously handling multiple infections, the dependence of B clones' activation on the number of receptors on clonal surface and the role of idiotypic interactions in enhancing parallel response to multiple infections.

One of the limitations of this model is that it works well only for some dilution regimes, in particular $\gamma < 1$. In the next chapter, we will analyse the regime $\gamma = 1$ (*finitely connected*) focusing on one particular effect introduced here, the idiotypic interactions. In addition, we will give formulae to compute the distributions of B clone sizes, which were integrated out from the beginning in the present approach, from the distribution of the overlaps. Being able to compute clonal distributions, indeed, constitutes a substantial step forward towards the comparisons with available data: nowadays these distributions can be directly measured exploiting new experimental techniques, such as high-throughput sequencing [67].

	Mathematical modelling	Summary of results
Receptor promiscuity	q_μ tuning patterns dilution in $\mathbb{P}(\boldsymbol{\xi}\|\mathbf{q})$	- B clones activation strength and T_c of activation increase with promiscuity. - Emergence of competition between activated clones.
B-B interactions	Interaction matrix $A_{\mu\nu}$	- Contribute to the overall network stability sustaining the signal propagation. - Widens the region where symmetric retrieval is stable (T_c increases with the B-B interaction strength k).
Antigen effect	External field ψ_μ	- Memory effect and hysteresis phenomena. - Increasing the number of infected clones the basal activity becomes more and more compromised.

Table 3.1: Summary of the main mechanisms introduced in the model, their mathematical implementation and the most interesting effects on the system's behaviour.

If I have seen further it is by standing on the shoulders of Giants.

Isaac Newton, Letter to Robert Hooke (15 February 1676)

4

Belief-propagation approach to the idiotypic network

Recently, following the development of experimental techniques such as high-throughput sequencing data [68] new interesting observables have become available, among which the clone size distribution. These techniques together with a complex computational analysis of the clonal genetic profile allow for the reconstruction of the clonal repertoire and the analysis of its properties in healthy and pathological conditions [67].

The clone size distribution encodes important information about the clones proliferation and activation. As already discussed, all clones in the immune system undergo a birth-death process, depending on the tasks to be performed: for example, in the presence of an antigen, complementary B clones start proliferating increasing the clone size, i.e. the number of clones sharing the same antigenic receptor. Monitoring the clone size distribution, thus, gives important information about the state of the immune system: for instance, it could help understand

how the organism reacts to antigens and vaccinations [12]. Experiments and analytical work are currently available for different cell types, both for humans and other species [12, 69, 70].

In this chapter, we will focus on the interacting system of B-T cells with idiotypic interactions in the *high storage* ($P = \alpha N$) and *finite connectivity* ($\gamma = 1$) regime. Using belief-propagation techniques [71] and tools imported from neural networks models [31, 72] and spin glasses [25, 32], we will be able to derive predictions for the B clone size distribution. The aim of this chapter is to understand the behaviour of the finitely connected system of B-T clones in the parameter space.

In particular, we want to study the effect of the idiotypic interactions on the activation properties of the system. We will show that they increase the stability of single B clones activation against biological noise while increasing the interference between different clones. We will derive analytically and via population dynamics simulations the critical line in the space of model parameters separating the region where each B clone would act independently from the others from the one where they feel a strong interference due to the presence of other active clones in the system.

The chapter is organised as follows. In sec. 4.1 we introduce the mathematical model, its players and assumptions. In sec. 4.2 we derive the recursive equations for the marginal distributions of T clones activation and we provide formulae for the overlap distributions of B clones activation and the B clone sizes distribution. We compute these distributions and inspect their transition from a single peaked to a bimodal distribution as the noise level is increased in the paramagnetic phase (sec. 4.4) and for ferromagnetic interactions (sec. 4.5), where one can proceed analytically. In sec. 4.6 we consider the most general case with disordered interactions.

4.1 The set-up

In this chapter, we study the system of interacting B-T clones via cytokines and subjected to idiotypic interactions. Here we will focus on the *high storage* regime where $P = \alpha N$ ($\delta = 1$) with a high degree of dilution $\gamma = 1$. The setting is similar to what has been introduced so far: T clones are binary variables $\sigma_i = \pm 1$, $i = 1, \dots, N$, where (± 1) depends on whether they are/are not secreting cytokines, while B clones are described by $b_\mu \in \mathbb{R}$, $\mu = 1, \dots, P$, characterising their log-concentrations. Moreover, they interact via cytokines ξ_i^μ, which are quenched i.i.d. random variables drawn from

$$\mathbb{P}(\xi_i^\mu = 1) = \mathbb{P}(\xi_i^\mu = -1) = \frac{c}{2N} , \quad \mathbb{P}(\xi_i^\mu = 0) = 1 - \frac{c}{N} , \tag{4.1.1}$$

with $c = \mathcal{O}(N^0)$, as required in the *finitely connected* regime. Both $\boldsymbol{\sigma}$'s and $\mathbf{b}$'s are random variable described at equilibrium at inverse noise level β [24], by the joint distribution $P(\boldsymbol{\sigma}, \mathbf{b}|\boldsymbol{\xi})$

$$P(\boldsymbol{\sigma}, \mathbf{b}|\boldsymbol{\xi}) = \frac{\mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{\sigma}, \mathbf{b}|\boldsymbol{\xi})}}{Z} , \tag{4.1.2}$$

with Hamiltonian

$$\mathcal{H}(\boldsymbol{\sigma}, \mathbf{b}|\boldsymbol{\xi}) = -\sum_{i,\mu=1}^{N,P} \xi_i^\mu \sigma_i b_\mu + \frac{1}{2} \sum_{\mu,\nu=1}^{P} b_\mu A_{\mu\nu} b_\nu , \tag{4.1.3}$$

where $\xi_i^\mu \in \{+1, -1, 0\}$ represent the interaction between T clone i and B clone μ, mediated by cytokines, which can be excitatory $(+1)$, inhibitory (-1) or absent (0). The matrix $\mathbf{A} = \{A_{\mu\nu}\}$, introduced in sec. 3.3, models the *idiotypic interactions*. As before, for the B-B interactions we use the following matrix

$$A_{\mu\nu} = \delta_{\mu\nu} - k\delta_{\mu,(\nu+P/2) \bmod P} , \tag{4.1.4}$$

with $k \in [0,1)$ representing the *strength* of idiotypic interactions. In the following, we will simplify the notation, dropping the explicit dependence on the ξ-interactions, in the Hamiltonian, which will be denoted by $\mathcal{H}(\boldsymbol{\sigma}, \mathbf{b})$, and the Boltzmann distribution, $P(\boldsymbol{\sigma}, \mathbf{b})$ (analogously for marginals).

Marginalising over the variables b_μ, as shown in sec. 3.3, the system can be described in terms of the distribution

$$P(\boldsymbol{\sigma}) \;=\; \frac{1}{Z'} \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{\sigma})} \;, \tag{4.1.5}$$

with effective Hamiltonian

$$\mathcal{H}(\boldsymbol{\sigma}) = -\frac{1}{2} \sum_{i,j=1}^{N} \sigma_i \sigma_j \sum_{\mu,\nu=1}^{P} \xi_i^\mu (\mathbf{A}^{-1})_{\mu\nu} \xi_j^\nu \tag{4.1.6}$$

and normalisation

$$Z' = \sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{\sigma})} \;. \tag{4.1.7}$$

This involves only interactions between T clones in the separable form $J_{ij} = \boldsymbol{\xi}_i^T \mathbf{A}^{-1} \boldsymbol{\xi}_j$, where $\boldsymbol{\xi}_i = (\xi_i^1, \ldots, \xi_i^P)^T$, and thus describes an associative network with diluted patterns $\{\boldsymbol{\xi}^\mu\}$. We can rewrite the Hamiltonian (4.1.6) as

$$\mathcal{H}(\boldsymbol{\sigma}) = -\frac{1}{2} \mathbf{m}^T(\boldsymbol{\sigma}) \mathbf{A}^{-1} \mathbf{m}(\boldsymbol{\sigma}) \tag{4.1.8}$$

in terms of the αN (non-normalised) order parameters $\mathbf{m} = (m_1, \ldots, m_P)^T$, where

$$m_\mu(\boldsymbol{\sigma}) = \sum_{i=1}^{N} \sigma_i \xi_i^\mu \tag{4.1.9}$$

quantifies the strength of the excitatory signal on B clone μ and thus its activation, and $\mathbf{m}^T$ denotes the transpose of $\mathbf{m}$. In the following we will show that important properties of the system are encoded in the system of interacting T clones only.

Experimentally, one has access to the B clone size distribution, rather than $P(\boldsymbol{\sigma})$. Hence, the relevant quantity to probe to compare the model behaviour with experimental data is

$$P(b) = \int \mathrm{d}\mathbf{b} \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}, \mathbf{b}) \frac{1}{P} \sum_{\mu=1}^{P} \delta(b - b_\mu) \ . \tag{4.1.10}$$

To this end, we rewrite (4.1.2) as

$$\begin{aligned} P(\boldsymbol{\sigma}, \mathbf{b}) &= \frac{1}{Z} \mathrm{e}^{-\frac{1}{2}\beta \mathbf{b}^T \mathbf{A} \mathbf{b} + \beta \mathbf{b}^T \mathbf{m}(\boldsymbol{\sigma})} & (4.1.11) \\ &= \frac{1}{Z'} P(\boldsymbol{\sigma}) P(\mathbf{b} | \mathbf{A}^{-1} \mathbf{m}(\boldsymbol{\sigma})) \ , & (4.1.12) \end{aligned}$$

where we have completed the square in (4.1.11) and introduced the joint distribution of the B clone sizes

$$P(\mathbf{b} | \mathbf{A}^{-1} \mathbf{m}(\boldsymbol{\sigma})) = \sqrt{\frac{\det \mathbf{A}}{(2\pi)^P}} \mathrm{e}^{-\frac{\beta}{2}(\mathbf{b} - \mathbf{A}^{-1}\mathbf{m}(\boldsymbol{\sigma}))^T \mathbf{A} (\mathbf{b} - \mathbf{A}^{-1}\mathbf{m}(\boldsymbol{\sigma}))} \ , \tag{4.1.13}$$

a multivariate Gaussian distribution. The marginal distribution for the μ-th B clone, obtained by integrating the above over all b_ν other than b_μ, will be Gaussian as well with average $(\mathbf{A}^{-1}\mathbf{m}(\boldsymbol{\sigma}))_\mu$ and variance $(\mathbf{A}^{-1})_{\mu\mu}$ i.e.

$$P(b_\mu | (\mathbf{A}^{-1}\mathbf{m}(\boldsymbol{\sigma}))_\mu) = \frac{1}{\sqrt{2\pi/\beta(1-k^2)}} \mathrm{e}^{-\frac{\beta(1-k^2)}{2}(b_\mu - (\mathbf{A}^{-1}\mathbf{m}(\boldsymbol{\sigma}))_\mu)^2} \ , \tag{4.1.14}$$

where we used $(\mathbf{A}^{-1})_{\mu\mu} = 1/(1-k^2) \ \forall \ \mu$. Hence, from (4.1.10) we get

$$P(b) = \frac{1}{Z'} \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) \frac{1}{P} \sum_{\mu=1}^{P} P(b | (\mathbf{A}^{-1}\mathbf{m}(\boldsymbol{\sigma}))_\mu) \ . \tag{4.1.15}$$

Upon defining $\tilde{\mathbf{m}}(\boldsymbol{\sigma}) = \mathbf{A}^{-1}\mathbf{m}(\boldsymbol{\sigma})$ we get

$$\begin{aligned} P(b) &= \frac{1}{Z'}\sum_{\tilde{m}}\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma})\frac{1}{P}\sum_{\mu=1}^{P}\delta(\tilde{m}-\tilde{m}_\mu(\boldsymbol{\sigma}))P(b|\tilde{m}) \\ &= \frac{1}{Z'}\int \mathrm{d}\tilde{m}P(\tilde{m})P(b|\tilde{m}) \ , \end{aligned} \tag{4.1.16}$$

where

$$P(\tilde{m}) = \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma})\frac{1}{P}\sum_{\mu=1}^{P}\delta(\tilde{m}-\tilde{m}_\mu(\boldsymbol{\sigma})) \tag{4.1.17}$$

and

$$P(b|\tilde{m}) = \frac{1}{\sqrt{2\pi/\beta(1-k^2)}}\mathrm{e}^{-\frac{\beta(1-k^2)}{2}(b-\tilde{m})^2} \ . \tag{4.1.18}$$

This shows that the distribution of B clone sizes $P(b)$ is readily determined once we know $P(\tilde{m})$, distribution of "rotated" overlaps $\tilde{\mathbf{m}}(\boldsymbol{\sigma})$ in the "marginalised" system involving only $\boldsymbol{\sigma}$ described by (4.1.8), where the $\mathbf{b}$ have been integrated out. Hence, the problem of computing B clone distributions, which are experimentally accessible in immunology, amounts to finding the overlap distribution in an associative memory with diluted and coupled patterns.

In the absence of idiotypic interactions, i.e. for uncoupled patterns, it is known that for ratios $\alpha = P/N$ not too large and below a critical value of the noise level T, non-zero values of the overlaps $\mathbf{m}$ emerge in the system, meaning that B clones are spontaneously activated [39]. The overlap distribution crosses-over from a unimodal distribution peaked at zero to a bimodal peaked at large values of the overlap [27]. In section 4.4.1 we will derive a tight bound on the noise level T at which this cross-over takes place. Interestingly, we will show that this coincides with the critical temperature at which the system undergoes a phase transition, from $\mathbf{m} = \mathbf{0}$ to $\mathbf{m} \neq \mathbf{0}$, in the regime of extremely diluted B-T interactions and sub-extensive number of B clones, analysed in sec. 3.3.

Since the B clones size distribution is a convolution of Gaussian distributions centred on the emerging values of the (rotated) overlaps, the phase where $\mathbf{m} \neq \mathbf{0}$ can be regarded as the healthy phase of the immune system, where cells numbers are sustained. This may provide a theoretical explanation for the need of a *basal* activity of the immune system, by which T clones send signals to B clones. This is experimentally observed, even in the absence of external pathogens, and is believed to be one of the mechanisms to accomplish a *homeostatic control* of cell numbers [53].

In a similar fashion, we can calculate the equilibrium concentration of any B clone μ and the activation of any T cell i via

$$\begin{aligned} P_{i\mu}(\sigma, b) &= \int \mathrm{d}\mathbf{b} \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}, \mathbf{b}) \delta(b - b_\mu) \delta_{\sigma,\sigma_i} \\ &= \frac{1}{Z} \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) \delta_{\sigma,\sigma_i} P(b|(\mathbf{A}^{-1}\mathbf{m}(\boldsymbol{\sigma}))_\mu) \\ &= \frac{1}{Z'} \int \mathrm{d}\tilde{m} P(b|\tilde{m}) \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) \delta_{\sigma,\sigma_i} \delta(\tilde{m} - \tilde{m}_\mu(\boldsymbol{\sigma})) \ , \end{aligned} \tag{4.1.19}$$

showing again that all the information about the physics of the system is encoded in the marginalised distribution $P(\boldsymbol{\sigma})$.

4.2 Factor graph representation

We can visualize our system of N T clones and P B clones as a bipartite graph, $\mathcal{G} = (\mathcal{T}, \mathcal{B})$, where T clones constitute the $\mathcal{T}$ party and B clones represent the $\mathcal{B}$ party, with $N = |\mathcal{T}|$ and $P = |\mathcal{B}|$. We indicate with $\partial\mu = \{i : \xi_i^\mu \neq 0\}$ the set of T clones connected to the B clone μ, and with $\partial i = \{\mu : \xi_i^\mu \neq 0\}$ the set of B clones connected to a particular clone i, i.e. $|\partial\mu|$ is the degree of a B clone μ and $|\partial i|$ is the degree of a T clone i in the bipartite graph $\mathcal{G}$ [73]. The bipartite graph is schematically represented in fig. 4.1.

In our analysis, we will consider random bipartite graph ensembles where

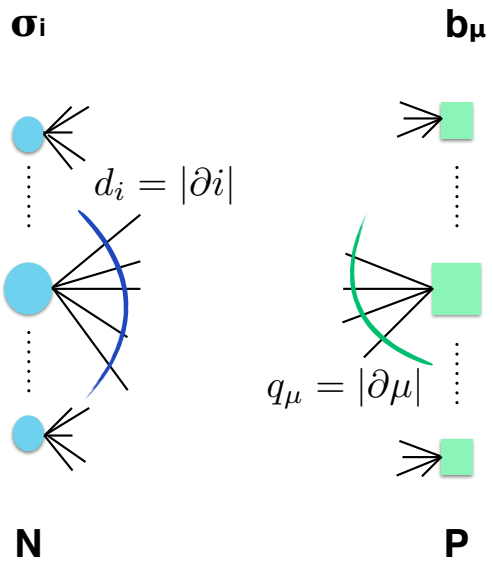

Figure 4.1: Schematic representation of the bipartite graph $\mathcal{G} = (\mathcal{T}, \mathcal{B})$ composed of P B clones and N T clones. $|\partial i| = d_i$ and $|\partial \mu| = q_\mu$ represent the degree of party $\mathcal{T}$ and $\mathcal{B}$ respectively.

the degrees $d_i = |\partial i|$ and $q_\mu = |\partial \mu|$ of the nodes in the two parties are drawn respectively from the distributions $P_d(d) = \frac{1}{N}\sum_{i=1}^{N} \delta_{d,|\partial i|}$, $P_q(q) = \frac{1}{P}\sum_{\mu=1}^{P} \delta_{q,|\partial \mu|}$, and links are i.i.d variables which take values $\xi_i^\mu \in \{+1, -1\}$. Conservation of links demands $\sum_{i=1}^{N} d_i = \sum_{\mu=1}^{P} q_\mu$, which gives, for large N, $N\langle d\rangle = P\langle q\rangle$, where averages are taken over $P_d(d)$ and $P_q(q)$. Different graph topologies have been considered, in the absence of idiotypic interactions, in [39], while here we mostly focus on regular graph topologies and on the role of B-B interactions.

In this section, we introduce a *factor graph representation* of the Boltzmann distribution for the marginalised system described by (4.1.8). In general, a factor graph is useful to represent multivariate distributions of n variables σ_i factorising into the product of local functions or factors, each dependent only on a subset of the n variables, as bipartite graphs. Hence, in the bipartite graph we can distinguish *variable nodes*, representing the σ_i's, and *factor nodes*, representing the local functions f_a, connected via an edge only if the variable σ_i is an argument of f_a. Let us consider, as an example, a simple factorising distribution over the

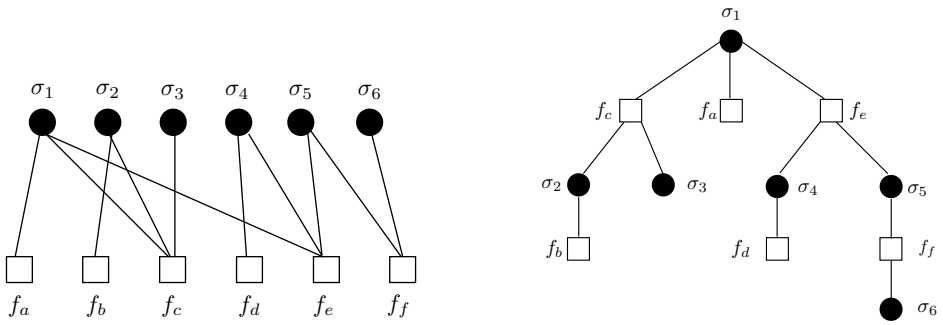

Figure 4.2: Factor graph representation of the distribution in (4.2.20). Black circles represent variable nodes and square represent factor nodes, edges are drawn if the variable σ_i is an argument of the factor f_a. We represent the distribution as a bipartite graph (left) and as a tree (right).

variables $\sigma_1, \ldots, \sigma_6$ of the form

$$P(\sigma_1, \ldots, \sigma_6) = f_a(\sigma_1) f_b(\sigma_2) f_c(\sigma_1, \sigma_2, \sigma_3) f_d(\sigma_4) f_e(\sigma_1, \sigma_4, \sigma_5) f_f(\sigma_5, \sigma_6) \ . \quad (4.2.20)$$

Using the above-given prescriptions, we can easily represent this distribution as a bipartite graph and, equivalently, as a tree, as shown in fig. 4.2.

For the general theory of factor representation we refer to [71, 75]. This representation will allow us to easily derive a set of recursive equations for marginal distributions of T clones activation and expressions for the distributions of local observables, which quantify B clones activation.

It is convenient to manipulate the $P(\boldsymbol{\sigma})$ in (4.1.5) to provide a suitable factor graph representation. First, we diagonalise the matrix $\mathbf{A}^{-1}$ transforming it via $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}^{-1}\mathbf{P}$, where $\mathbf{D}$ is the diagonal matrix constructed from the eigenvalues $\{\lambda_\mu\}_{\mu=1}^{P}$ of $\mathbf{A}^{-1}$ and $\mathbf{P}$ is the orthogonal matrix of eigenvectors. We can rewrite the Hamiltonian in terms of the transformed vector $\mathbf{v}(\boldsymbol{\sigma}) = \mathbf{P}^{-1}\mathbf{m}(\boldsymbol{\sigma})$

$$\mathcal{H}(\boldsymbol{\sigma}) = -\frac{1}{2}\mathbf{v}^T(\boldsymbol{\sigma})\mathbf{D}^{-1}\mathbf{v}(\boldsymbol{\sigma}) = -\frac{1}{2}\sum_\mu v_\mu^2(\boldsymbol{\sigma})\frac{1}{\lambda_\mu} \ . \quad (4.2.21)$$

Hence, the $P(\boldsymbol{\sigma})$ defined in (4.1.5) clearly factorizes over μ

$$P(\boldsymbol{\sigma}) = \frac{1}{Z'} \prod_{\mu=1}^{P} F_\mu(\boldsymbol{\sigma}) \ , \tag{4.2.22}$$

with normalisation Z' defined in (4.1.7) and factors

$$F_\mu(\boldsymbol{\sigma}) = \mathrm{e}^{\beta v_\mu^2(\boldsymbol{\sigma})/(2\lambda_\mu)} = \left\langle \mathrm{e}^{v_\mu(\boldsymbol{\sigma}) z \sqrt{\beta/\lambda_\mu}} \right\rangle_z \ , \tag{4.2.23}$$

where

$$\langle f(z) \rangle_z = \int_{-\infty}^{+\infty} \frac{\mathrm{d}z}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}z^2} f(z) \ . \tag{4.2.24}$$

We recall that $\mathbf{v}(\boldsymbol{\sigma}) = \mathbf{P}^{-1}\mathbf{m}(\boldsymbol{\sigma})$ and $\mathbf{P}$ is the orthogonal matrix whose columns are the eigenvectors $\{\mathbf{n}^\mu\}_{\mu=1}^P$ of the matrix $\mathbf{A}$ defined in (4.1.4). We can write

$$v_\mu(\boldsymbol{\sigma}) = \sum_{\nu=1}^{P} (\mathbf{P}^T)_{\mu\nu} m_\nu(\boldsymbol{\sigma}) = \sum_{\nu=1}^{P} n_\nu^\mu m_\nu(\boldsymbol{\sigma}) = \sum_{\nu=1}^{P} n_\nu^\mu \sum_{i \in \partial\nu} \xi_i^\nu \sigma_i \ , \tag{4.2.25}$$

where $\partial\nu = \{i : \xi_i^\nu \neq 0\}$ and n_ν^μ is the ν-component of the eigenvector $\mathbf{n}^\mu$ of $\mathbf{A}$, associated to the eigenvalue λ_μ. The eigenvalues of $\mathbf{A}$, each with degeneracy $P/2$, are

$$\lambda_1 = 1 - k \qquad \text{and} \qquad \lambda_2 = 1 + k \ . \tag{4.2.26}$$

The $P/2$ eigenvectors $\mathbf{n}^\mu$, $\mu = 1, \ldots, P/2$, associated to λ_1, have components

$$n_\nu^\mu = \frac{1}{\sqrt{2}} \left(\delta_{\nu,\mu} + \delta_{\nu,\mu+P/2} \right) \ , \quad \mu = 1, \ldots, P/2 \ , \tag{4.2.27}$$

whereas the $P/2$ eigenvectors $\mathbf{n}^\mu$, $\mu = P/2+1, \ldots, P$, associated to λ_2, have components

$$n_\nu^\mu = \frac{1}{\sqrt{2}} \left(\delta_{\nu,\mu} - \delta_{\nu,\mu+P/2} \right) \ , \quad \mu = P/2+1, \ldots, P \ . \tag{4.2.28}$$

Hence, there are only two contributions to the sum over ν in (4.2.25)

$$v_\mu(\boldsymbol{\sigma}) = \frac{1}{\sqrt{2}}\left(n_\mu^\mu \sum_{k\in\partial\mu}\xi_k^\mu\sigma_k + n_{\mu+P/2}^\mu \sum_{\ell\in\partial(\mu+P/2)}\xi_\ell^{\mu+P/2}\sigma_\ell\right). \quad (4.2.29)$$

Defining the scaled eigenvector $\mathbf{x}^\mu = \frac{\mathbf{n}^\mu}{\sqrt{2\lambda_\mu}}$, we obtain

$$F_\mu(\boldsymbol{\sigma}) = \left\langle \exp\left[\sqrt{\beta}z x_\mu^\mu \sum_{k\in\partial\mu}\xi_k^\mu\sigma_k + \sqrt{\beta}z x_{\mu+P/2}^\mu \sum_{\ell\in\partial(\mu+P/2)}\xi_\ell^{\mu+P/2}\sigma_\ell\right]\right\rangle_z . \quad (4.2.30)$$

We can proceed by splitting the product over μ in (4.2.22), to separate the contributions from $\mu \le P/2$ and $\mu > P/2$

$$\begin{aligned} P(\boldsymbol{\sigma}) &= \frac{1}{Z'}\prod_{\mu=1}^{P/2}\langle e^{v_\mu(\boldsymbol{\sigma})z\sqrt{\beta/\lambda_\mu}}\rangle_z \prod_{\nu=P/2+1}^{P}\langle e^{v_\nu(\boldsymbol{\sigma})z\sqrt{\beta/\lambda_\nu}}\rangle_z \\ &= \frac{1}{Z'}\prod_{\mu=1}^{P/2}\left\langle \exp\left[\sqrt{\beta}z\left(x_\mu^\mu \sum_{k\in\partial\mu}\xi_k^\mu\sigma_k + x_{\mu+P/2}^\mu \sum_{\ell\in\partial(\mu+P/2)}\xi_\ell^{\mu+P/2}\sigma_\ell\right)\right]\right\rangle \\ &\quad \times \prod_{\nu=P/2+1}^{P}\left\langle \exp\left[\sqrt{\beta}z\left(x_\nu^\nu \sum_{k\in\partial\nu}\xi_k^\nu\sigma_k + x_{\nu+P/2}^\nu \sum_{\ell\in\partial(\nu+P/2)}\xi_\ell^{\nu+P/2}\sigma_\ell\right)\right]\right\rangle . \end{aligned} \quad (4.2.31)$$

Setting $\nu = (\mu + P/2) \bmod P$ we have

$$\begin{aligned} P(\boldsymbol{\sigma}) &= \frac{1}{Z'}\prod_{\mu=1}^{P/2}\left\langle \exp\left[\sqrt{\beta}z\left(x_\mu^\mu \sum_{k\in\partial\mu}\xi_k^\mu\sigma_k + x_{\mu+P/2}^\mu \sum_{\ell\in\partial(\mu+P/2)}\xi_\ell^{\mu+P/2}\sigma_\ell\right)\right]\right\rangle \times \\ &\quad \times \prod_{\mu=1}^{P/2}\left\langle \exp\left[\sqrt{\beta}z\left(x_\mu^{\mu+P/2} \sum_{k\in\partial\mu}\xi_k^\mu\sigma_k + x_{\mu+P/2}^{\mu+P/2} \sum_{\ell\in\partial(\mu+P/2)}\xi_\ell^{\mu+P/2}\sigma_\ell\right)\right]\right\rangle . \end{aligned} \quad (4.2.32)$$

and combining the averages yields

$$P(\boldsymbol{\sigma}) = \frac{1}{Z'} \prod_{\mu=1}^{P/2} \left\langle \exp\left[\sqrt{\beta} \left(z_1 x_\mu^\mu + z_2 x_\mu^{\hat{\mu}} \right) \sum_{k \in \partial\mu} \xi_k^\mu \sigma_k + \sqrt{\beta} \left(z_1 x_{\hat{\mu}}^\mu + z_2 x_{\hat{\mu}}^{\hat{\mu}} \right) \sum_{\ell \in \partial\hat{\mu}} \xi_\ell^{\hat{\mu}} \sigma_\ell \right] \right\rangle_{z_1, z_2} , \tag{4.2.33}$$

where $\hat{\mu} = (\mu + P/2) \bmod P$ and $\langle \cdot \rangle_{z_1,z_2}$ denotes the average over the distribution $f(z_1) \cdot f(z_2)$ where f is a Gaussian distribution with zero mean and unit variance as in (4.2.24). For all $\mu = 1, \ldots, P/2$ we have

$$z_1 x_\mu^\mu + z_2 x_\mu^{\hat{\mu}} = \frac{1}{\sqrt{2(1-k)}} z_1 - \frac{1}{\sqrt{2(1+k)}} z_2 , \tag{4.2.34}$$

$$z_1 x_{\hat{\mu}}^\mu + z_2 x_{\hat{\mu}}^{\hat{\mu}} = \frac{1}{\sqrt{2(1-k)}} z_1 + \frac{1}{\sqrt{2(1+k)}} z_2 . \tag{4.2.35}$$

Applying to the variables z_1, z_2 the transformation $\mathbf{y} = \mathbf{X}\mathbf{z}$ with

$$\mathbf{X} = \begin{pmatrix} \frac{1}{\sqrt{2(1-k)}} & -\frac{1}{\sqrt{2(1+k)}} \\ \frac{1}{\sqrt{2(1-k)}} & \frac{1}{\sqrt{2(1+k)}} \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \qquad \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} , \tag{4.2.36}$$

we can rewrite (4.2.33) in terms of (y_1, y_2)

$$P(\boldsymbol{\sigma}) = \frac{1}{Z'} \prod_{\mu=1}^{P/2} \left\langle \exp\left[\sqrt{\beta} y_1 \sum_{k \in \partial\mu} \xi_k^\mu \sigma_k + \sqrt{\beta} y_2 \sum_{\ell \in \partial\hat{\mu}} \xi_\ell^{\hat{\mu}} \sigma_\ell \right] \right\rangle_{\mathbf{y}} . \tag{4.2.37}$$

In the above expression we used

$$\langle (\cdots) \rangle_{\mathbf{y}} = \sqrt{1-k^2} \int \frac{\mathrm{d}y_1 \mathrm{d}y_2}{2\pi} (\cdots) \mathrm{e}^{-\frac{1}{2}\mathbf{y}^T \mathbf{C}^{-1} \mathbf{y}} \tag{4.2.38}$$

and

$$\mathbf{C}^{-1} = (\mathbf{X}^{-1})^T \mathbf{X}^{-1} = \begin{pmatrix} 1 & -k \\ -k & 1 \end{pmatrix} . \tag{4.2.39}$$

We can finally write $P(\boldsymbol{\sigma})$ as a product of $P/2$ factors

$$P(\boldsymbol{\sigma}) = \frac{1}{Z'} \prod_{\mu=1}^{P/2} f_{\mu\hat{\mu}}(\{\sigma_k, k \in \partial\mu\}, \{\sigma_\ell, \ell \in \partial\hat{\mu}\}) \ , \tag{4.2.40}$$

each involving a pair of (complementary) clones $\mu, \hat{\mu}$

$$f_{\mu\hat{\mu}}(\{\sigma_k, k \in \partial\mu\}, \{\sigma_\ell, \ell \in \partial\hat{\mu}\}) = \left\langle \exp\left[\sqrt{\beta}\left(y_1 \sum_{k\in\partial\mu} \xi_k^\mu \sigma_k + y_2 \sum_{\ell\in\partial\hat{\mu}} \xi_\ell^{\hat{\mu}} \sigma_\ell\right)\right]\right\rangle_{\mathbf{y}} . \tag{4.2.41}$$

Each factor $f_{\mu\hat{\mu}}$ is a function of spins (T clones) $\{\sigma_k\}$ connected to the μ-th B clone and $\{\sigma_\ell\}$ connected to $\hat{\mu} = (\mu + P/2) \ \mathrm{mod} P$, which is the B clone complementary to μ. Via this representation, we note that the effect of the B-B interactions is equivalent to connecting together T clones signalling to complementary B clones.

4.3 Derivation of the cavity equations

The cavity equations and the belief-propagation approach are useful to recursively compute marginals of joint distributions of N variables $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)$ of the form (4.2.40) and the associated observables of the system, such as the overlap distribution quantifying B clones activation, as shown in sec. 4.3.1.

These types of factorised distributions can be graphically represented on bipartite graphs (as the graph shown in fig. 4.1) or *factor graphs* containing N variable nodes, representing the variables σ_i's, and $\frac{P}{2}$ factor nodes associated to the factors $f_{\mu\hat{\mu}}$ in (4.2.41). For sparse interactions $\{\xi_i^\mu\}$, our factor graph will be locally tree-like in the thermodynamic limit, with typical loop lengths diverging (logarithmically) with N. This will allow us to use the Bethe-Peierls approximation [71, 74] to find the cavity distributions in a recursive fashion. Compared to the naive algorithm to compute marginals of joint distributions, which would

sum over all the possible configurations for a total of order 2^N operations, computations on trees can reduce the complexity of the operations to just a linear growth in N.

To better understand this point, it is instructive to consider a simple example with $P(\boldsymbol{\sigma})$ of the form

$$P(\sigma_1, \dots, \sigma_N) = f_{1,2}(\sigma_1, \sigma_2) f_{2,3}(\sigma_2, \sigma_3) \cdots f_{N-1,N}(\sigma_{N-1}, \sigma_N) \ . \tag{4.3.42}$$

In a factor graph representation, this distribution corresponds to a chain, as each variable is common to at most two factors. To compute the marginal

$$P(\sigma_1) = \sum_{\sigma_2, \dots, \sigma_N} f_{1,2}(\sigma_1, \sigma_2) f_{2,3}(\sigma_2, \sigma_3) \cdots f_{N-1,N}(\sigma_{N-1}, \sigma_N) \ , \tag{4.3.43}$$

the naive strategy would require summing over 2^{N-1} configurations. A more efficient procedure would be to write the sums as follows

$$P(\sigma_1) = \sum_{\sigma_2} f_{1,2}(\sigma_1, \sigma_2) \left[\sum_{\sigma_3} f_{2,3}(\sigma_2, \sigma_3) \cdots \left[\sum_{\sigma_N} f_{N-1,N}(\sigma_{N-1}, \sigma_N) \right] \right] \ , \tag{4.3.44}$$

and start computing them from the innermost sum. The latter can be represented as a product between a 2×2 matrix $\mathbf{S}$, whose entries are $S_{\sigma\sigma'} = f(\sigma, \sigma')$ and a vector $\mathbf{v} = (1, 1)^T$. Each subsequent sum is again a matrix multiplication of $\mathbf{S}$ with the resulting vector from the previous step, for a total of 2^2 elementary multiplications at each step. Hence, to compute the marginal, we must perform $N - 1$ matrix multiplications for a total of $(N - 1)2^2 = \mathcal{O}(N)$ operations.

This reasoning can be easily applied not only to chains, but also to more complicated factor graphs as long as they are trees. For instance, we may consider the example given in the previous section where the distribution (4.2.20) can be represented as a tree (see fig. 4.2 (right)) and compute the marginal $P(\sigma_1) = \sum_{\sigma_2, \dots, \sigma_6} f_a(\sigma_1) f_b(\sigma_2) f_c(\sigma_1, \sigma_2, \sigma_3) f_d(\sigma_4) f_e(\sigma_1, \sigma_4, \sigma_5) f_f(\sigma_5, \sigma_6)$ in the more

efficient way

$$P(\sigma_1) = f_a(\sigma_1)\Big[\sum_{\sigma_2} f_b(\sigma_2)\Big[\sum_{\sigma_3} f_c(\sigma_1,\sigma_2,\sigma_3)\Big]\Big]\times$$
$$\times\Big[\sum_{\sigma_4} f_d(\sigma_4)\Big[\sum_{\sigma_5} f_e(\sigma_1,\sigma_4,\sigma_5)\Big[\sum_{\sigma_6} f_f(\sigma_5,\sigma_6)\Big]\Big]\Big]\,, \tag{4.3.45}$$

exploiting the factorisation properties of the distribution, as shown for the case of the chain.

Coming back to our distribution $P(\boldsymbol{\sigma})$ in (4.2.40) we denote with $P_{\mu\hat{\mu}}(\sigma_i)$ the marginals of σ_i, when the i-th clone is coupled to the factors $\mu\hat{\mu}$, which "measure" the dependence of the factors $\mu\hat{\mu}$ on the variable σ_i. The so-called *cavity marginals* $P_{\backslash\mu\hat{\mu}}(\sigma_i)$ of σ_i when coupled to all factors except μ and its complement $\hat{\mu}$ (also called *messages* from node i to factors $\mu,\hat{\mu}$), instead quantify the influence of all the neighbouring factors except $\mu,\hat{\mu}$ on σ_i.

Exploiting the tree structure of the graph, we notice that the message from factors $\mu,\hat{\mu}$ to node i, encoded in the marginals $P_{\mu\hat{\mu}}(\sigma_i)$ depends on the incoming messages from the nodes $j\neq i$ connected to the factors $\mu\hat{\mu}$. Equivalently, the message from node i to factors $\mu,\hat{\mu}$ (when the edge $i\to\mu\hat{\mu}$ has been removed) can be computed from the messages sent to node i from all the neighbouring factors $\rho,\hat{\rho}\neq\mu,\hat{\mu}$, as summarised in the scheme in fig. 4.3.

Hence, messages can be updated using *local* computations performed at each node of the factor graph, i.e. using the incoming messages from a previous iteration to compute the new outgoing ones. For a more detailed account of main features and limitations of the method we refer to the standard books [71, 75], while for applications to spin glasses we point the reader to [74].

We now proceed to the formal derivation of the recursive equations considering the tree schematically represented in fig. 4.4. We start by removing the factor μ and its complementary factor $\hat{\mu}$ and we calculate the partition function of the

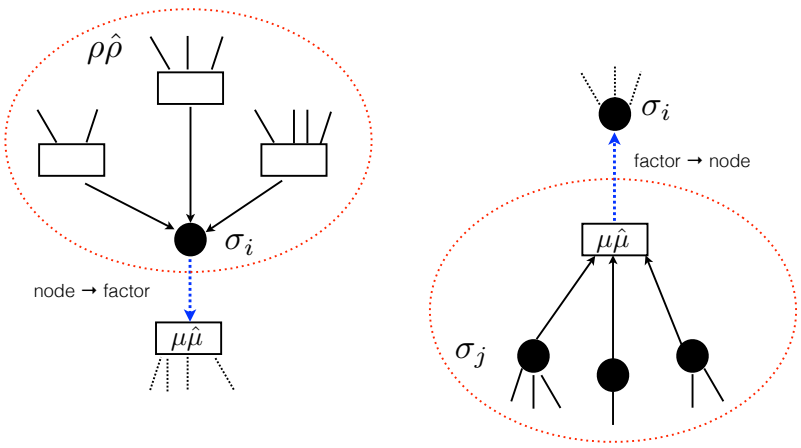

Figure 4.3: Scheme representing the incoming and outgoing node-factor and factor node messages used in local computations of marginal distributions on a tree.

tree $T_i(r)$ rooted in i with depth r, in the absence of the pair $\mu, \hat{\mu}$

$$Z^{(0)}_{i\mu\hat{\mu}} = \sum_{\{\sigma_{n\in T_i(r)}\}} \prod_{\nu\in T_i(r)\setminus\mu,\hat{\mu}} f_{\nu\hat{\nu}}(\{\sigma_{k\in\partial\nu}\},\{\sigma_{\ell\in\partial\hat{\nu}}\}) \ , \tag{4.3.46}$$

where the 0-index highlights that the root is the 0-th layer of the tree. We can rewrite the above in a more compact way as

$$Z^{(0)}_{i\mu\hat{\mu}} = \sum_{\{\sigma_{n\in T_i(r)}\}} \prod_{\nu\in T_i(r)\setminus\mu,\hat{\mu}} f_{\nu\hat{\nu}}(\{\sigma_{k\in\partial\nu\hat{\nu}}\}) \ , \tag{4.3.47}$$

where we introduced the notation $\partial\nu\hat{\nu} = \partial\nu \cup \partial\hat{\nu}$ and σ_k denotes any spin interacting with either ν or $\hat{\nu}$. We proceed by computing the same quantity as above but fixing the spin σ_i. This gives the constrained partition function $Z^{(0)}_{i\mu\hat{\mu}}(\sigma_i)$

$$Z^{(0)}_{i\mu\hat{\mu}}(\sigma_i) = \sum_{\{\sigma_{n\setminus i\in T_i(r)}\}} \prod_{\nu\in\partial i\setminus\mu,\hat{\mu}} \left[f_{\nu\hat{\nu}}(\sigma_i,\{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\}) \prod_{k\in\partial\nu\hat{\nu}\setminus i} \prod_{\rho\in T_k(r-1)} f_{\rho\hat{\rho}}(\{\sigma_{s\in\partial\rho\hat{\rho}\setminus k}\}) \right] \ , \tag{4.3.48}$$

which represents the un-normalised marginal distribution of σ_i in the absence of factors $\mu, \hat{\mu}$. Its normalised version gives the cavity distribution $P_{\setminus\mu\hat{\mu}}(\sigma_i)$, also defined as the "message" that site i sends to factors $\mu, \hat{\mu}$. In (4.3.48) we separated

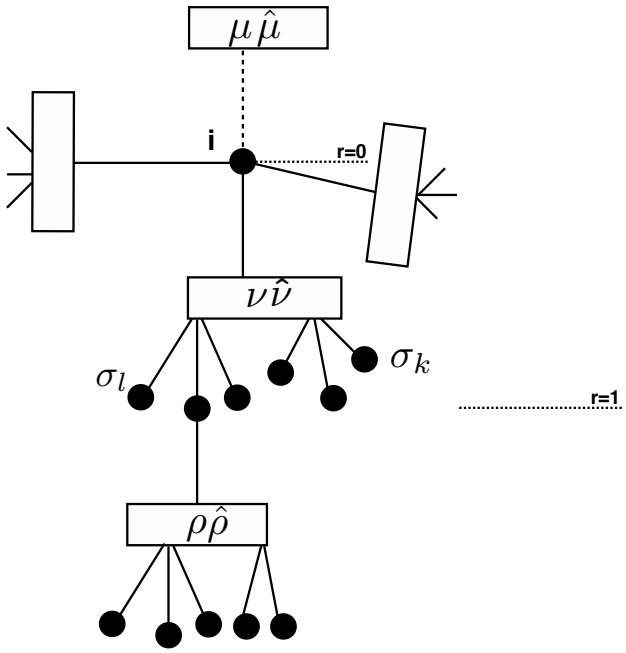

Figure 4.4: Schematic representation of the factor tree $T_i(r)$, i.e. the tree rooted in the spin σ_i with radius r. Rectangles represent factors $f_{\mu\hat{\mu}}$ defined in (4.2.41) and filled circle the spin variables the factors depend on.

the contributions of the factors directly attached to i from those in the rest of the tree, which are independent of σ_i. Rearranging the sums in the expression above we obtain

$$Z^{(0)}_{i\mu\hat{\mu}}(\sigma_i) = \prod_{\nu\in\partial i\setminus\mu\hat{\mu}} \left[\sum_{\{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\}} f_{\nu\hat{\nu}}(\sigma_i, \{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\}) \prod_{k\in\partial\nu\hat{\nu}\setminus i} \sum_{\{\sigma_{n\in T_k(r-1)\setminus k}\}} \prod_{\rho\in T_k(r-1)\setminus\nu,\hat{\nu}} f_{\rho\hat{\rho}}(\{\sigma_{s\in\partial\rho\hat{\rho}}\}) \right] . \tag{4.3.49}$$

The last sum in the expression represents exactly the partition function $Z^{(1)}_{k\nu\hat{\nu}}(\sigma_k)$ of the sub-tree rooted in k of depth $r-1$, where the spin σ_k has been fixed. This leads to the following recursive relation

$$Z^{(0)}_{i\mu\hat{\mu}}(\sigma_i) = \prod_{\nu\in\partial i\setminus\mu\hat{\mu}} \left[\sum_{\{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\}} f_{\nu\hat{\nu}}(\sigma_i, \{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\}) \prod_{k\in\partial\nu\hat{\nu}\setminus i} Z^{(1)}_{k\nu\hat{\nu}}(\sigma_k) \right] , \tag{4.3.50}$$

and more generally to the recursion for the (un-normalised) cavity distribution

at layer r in terms of those at layer $r+1$

$$Z^{(r)}_{i\mu\hat{\mu}}(\sigma_i) \;=\; \prod_{\nu\in\partial i\setminus\mu\hat{\mu}}\left[\sum_{\{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\}} f_{\nu\hat{\nu}}(\sigma_i,\{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\})\prod_{k\in\partial\nu\hat{\nu}\setminus i} Z^{(r+1)}_{k\nu\hat{\nu}}(\sigma_k)\right] . \tag{4.3.51}$$

The stationary solutions of the recursion equations then satisfy

$$Z_{i\mu\hat{\mu}}(\sigma_i) \;=\; \prod_{\nu\in\partial i\setminus\mu\hat{\mu}}\left[\sum_{\{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\}} f_{\nu\hat{\nu}}(\sigma_i,\{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\})\prod_{k\in\partial\nu\hat{\nu}\setminus i} Z_{k\nu\hat{\nu}}(\sigma_k)\right] . \tag{4.3.52}$$

Recalling the meaning of $Z_{i\mu\hat{\mu}}(\sigma_i)$ as the un-normalised marginal distribution of σ_i where factors $\mu,\hat{\mu}$ have been removed, one finally obtains the recursive equations for the cavity marginals

$$P_{\setminus\mu\hat{\mu}}(\sigma_i) \;=\; \prod_{\nu\in\partial i\setminus\mu\hat{\mu}} P_{\nu\hat{\nu}}(\sigma_i) , \tag{4.3.53}$$

$$P_{\nu\hat{\nu}}(\sigma_i) \;=\; \frac{1}{\mathcal{Z}_{\nu\hat{\nu}}}\sum_{\{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\}} f_{\nu\hat{\nu}}(\sigma_i,\{\sigma_{k\in\partial\nu\hat{\nu}\setminus i}\})\prod_{k\in\partial\nu\hat{\nu}\setminus i} P_{\setminus\nu\hat{\nu}}(\sigma_k) , \tag{4.3.54}$$

where $\mathcal{Z}_{\nu\hat{\nu}}$ ensures the normalisation of $P_{\nu\hat{\nu}}(\sigma_i)$. Iterating these equations until convergence, we obtain exact solutions on trees and approximately exact solutions on locally tree-like graph [71].

4.3.1 Distribution of overlaps

To study the activation properties of the system, i.e. its ability to retrieve stored patterns of clonal activation, encoded in the pattern overlaps $m_\mu = \sum_{k\in\partial\mu}\xi^\mu_k\sigma_k$, we look at the joint distribution of complementary clones activation

$$P_{\mu\hat{\mu}}(m,\hat{m}) = \langle\delta_{m,m_\mu(\boldsymbol{\sigma})}\delta_{\hat{m},m_{\hat{\mu}}(\boldsymbol{\sigma})}\rangle = \sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma})\delta_{m,m_\mu(\boldsymbol{\sigma})}\delta_{\hat{m},m_{\hat{\mu}}(\boldsymbol{\sigma})} . \tag{4.3.55}$$

Inserting (4.2.40) in (4.3.55) and splitting the sums over spins attached to factors $\mu, \hat{\mu}$ from the sums over spins in the rest of the (tree-like) graph (which give the cavity marginals in the absence of the factors $\mu, \hat{\mu}$) we obtain

$$P_{\mu\hat{\mu}}(m,\hat{m}) = \frac{\sum_{\{\sigma_k \in \partial\mu\hat{\mu}\}} \int \mathcal{D}\mathbf{y} \mathrm{e}^{\sqrt{\beta}(y_1 m_\mu(\boldsymbol{\sigma}) + y_2 m_{\hat{\mu}}(\boldsymbol{\sigma}))} \delta_{m,m_\mu(\boldsymbol{\sigma})} \delta_{\hat{m},m_{\hat{\mu}}(\boldsymbol{\sigma})} \prod_{k\in\partial\mu\hat{\mu}} P_{\backslash\mu\hat{\mu}}(\sigma_k)}{\sum_{\{\tilde{\sigma}_k \in \partial\mu\hat{\mu}\}} \int \mathcal{D}\mathbf{y} \mathrm{e}^{\sqrt{\beta}(y_1 m_\mu(\tilde{\boldsymbol{\sigma}}) + y_2 m_{\hat{\mu}}(\tilde{\boldsymbol{\sigma}})} \prod_{k\in\partial\mu\hat{\mu}} P_{\backslash\mu\hat{\mu}}(\tilde{\sigma}_k)} \,, \tag{4.3.56}$$

where $\mathcal{D}\mathbf{y} = \sqrt{1-k^2}\frac{\mathrm{d}y_1 \mathrm{d}y_2}{2\pi} \mathrm{e}^{-\frac{1}{2}\mathbf{y}^T \mathbf{C}^{-1}\mathbf{y}}$ and the cavity marginals $P_{\backslash\mu\hat{\mu}}(\sigma_k)$ must be computed recursively from equations (4.3.53, 4.3.54). Summing over m or $\hat{m}$, we obtain the marginalised distributions $P_{\mu\hat{\mu}}(m)$ and $P_{\mu\hat{\mu}}(\hat{m})$ respectively. Marginalised distributions are particularly useful to monitor the single-cluster behaviour varying the parameters T, k.

Finally, to compute the B clone size distribution $P(b)$ given in (4.1.16), we need the distribution of the rotated overlaps $\tilde{m}_\mu = \frac{m_\mu}{1-k^2} + \frac{k m_{\hat{\mu}}}{1-k^2}$

$$P_{\mu\hat{\mu}}(\tilde{m}) = \sum_{m,\hat{m}} P_{\mu\hat{\mu}}(m,\hat{m}) \delta\left(\tilde{m} - \frac{m}{1-k^2} - \frac{k\hat{m}}{1-k^2}\right). \tag{4.3.57}$$

In the following sections, we will solve the recursive equations (4.3.53, 4.3.54) in two cases that can be treated analytically, namely the paramagnetic phase (sec. 4.4) and the case with ferromagnetic interactions (sec. 4.5) and we will compute the overlaps and clone sizes distributions in different regimes of the model parameters. We will also specialise them for a particular graph distribution, focusing mainly on the regular graph topology. In sec. 4.6 we numerically solve the recursive equations for the general case with disordered interactions using a population dynamics algorithm, described in appendix A.2.

4.4 Paramagnetic phase

One can easily see that $P_{\setminus\mu\hat{\mu}}(\sigma_k) = 1/2 \ \forall \ \mu, \hat{\mu}, k$ is always a solution of the recursive equations (4.3.53, 4.3.54). We refer to this solution as the paramagnetic phase, where spins have probability 1/2 to be ± 1. In this phase, the distribution of overlaps (4.3.56) for the clusters $\mu, \hat{\mu}$ simplifies to

$$P_{\mu\hat{\mu}}(m, \hat{m}) = \frac{\sum_{\{\sigma_k \in \partial\mu\hat{\mu}\}} \int \mathcal{D}\mathbf{y} \mathrm{e}^{\sqrt{\beta}(y_1 m_\mu(\boldsymbol{\sigma}) + y_2 m_{\hat{\mu}}(\boldsymbol{\sigma}))} \delta_{m, m_\mu(\boldsymbol{\sigma})} \delta_{\hat{m}, m_{\hat{\mu}}(\boldsymbol{\sigma})}}{\sum_{\{\tilde{\sigma}_{k \in \partial\mu\hat{\mu}}\}} \int \mathcal{D}\mathbf{y} \mathrm{e}^{\sqrt{\beta}(y_1 m_\mu(\tilde{\boldsymbol{\sigma}}) + y_2 m_{\hat{\mu}}(\tilde{\boldsymbol{\sigma}}))}} \ , \quad (4.4.58)$$

thus losing the dependence on the interaction with the other clusters in the graph through the cavity marginals $P_{\setminus\mu\hat{\mu}}(\sigma_k)$. This means that in the paramagnetic phase each cluster in the graph behaves as if it were in isolation. Hence, analysing (4.4.58) and the other related overlap distribution in this phase, we can obtain information about *single clusters* activation.

We use the delta constraints to remove the spin dependence from the exponential in the numerator, then we perform the integration over y_1, y_2 and the sum over the spins

$$\begin{aligned}
&\sum_{\{\sigma_{k\in\partial\mu}, \sigma_{\ell\in\partial\hat{\mu}}\}} \delta_{m, m_\mu(\boldsymbol{\sigma})} \delta_{\hat{m}, m_{\hat{\mu}}(\boldsymbol{\sigma})} = \\
&= \frac{1}{4\pi^2} \int \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{e}^{\mathrm{i}x_1 m + \mathrm{i}x_2 \hat{m}} \sum_{\{\sigma_{k\in\partial\mu}\}} \mathrm{e}^{-\mathrm{i}x_1 \sum_{k\in\partial\mu} \xi_k^\mu \sigma_k} \sum_{\{\sigma_{\ell\in\partial\hat{\mu}}\}} \mathrm{e}^{-\mathrm{i}x_2 \sum_{\ell\in\partial\hat{\mu}} \xi_\ell^{\hat{\mu}} \sigma_\ell} \\
&= \frac{1}{4\pi^2} \int \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{e}^{\mathrm{i}x_1 m + \mathrm{i}x_2 \hat{m}} (2\cos x_1)^{|\partial\mu|} (2\cos x_2)^{|\partial\hat{\mu}|} \\
&= \sum_{r=0}^{|\partial\mu|} \binom{|\partial\mu|}{r} \delta_{m, 2r - |\partial\mu|} \sum_{j=0}^{|\partial\hat{\mu}|} \binom{|\partial\hat{\mu}|}{j} \delta_{\hat{m}, 2j - |\partial\hat{\mu}|} \ , \qquad (4.4.59)
\end{aligned}$$

where we used the Fourier representation of the Kronecker δ's, the parity of the cosine function to drop the ξ's from its argument, the binomial expansion of the powers of cosine and finally carried out the integrations. This leads us to the

discrete distribution

$$P_{\mu\hat{\mu}}(m,\hat{m}) = \frac{1}{\mathcal{Z}_{\mu\hat{\mu}}} \mathrm{e}^{\frac{\beta}{2(1-k^2)}(m^2+2km\hat{m}+\hat{m}^2)} \sum_{\ell=0}^{|\partial\mu|} \binom{|\partial\mu|}{\ell} \delta_{m,2\ell-|\partial\mu|} \sum_{j=0}^{|\partial\hat{\mu}|} \binom{|\partial\hat{\mu}|}{j} \delta_{\hat{m},2j-|\partial\hat{\mu}|} \ , \tag{4.4.60}$$

where

$$\mathcal{Z}_{\mu\hat{\mu}} = \sum_{\ell=0}^{|\partial\mu|} \binom{|\partial\mu|}{\ell} \sum_{j=0}^{|\partial\hat{\mu}|} \binom{|\partial\hat{\mu}|}{j} \mathrm{e}^{\frac{\beta}{2(1-k^2)}[(2\ell-|\partial\mu|)^2+2k(2\ell-|\partial\mu|)(2j-|\partial\hat{\mu}|)+(2j-\partial\hat{\mu})^2]} \tag{4.4.61}$$

ensures normalization of (4.4.60) over the support $m \in \{-|\partial\mu|, |\partial\mu|+2, \ldots, |\partial\mu|-2, |\partial\mu|\}$ and similarly for $\hat{m}$.

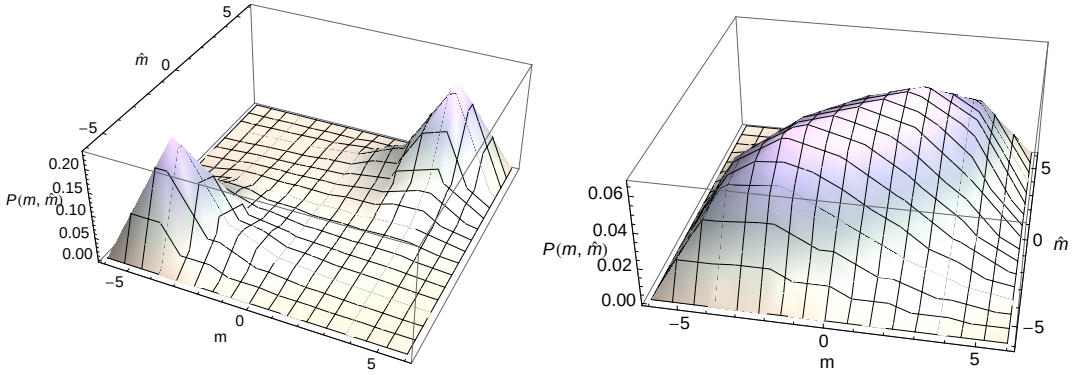

Figure 4.5: 3D plot of the joint distribution of complementary B-clones overlaps $P(m,\hat{m}|K)$ (4.4.65) for $T = 5$ (left) and $T = 10$ (right) and $k = 0.5$. The distribution is computed for a regular graph with degree $K = \hat{K} = 4$ in the paramagnetic phase. Note that here and elsewhere the support of the distribution is discrete, $m, \hat{m} \in \{-4, -2, 0, 2, 4\}$, and a continuous interpolating function has been used to guide the eye.

It is worth stressing that in the paramagnetic phase the overlap distribution, from which our observables of interest can be derived, does not depend on the nature of the interactions $\{\xi_i^\mu\}$, and $\mathcal{Z}_{\mu\hat{\mu}}$ depends on $\mu, \hat{\mu}$ only through $|\partial\mu|, |\partial\hat{\mu}|$.

The distribution of rotated overlap follows as

$$P_{\mu\hat{\mu}}(\tilde{m}) = \sum_{m,\hat{m}} \frac{\mathrm{e}^{\frac{\beta}{2(1-k^2)}(m^2+2km\hat{m}+\hat{m}^2)}}{\mathcal{Z}'_{\mu\hat{\mu}}} \sum_{\ell=0}^{|\partial\mu|} \binom{|\partial\mu|}{\ell} \delta_{m,2\ell-|\partial\mu|} \times \\ \times \sum_{j=0}^{|\partial\hat{\mu}|} \binom{|\partial\hat{\mu}|}{j} \delta_{\hat{m},2j-|\partial\hat{\mu}|} \delta\left(\tilde{m} - \frac{m}{1-k^2} - \frac{k\hat{m}}{1-k^2}\right) . \quad (4.4.62)$$

Averaging eq. (4.4.60) and (4.4.62) over graphs with factor degree distribution $P_q(\kappa)$ we obtain respectively

$$P(m,\hat{m}|P_q) = \sum_{\kappa,\tilde{\kappa}\geq 1} P_q(\kappa) P_q(\tilde{\kappa}) \frac{1}{\mathcal{Z}_{\kappa,\tilde{\kappa}}} \mathrm{e}^{\frac{\beta}{2(1-k^2)}(m^2+2km\hat{m}+\hat{m}^2)} \sum_{\ell=0}^{\kappa} \binom{\kappa}{\ell} \delta_{m,2l-\kappa} \sum_{j=0}^{\tilde{\kappa}} \binom{\tilde{\kappa}}{j} \delta_{\hat{m},2j-\tilde{\kappa}} \quad (4.4.63)$$

and

$$P(\tilde{m}|P_q) = \sum_{\kappa,\tilde{\kappa}\geq 1} P_q(\kappa) P_q(\tilde{\kappa}) \sum_{m,\hat{m}} \frac{\mathrm{e}^{\frac{\beta}{2(1-k^2)}(m^2+2km\hat{m}+\hat{m}^2)}}{\mathcal{Z}'_{\kappa,\tilde{\kappa}}} \sum_{\ell=0}^{\kappa} \binom{\kappa}{\ell} \delta_{m,2\ell-\kappa} \times \\ \times \sum_{j=0}^{\tilde{\kappa}} \binom{\tilde{\kappa}}{j} \delta_{\hat{m},2j-\tilde{\kappa}} \delta\left(\tilde{m} - \frac{m}{1-k^2} - \frac{k\hat{m}}{1-k^2}\right) . \quad (4.4.64)$$

In addition, we can obtain the marginalised distributions $P(m|P_q) = \sum_{\hat{m}} P(m,\hat{m}|P_q)$ and $P(\hat{m}|P_q) = \sum_m P(m,\hat{m}|P_q)$. In the following we consider the regular graph case, i.e. $P_q(\kappa) = \delta_{\kappa,K}$ for which the distributions reduce to

$$P(m,\hat{m}|K) = \frac{\mathrm{e}^{\frac{\beta}{2(1-k^2)}(m^2+2km\hat{m}+\hat{m}^2)}}{\mathcal{Z}} \sum_{\ell=0}^{K} \binom{K}{\ell} \delta_{m,2\ell-K} \sum_{j=0}^{K} \binom{K}{j} \delta_{\hat{m},2j-K} \quad (4.4.65)$$

and

$$P(\tilde{m}|K) = \sum_{m,\hat{m}} \frac{\mathrm{e}^{\frac{\beta}{2(1-k^2)}(m^2+2km\hat{m}+\hat{m}^2)}}{\mathcal{Z}'} \sum_{\ell=0}^{K} \binom{K}{\ell} \delta_{m,2\ell-K} \times \\ \times \sum_{j=0}^{K} \binom{K}{j} \delta_{\hat{m},2j-K} \delta\left(\tilde{m} - \frac{m}{1-k^2} - \frac{k\hat{m}}{1-k^2}\right) , \quad (4.4.66)$$

where $\mathcal{Z}, \mathcal{Z}'$ are the normalising constants. Analogously, we can define the marginalised distribution for a regular graph topology, which we will indicate with $P(m|K), P(\hat{m}|K)$.

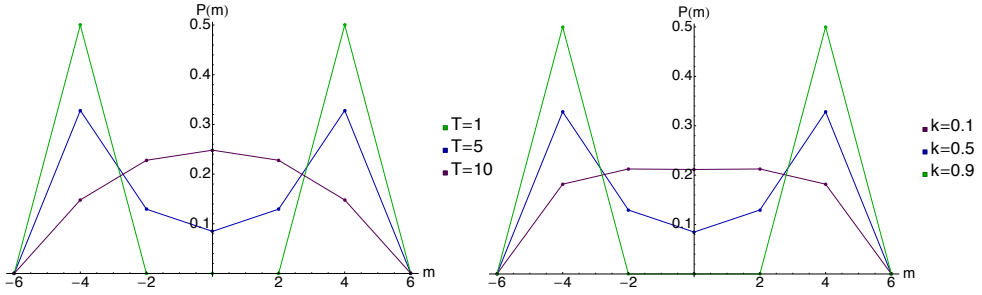

Figure 4.6: Plot of $P(m|K)$ in the paramagnetic phase for a regular graph with degree $K = 4$. Left panel: $P(m|K)$ for different temperatures $T = 1, 5, 10$ at fixed $k = 0.5$. Right panel: $P(m|K)$ for different B-B interaction strengths $k = 0.1, 0.5, 0.9$ at fixed $T = 5$. Note that the support of the distribution is discrete (markers), $m \in \{-4, -2, 0, 2, 4\}$.

Our aim is now to investigate the system's behaviour varying the parameters T and k. To this end, we first we plot $P(m, \hat{m}|K)$ in figure 4.5 for different values of the noise level T. At high temperature, the distribution is peaked around zero, suggesting that in the absence of an antigenic field there is no clonal expansion. Decreasing the temperature, the overlap distribution becomes peaked at equal values of $m = \hat{m} = \pm K$: this means that pairs of complementary clones are likely to receive the same signals, either both excitatory or both inhibitory. The system will fluctuate from one peak to the other with a timescale τ proportional to the exponential of the free-energy barrier between $(m, \hat{m}) = \pm(K, K)$ and $(0, 0)$, hence exponentially large in the finite size K of the two clones $\tau \sim \mathrm{e}^{\beta K^2/(1-k)}$. It follows that any clone that is initially expanded, will eventually undergo a contraction in the absence of an antigen, over a typical timescale that increases with the size K and the strength k of the idiotypic interactions. Hence, one of the roles of idiotypic interactions is to prolong the short-term memory of the encountered antigen in the system.

To better understand the dependence of the activation on the system's parameters, we plot $P(m|K)$ in fig. 4.6 for different values of the temperature T

and of the B-B interaction strength k. The effect of increasing the B-B interactions strength k is qualitatively similar to what we observe when decreasing the temperature. In the next section (4.4.1), we will discuss in more detail the dependence of the transition from bimodal to unimodal overlap distribution on the parameters T and k .

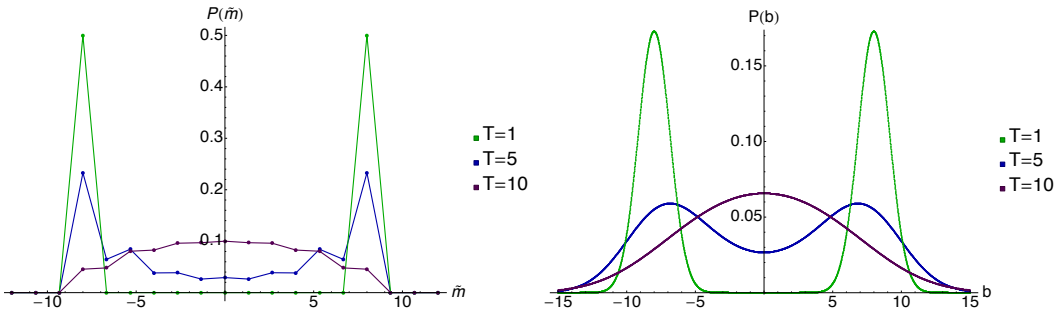

Figure 4.7: Plot of $P(\tilde{m}|K)$ (left) and the associated B clone sizes distribution $P(b)$ (right) for different temperatures $T = 1, 5, 10$, and fixed $k = 0.5$ for a regular graph with degree $K = 4$ in the paramagnetic phase.

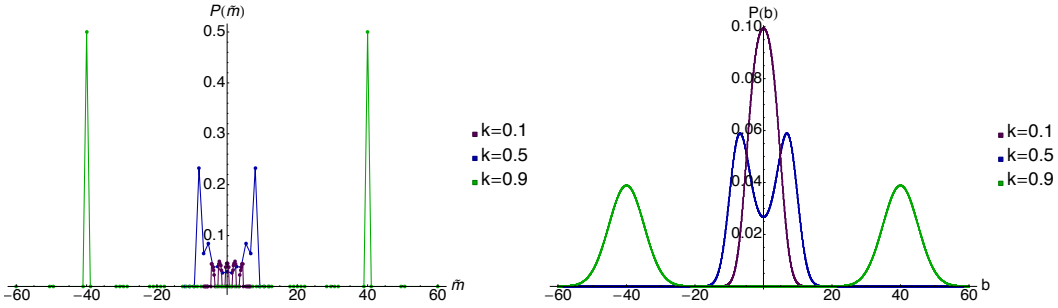

Figure 4.8: Plot of $P(\tilde{m}|K)$ (left) and the associated B clone sizes distribution $P(b)$ (right) for different strengths of idiotypic interactions $k = 0.1, 0.5, 0.9$, and temperature $T = 5$ for a regular graph with degree $K = 4$ in the paramagnetic phase.

Finally, we plot the distribution of rotated overlaps that is needed to derive the B clone size distribution. In fig. 4.7 (left panel) we show the behaviour of $P(\tilde{m}|K)$ when varying the temperature, while in fig. 4.8 (left panel) we plot it for different values of k. The behaviour is qualitatively similar to what we discussed for $P(m|K)$ and directly affects the B clones size distribution. The latter can be computed from $P(\tilde{m}|K)$ by using (4.1.16, 4.1.17, 4.1.18), and is shown in the right panels of the same figures. Fig. 4.7 (right panel) shows that at high

temperature the B clone size distribution is peaked at zero, meaning that no clonal expansion or contraction take place in the system. Lowering the temperature, the distribution develops two peaks leading the system to alternate, in the absence of an antigenic field, between "memorised" (expanded or contracted) states, where B clones are boosted or suppressed, respectively. An antigenic field will force the system to remain in the expanded state.

In fig. 4.8 (right panel) one notices that the peaks shift to larger clonal sizes when k is increased, showing that idiotypic interactions may help boost the proliferation of B clones. Furthermore, decreasing k at fixed T the distribution crosses over from a bimodal to a unimodal distribution peaked at zero, meaning that clonal expansions and contractions are more resilient to noise in the presence of idiotypic interactions.

4.4.1 Crossover transition

In this section, we study the dependence of the crossover transition of the overlap distribution from unimodal to bimodal (as shown in fig. 4.5 and 4.6) on the noise level T and the idiotypic interaction strength k .

We note that this crossover transition is *not* a true phase transition, due to the effective finite size K of the system. In contrast with the case $K \sim N^{1-\gamma}$, with $\gamma < 1$ (and $P \sim N^{\gamma}$) analysed in chapter 3, there is no order parameter that becomes non-zero at the crossover. Before the crossover, one has a broad distribution of the order parameter, peaked at zero, while after the crossover one has a typical timescale for the system to make a transition between one peak and the other, rather than a full ergodicity breaking. We will refer to the line in the (T, k) plane where the crossover takes place as the *single cluster activation line*, since it represents the onset of B clonal activation in the paramagnetic phase.

In order to derive the activation line, we consider the distribution of the magnetisations $m \in \{-K, \dots, K\}$ and $\hat{m} \in \{-\hat{K}, \dots, \hat{K}\}$ in the cluster of size

$K+\hat{K}$ defined in (4.4.65) for the case $K=\hat{K}$. This distribution is of the form

$$P(m,\hat{m}|K,\hat{K}) = \frac{1}{\mathcal{Z}} \mathrm{e}^{\frac{\beta}{2(1-k^2)}\left(m^2+2km\hat{m}+\hat{m}^2\right)} \binom{K}{\frac{K+m}{2}} \binom{\hat{K}}{\frac{\hat{K}+\hat{m}}{2}} , \tag{4.4.67}$$

where we introduced the binomial coefficients, which are meant to be zero for non-integer arguments. The normalising constant, which reads

$$\mathcal{Z} = \sum_{m,\hat{m}} \mathrm{e}^{\frac{\beta}{2(1-k^2)}\left(m^2+2km\hat{m}+\hat{m}^2\right)} \binom{K}{\frac{K+m}{2}} \binom{\hat{K}}{\frac{\hat{K}+\hat{m}}{2}} , \tag{4.4.68}$$

gives the system's partition function. Hence, we can derive the free energy $F = -\frac{1}{\beta}\log\mathcal{Z}$, which is expected to give information on the "critical behaviour" of the cluster. We note that F cannot be directly computed but we can obtain bounds on this function by considering bounds for $\mathcal{Z}$. In particular, we can exploit the inequalities

$$\frac{1}{K+1}\mathrm{e}^{K\mathcal{S}\left(\frac{r}{K}\right)} \leq \binom{K}{r} \leq \mathrm{e}^{K\mathcal{S}\left(\frac{r}{K}\right)} , \tag{4.4.69}$$

where $\mathcal{S}(p) = -p\log p-(1-p)\log(1-p)$. The second inequality, after the change of variables $m=\frac{m}{K}$ and $\hat{m}=\frac{\hat{m}}{\hat{K}}$, gives us the upper bound

$$\begin{aligned}\mathcal{Z} &\leq \sum_{m,\hat{m}} \mathrm{e}^{\frac{\beta}{2(1-k^2)}\left(K^2m^2+2K\hat{K}km\hat{m}\hat{K}^2+\hat{m}^2\right)+K\mathcal{S}\left(\frac{1+m}{2}\right)+\hat{K}\mathcal{S}\left(\frac{1+\hat{m}}{2}\right)} \\ &\leq (2K+1)(2\hat{K}+1)\mathrm{e}^{\sup_{m,\hat{m}\in[-1,1]}\phi_1(m,\hat{m})} ,\end{aligned} \tag{4.4.70}$$

where we have defined the function

$$\phi_1(m,\hat{m}) = \frac{\beta\left(K^2m^2+2kK\hat{K}m\hat{m}+\hat{K}^2\hat{m}^2\right)}{2(1-k^2)} + K\mathcal{S}\left(\frac{1+m}{2}\right) + \hat{K}\mathcal{S}\left(\frac{1+\hat{m}}{2}\right) . \tag{4.4.71}$$

The lower bound on the free energy follows from (4.4.71) and reads

$$F \geq \frac{1}{\beta}\left[-\sup_{m,\hat{m}\in[-1,1]}\phi_1(m,\hat{m}) - \log\Big((2K+1)(2\hat{K}+1)\Big)\right] . \tag{4.4.72}$$

Let us now consider the function $\phi_1(m, \hat{m})$. The stationary points of this function satisfy the "mean-field" equations

$$m = \tanh\left(\frac{\beta}{1-k^2}\left(Km + \hat{K}k\hat{m}\right)\right) , \tag{4.4.73}$$

$$\hat{m} = \tanh\left(\frac{\beta}{1-k^2}\left(\hat{K}\hat{m} + Kkm\right)\right) . \tag{4.4.74}$$

In this approximation, eq. (4.4.73, 4.4.74) qualitatively correspond to the steady state of the equations (3.3.106) derived in sec. 3.3. We note that the "paramagnetic" point $(m, \hat{m}) = (0, 0)$ is always a solution to the above equations. However, the point $(0, 0)$ becomes unstable when the largest eigenvalue of the Jacobian J of the system (4.4.73, 4.4.74) evaluated in $(0, 0)$

$$J = \frac{\beta}{1-k^2}\begin{pmatrix} K & \hat{K}k \\ Kk & \hat{K} \end{pmatrix} \tag{4.4.75}$$

becomes greater than one. This happens for

$$\beta \geq \frac{2(1-k^2)}{K + \hat{K} + \sqrt{K^2 + 2K\hat{K}\left(2k^2 - 1\right) + \hat{K}^2}} , \tag{4.4.76}$$

which for $K = \hat{K}$ reduces to $\beta \geq \frac{1-k}{K}$. In this regime, the point $(0, 0)$ ceases to be a maximum and becomes a saddle-point of the function $\phi_1(m, \hat{m})$. Interestingly, the line $T = K/(1-k)$ coincides with the critical line of the real phase transition occurring in the system with a sub-extensive number $P \sim N^\gamma$ of extremely diluted patterns, with $\gamma < 1$, analysed in chapter 3, where the typical size of a cluster is $K \sim N^{1-\gamma}$.

This connection is best understood by looking also at the first inequality in (4.4.69), which gives an upper bound on F. Let us define α such that $\hat{K} = \alpha K$

and set $\beta = \frac{\tilde{\beta}}{K}$ then

$$\begin{aligned}\mathcal{Z} &\geq \frac{1}{(K+1)(\hat{K}+1)} \sum_{m,\hat{m}} \mathrm{e}^{\frac{\beta}{2(1-k^2)}\left(K^2m^2+2kK\hat{K}m\hat{m}+\hat{K}^2\hat{m}^2\right)+K\mathcal{S}\left(\frac{1+m}{2}\right)+\hat{K}\mathcal{S}\left(\frac{1+\hat{m}}{2}\right)} \\ &= \frac{1}{(K+1)(\hat{K}+1)} \sum_{m,\hat{m}} \mathrm{e}^{K\phi_2(m,\hat{m})} \\ &= \frac{\mathrm{e}^{K \sup_{m_1,m_2\in[-1,1]} \phi_2(m_1,m_2)}}{(K+1)(\hat{K}+1)} \sum_{m,\hat{m}} \mathrm{e}^{-K\left(\sup_{m_1,m_2\in[-1,1]} \phi_2(m_1,m_2)-\phi_2(m,\hat{m})\right)} ,\end{aligned} \tag{4.4.77}$$

where we have defined the function

$$\phi_2(m,\hat{m}) = \frac{\tilde{\beta}}{2(1-k^2)}\left(m^2+2k\alpha m\hat{m}+\alpha^2\hat{m}^2\right)+\mathcal{S}\left(\frac{1+m}{2}\right)+\alpha\mathcal{S}\left(\frac{1+\hat{m}}{2}\right) . \tag{4.4.78}$$

From the above lower bound on $\mathcal{Z}$ we get the upper bound

$$\begin{aligned}F \leq \frac{1}{\beta}\Bigg[&- K \sup_{m,\hat{m}\in[-1,1]} \phi_2(m,\hat{m}) - \log\left(\sum_{m,\hat{m}} \mathrm{e}^{-K\left(\sup_{m_1,m_2\in[-1,1]} \phi_2(m_1,m_2)-\phi_2(m,\hat{m})\right)}\right) \\ &+ \log\left((K+1)(\alpha K+1)\right)\Bigg] .\end{aligned} \tag{4.4.79}$$

The stationary points of the function $\phi_2(m,\hat{m})$ satisfy the equations

$$m = \tanh\left(\frac{\tilde{\beta}}{1-k^2}\left(m+\alpha k\hat{m}\right)\right) , \tag{4.4.80}$$

$$\hat{m} = \tanh\left(\frac{\tilde{\beta}}{1-k^2}\left(\alpha\hat{m}+km\right)\right) , \tag{4.4.81}$$

which, if we reverse the transformations $\hat{K} = \alpha K$ and $\beta = \frac{\tilde{\beta}}{K}$, give us again the mean-field equations (4.4.73, 4.4.74). We note that $\phi_1(m,\hat{m}) = K\phi_2(m,\hat{m})$ when $\hat{K} = \alpha K$ and $\beta = \frac{\tilde{\beta}}{K}$, hence in the limit of large K and assuming $\alpha < \infty$, the upper bound (4.4.79) and the lower bound (4.4.72) give for the free energy density $\frac{F}{K} \to -\frac{1}{\beta}\sup_{m,\hat{m}\in[-1,1]} \phi_2(m,\hat{m})$ as $K \to \infty$. This suggests that the result

(4.4.73, 4.4.74) can be interpreted as the infinite size approximation of the finite size system (4.4.67), which gives the connection with results derived in ch. 3 [61]. This system becomes "critical", i.e. develops bi-stability, when the equality in (4.4.76) is satisfied.

Numerically, it is possible to identify a crossover temperature by looking at the height difference between the peaks in 0 and K. This method is exact in the case $K = 2$, where the support consists of three points, and will provide a lower bound for the crossover temperature for $K > 2$ (an exact numerical method would consist in locating the temperature at which the peak in 0 stops being the maximum over the support of the distribution).

We consider $P(m|K, P_q) = \sum_{\hat{K}\geq 0} P_q(\hat{K}) \sum_{\hat{m}} P(m, \hat{m}|K, P_q)$, which can be written as

$$P(m|K,P_q) = \sum_{\hat{K}\geq 0} P_q(\hat{K}) \left(\frac{\sum_{j=0}^{\hat{K}} \binom{\hat{K}}{j} \mathrm{e}^{\frac{\beta}{2(1-k^2)}[m^2+(2j-\hat{K})^2+2km(2j-\hat{K})]} \binom{K}{\frac{m+K}{2}}}{\sum_{r=0}^{K} \binom{K}{r} \sum_{s=0}^{\hat{K}} \binom{\hat{K}}{s} \mathrm{e}^{\frac{\beta}{2(1-k^2)}[(2r-K)^2+2k(2r-K)(2s-\hat{K})+(2s-\hat{K})]}} \right), \tag{4.4.82}$$

where we have fixed the size of one clone to K and marginalised over the overlap $\hat{m}$ of the other clone. Its size $\hat{K}$ is assumed, in what follows, to be drawn from a Poisson distribution π_K with average K. We plot in fig. 4.9 the peak difference $\Delta = P(K|K, \pi_K) - P(0|K, \pi_K)$ as a function of T and k and we show in fig. 4.10 (left panel) the line, in the plane (T, k), where Δ first becomes positive. For low values of K, this is expected to give a good approximation for the activation line where the crossover transition occurs, and is in good agreement with the critical line $T_{\text{clust}} = K/(1-k)$, theoretically predicted for regular graph topology and large size K (dashed line).

Alternatively, for a regular topology with degree K one can monitor the total cluster magnetisation $m_t = m + \hat{m}$ via $P(m_t|K) = \sum_{m,\hat{m}} P(m, \hat{m}|K)\delta_{m_t, m+\hat{m}}$, although we expect here the peak difference $P(2K|K) - P(0|K)$ to give a worse estimate of the transition line, given the broader support of the distribution. The

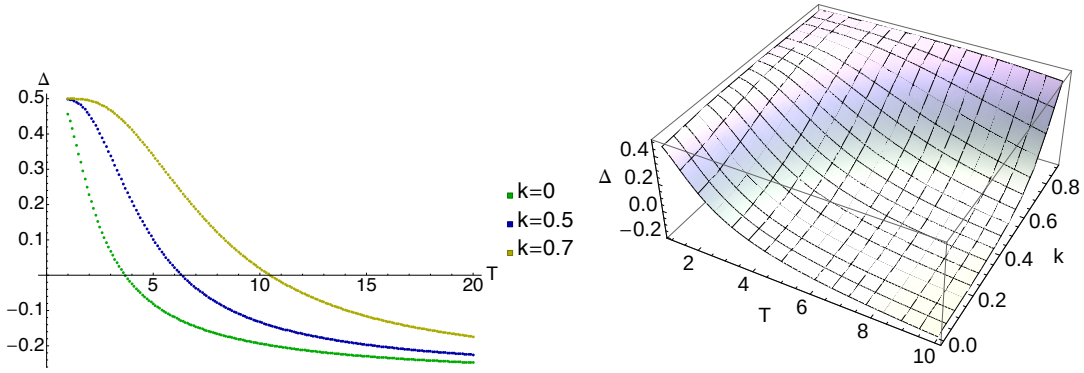

Figure 4.9: Plot of the peak difference $\Delta = P(K|K,\pi_K) - P(0|K,\pi_K)$ for the distribution defined in (4.4.82) as a function of the temperature (left) and 3D plot (right) with $K = 3$.

latter can be written as

$$P(m_t|K) = \sum_{m,\hat{m}} \frac{\mathrm{e}^{\frac{\beta}{2(1-k^2)}(m^2+2km\hat{m}+\hat{m}^2)} \binom{K}{\frac{m+K}{2}} \binom{K}{\frac{\hat{m}+K}{2}}}{\sum_{\ell=0}^{K}\binom{K}{\ell}\sum_{j=0}^{K}\binom{K}{j}\mathrm{e}^{\frac{\beta}{2(1-k^2)}((2\ell-K)^2+2k(2\ell-K)(2j-K)+(2j-K)^2}} \delta_{m_{tot},m+\hat{m}} \tag{4.4.83}$$

and the line where the peak difference first becomes positive is plotted in fig. 4.10 (right panel). This indeed provides a lower bound on the activation temperature $T_{\text{clust}} = K/(1-k)$ shown in the same plot, for guidance, as a dashed line.

4.5 Ferromagnetic interactions

Cross-talks effects between clusters can be studied in the case of ferromagnetic interactions $\xi_i^\mu = \xi_i^{\hat{\mu}} = 1$, $\forall\ i,\mu,\hat{\mu}$, where we can progress analytically even away from the paramagnetic phase. To this end, we consider spins interacting on a random regular factor-graph with $|\partial i| = L$ and $|\partial\mu| = |\partial\hat{\mu}| = K$. In the thermodynamic limit $N \to \infty$, the graph is tree-like and due to the ferromagnetic nature of the interactions all factors (4.2.41) are equivalent. It follows that all the cavity distributions are equivalent $P_{\setminus\mu\hat{\mu}}(\sigma) = P^c(\sigma)$, $\forall\ \mu,\hat{\mu}$. Parametrizing them as $P^c(\sigma_i) \propto \mathrm{e}^{\beta\phi\sigma_i}$ and denoting $\mathcal{D}\mathbf{y} = \sqrt{1-k^2}\frac{\mathrm{d}y_1\mathrm{d}y_2}{2\pi}\mathrm{e}^{\frac{-1}{2}\mathbf{y}^T\mathbf{C}^{-1}\mathbf{y}}$, we can write the

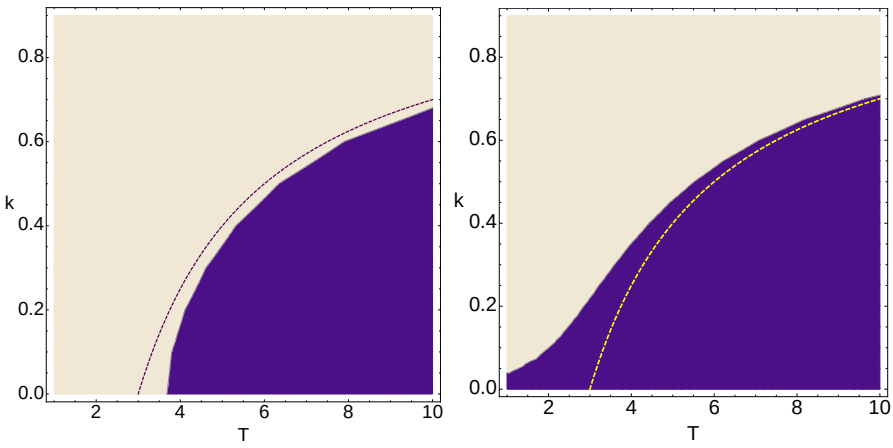

Figure 4.10: Transition line in the plane (T, k) from unimodal to bimodal distribution. Left: Zero-contour plot of the peak difference Δ in the plane (T, k) monitoring $P(m|K, P_q)$ for a regular graph with degree $K = 3$ when coupled to a factor with Poisson distribution (average degree $K = 3$). Right: Zero-contour plot of the peak difference monitoring $P(m_{tot}|K)$ with $K = 3$. The dashed lines represent the activation line $T_{\text{clust}} = \frac{K}{1-k}$ predicted for large K, consistently with results in [61].

recursive equation (4.3.53) as

$$P^c(\sigma) = \frac{1}{\mathcal{Z}} \left[\int_{-\infty}^{+\infty} \mathcal{D}\mathbf{y} \sum_{\{\sigma_k\},\{\tau_k\}} \mathrm{e}^{\sqrt{\beta}\sigma y_1 + y_1\left(\sqrt{\beta}\sum_{k=1}^{K-1}\sigma_k\right) + y_2\left(\sqrt{\beta}\sum_{k=1}^{K}\tau_k\right) + \beta\phi\left(\sum_{k=1}^{K-1}\sigma_k + \sum_{k=1}^{K}\tau_k\right)} \right]^{L-1} , \tag{4.5.84}$$

where we denoted by σ the spins attached to factor μ and by τ those attached to $\hat{\mu}$. We also assumed that, due to sparsity of interactions, σ (where the tree is rooted) is attached to factor μ only.

Considering $\phi = 0$ in (4.5.84) we retrieve the paramagnetic phase analysed in sec. 4.4 with $P^c(\sigma_i) = 1/2 \ \forall \ \sigma_i = \pm 1$, while for $\phi \neq 0$ each cluster will receive a signal from the others, acting as a field. We can manipulate (4.5.84) by summing

over $\{\sigma_k\}, \{\tau_k\}$ to obtain

$$P^c(\sigma) = \frac{\left[\int_{-\infty}^{+\infty} \frac{\mathrm{d}\mathbf{y}}{2\pi} \mathrm{e}^{\frac{-\mathbf{y}^T \mathbf{C}^{-1} \mathbf{y}}{2}} \mathrm{e}^{\sqrt{\beta}\sigma y_1} \left[\cosh(\sqrt{\beta} y_1 + \beta\phi)\right]^{K-1} \left[\cosh(\sqrt{\beta} y_2 + \beta\phi)\right]^K\right]^{L-1}}{\sum_{\tilde{\sigma}} \left[\int_{-\infty}^{+\infty} \frac{\mathrm{d}\mathbf{y}}{2\pi} \mathrm{e}^{\frac{-\mathbf{y}^T \mathbf{C}^{-1} \mathbf{y}}{2}} \mathrm{e}^{\sqrt{\beta}\tilde{\sigma} y_1} \left[\cosh(\sqrt{\beta} y_1 + \beta\phi)\right]^{K-1} \left[\cosh(\sqrt{\beta} y_2 + \beta\phi)\right]^K\right]^{L-1}} \tag{4.5.85}$$

and using $P^c(1) \propto \mathrm{e}^{\beta\phi}$ and $P^c(-1) \propto \mathrm{e}^{-\beta\phi}$, we get

$$\phi = \frac{1}{2\beta} \log \frac{P^c(1)}{P^c(-1)} \ . \tag{4.5.86}$$

This leads to the following self-consistency equation for ϕ

$$\phi = \frac{L-1}{2\beta} \log \left(\frac{\int_{-\infty}^{+\infty} \frac{\mathrm{d}\mathbf{y}}{2\pi} \mathrm{e}^{\frac{-\mathbf{y}^T \mathbf{C}^{-1} \mathbf{y}}{2}} \mathrm{e}^{\sqrt{\beta} y_1} \left[\cosh(\sqrt{\beta} y_1 + \beta\phi)\right]^{K-1} \left[\cosh(\sqrt{\beta} y_2 + \beta\phi)\right]^K}{\int_{-\infty}^{+\infty} \frac{\mathrm{d}\mathbf{y}}{2\pi} \mathrm{e}^{\frac{-\mathbf{y}^T \mathbf{C}^{-1} \mathbf{y}}{2}} \mathrm{e}^{-\sqrt{\beta} y_1} \left[\cosh(\sqrt{\beta} y_1 + \beta\phi)\right]^{K-1} \left[\cosh(\sqrt{\beta} y_2 + \beta\phi)\right]^K} \right) \ . \tag{4.5.87}$$

Clearly $\phi = 0$ is always a solution, but we expect it to become unstable at low temperature. With simple manipulations we can rewrite (4.5.87) in the following form

$$\phi = \frac{L-1}{2} \log \left(\frac{\sum_{\ell=0}^{K-1} \binom{K-1}{\ell} \sum_{p=0}^{K} \binom{K}{p} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\mathbf{y}}{2\pi} \mathrm{e}^{\frac{-\mathbf{y}^T \mathbf{C}^{-1} \mathbf{y}}{2}} \mathrm{e}^{\mathbf{J}^T \mathbf{y} + \beta\phi(2\ell + 2p - 2K + 1)}}{\sum_{\tilde{\ell}=0}^{K-1} \binom{K-1}{\tilde{\ell}} \sum_{\tilde{p}=0}^{K} \binom{K}{\tilde{p}} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\tilde{\mathbf{y}}}{2\pi} \mathrm{e}^{\frac{-\tilde{\mathbf{y}}^T \mathbf{C}^{-1} \tilde{\mathbf{y}}}{2}} \mathrm{e}^{\mathbf{M}^T \tilde{\mathbf{y}} + \beta\phi(2\tilde{\ell} + 2\tilde{p} - 2K + 1)}} \right) \ , \tag{4.5.88}$$

where $\mathbf{J}^T = (\sqrt{\beta}(2 + 2\ell - K), \sqrt{\beta}(2p - K))$ and $\mathbf{M}^T = (\sqrt{\beta}(2\tilde{\ell} - K), \sqrt{\beta}(2\tilde{p} - K))$. Integrating over $\mathbf{y}$, we obtain

$$\phi = \frac{L-1}{2} \log \left(\frac{\sum_{\ell=0}^{K-1} \binom{K-1}{\ell} \sum_{p=0}^{K} \binom{K}{p} \mathrm{e}^{\frac{\beta((2\ell - K + 2)^2 + (2p - K)^2 + 2k(2\ell - K + 2)(2p - K))}{1 - k^2} + \beta\phi(2\ell + 2p - 2K + 1)}}{\sum_{\tilde{\ell}=0}^{K-1} \binom{K-1}{\tilde{\ell}} \sum_{\tilde{p}=0}^{K} \binom{K}{\tilde{p}} \mathrm{e}^{\frac{\beta((2\tilde{\ell} - K)^2 + (2\tilde{p} - K)^2 + 2k(2\tilde{\ell} - K)(2\tilde{p} - K))}{1 - k^2} + \beta\phi(2\tilde{\ell} + 2\tilde{p} - 2K + 1)}} \right) \ . \tag{4.5.89}$$

Via bifurcation analysis we can determine the critical temperature T_c at which

cavity fields bifurcate to a non-zero value (we will provide full details for general types of interactions in sec. 4.6). In fig. 4.11 (left), we plot the critical temperature T_c as a function of the vertex and factor degree for two different values of the B-B interaction strength k. Increasing k widens the region where the cavity fields are non-zero, the so-called *interference region*. Given that in the ferromagnetic case the cavity fields are homogeneous, the interference between cytokine patterns is constructive and does not disrupt the system's parallel retrieval of information.

In addition, in the right panel of fig. 4.11 we show a comparison between the critical temperature T_c, at which the cavity field bifurcate to non-zero values, and the single cluster activation temperature T_{clust} (computed in sec. 4.4.1) for fixed values of $K = 4$ and $L = 2$.

As in the paramagnetic case, we compute the overlap distribution, defined in (4.3.55), for a regular factor graph with degree K (generalizations to non-regular topologies are straightforward) from

$$P(m,\hat{m}|K) = \frac{\mathrm{e}^{\frac{\beta}{2(1-k^2)}(m^2+\hat{m}^2+2km\hat{m})+\beta\phi(m+\hat{m})}\sum_{\ell=0}^{K}\binom{K}{\ell}\sum_{p=0}^{K}\binom{K}{p}\delta_{m,2\ell-K}\delta_{\hat{m},2p-K}}{\sum_{\tilde{\ell}=0}^{K}\binom{K}{\tilde{\ell}}\sum_{\tilde{p}=0}^{K}\binom{K}{\tilde{p}}\mathrm{e}^{\frac{\beta((2\tilde{\ell}-K)^2+(2\tilde{p}-K)^2+2k(2\tilde{\ell}-K)(2\tilde{p}-K)^2)}{2(1-k^2)}+\beta\phi(2\tilde{\ell}+2\tilde{p}-2K+1)}}\,, \tag{4.5.90}$$

where ϕ is self-consistently obtained from eq. (4.5.89). In fig. 4.12 we plot the distribution (4.5.90) for different values of the temperature and fixed degrees $K = 4$, $L = 2$. For $T > T_c$ where $\phi = 0$ (see fig. 4.11, right), the overlap distribution (4.5.90) reduces to (4.4.65), which describes the behaviour of a single cluster in the paramagnetic phase and does not depend on the nature (e.g. ferromagnetic or disordered) of the interactions. Since T_c in this case is larger than the single cluster activation temperature (see fig. 4.11, right), the system is found in the high-temperature regime of the paramagnetic phase and displays an overlap distribution peaked at zero.

Lowering the temperature, ϕ becomes non-zero: hence, we expect that the predictions obtained in the paramagnetic phase become inaccurate in this regime. Numerical evaluation of (4.5.90) shows that at low temperature one has again a crossover of the overlap distribution to an activated regime, where it displays a single peak at either positive or negative values of the magnetisation, depending on the initial conditions (see fig. 4.12, left). Indeed, clonal cross-talk arises, in this regime, as a field that pins the system at one peak or the other, and the system develops a non-zero global magnetisation (as for ferromagnetic interactions the field is homogeneous).

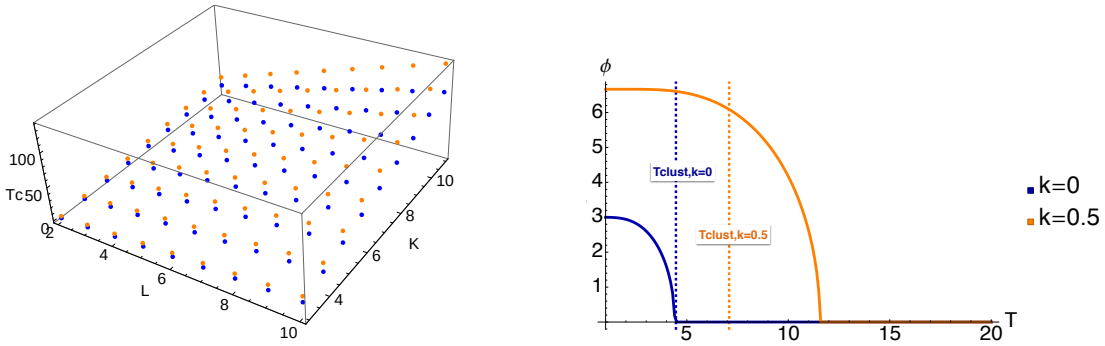

Figure 4.11: Left: Critical temperature T_c for the transition $\phi = 0 \to \phi \neq 0$ for a regular graph with degrees K, L for the values of the B-B interaction strength $k = 0.5$ (orange) and $k = 0$ (blue). Right: Plot of ϕ as a function of the temperature T for $K = 4$ and $L = 2$. Dashed lines represent the single cluster activation temperature (sec. 4.4).

Marginalising over $\hat{m}$, we can get information about the overlap distribution of the μ-th B clone given that it interacts with its complement $\hat{\mu}$: in fig. 4.13 we plot it for different values of the temperature and the B-B interaction strength, and observe qualitatively the same transition described above. From the joint distribution $P(m, \hat{m}|K)$, we can compute the distribution $P(\tilde{m}|K)$ of rotated overlaps $\tilde{m}_\mu = \frac{m_\mu}{1-k^2} + \frac{k m_{\hat{\mu}}}{1-k^2}$, which is plotted in fig. 4.14 (left panel) for different temperatures and in fig. 4.15 (left panel) for different values of k.

We can finally use the rotated overlap distribution $P(\tilde{m})$ to derive the B clones size distribution, as shown in sec. 4.1. In fig. 4.14 and 4.15 (right panels) we

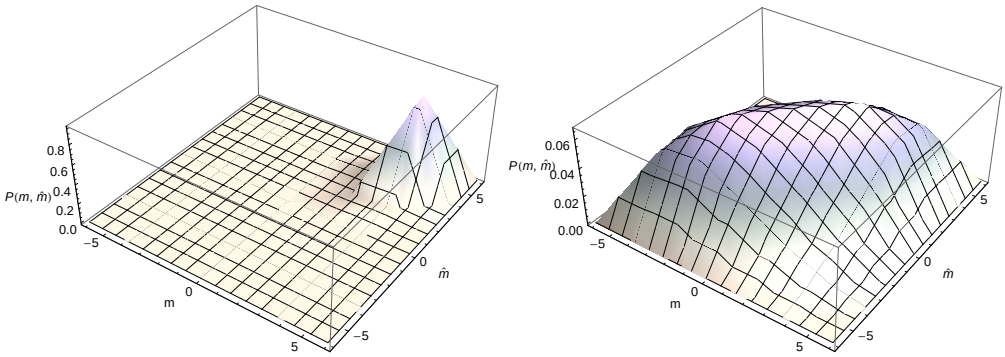

Figure 4.12: Plot of $P(m, \hat{m}|K)$ defined in (4.5.90) for $T = 5$ (left panel) and $T = 15$ (right panel) fixing $k = 0.5$, for a regular graph with $K = 4, L = 2$ with ferromagnetic interactions.

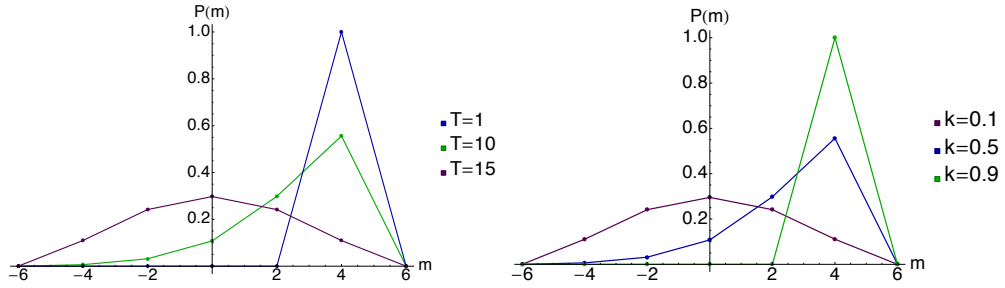

Figure 4.13: Plot of $P(m|K)$ for a regular graph with $K = 4, L = 2$ with ferromagnetic interactions. Left: $P(m|K)$ varying the temperature $T = 1, 10, 15$ with $k = 0.5$. Right: $P(m)$ varying B-B interaction strength $k = 0.1, 0.5, 0.9$ for $T = 10$.

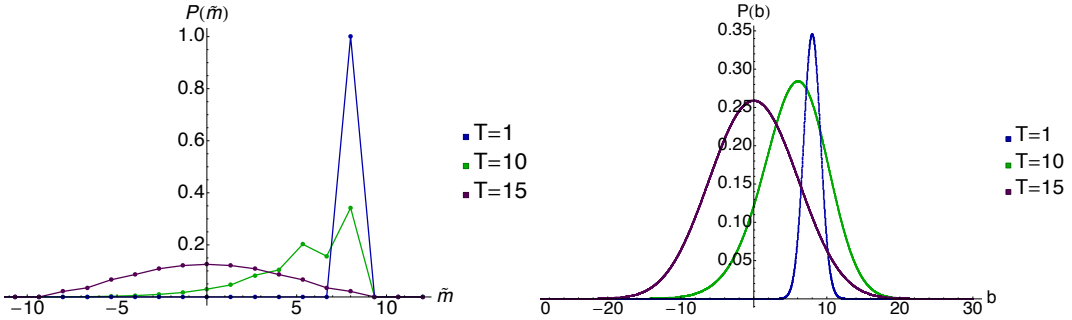

Figure 4.14: Plot of $P(\tilde{m}|K)$ (left) and the corresponding B clones distribution $P(b)$ (right) for different temperatures $T = 1, 5, 10$ for a regular graph with $K = 4, L = 2$ and $k = 0.5$ with ferromagnetic interactions.

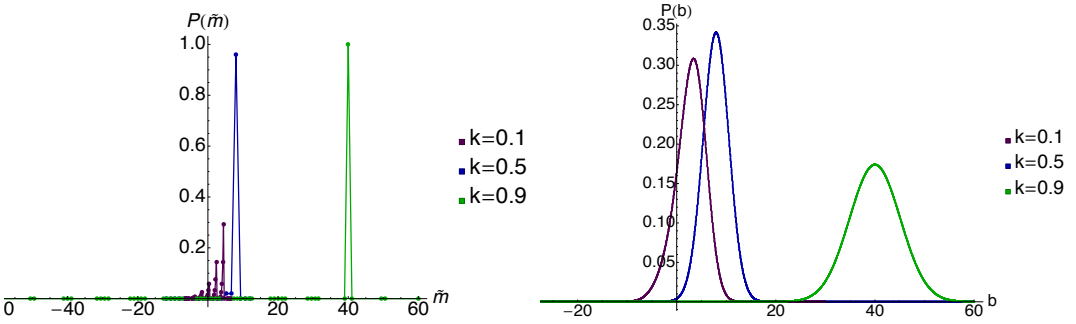

Figure 4.15: Plot of $P(\tilde{m}|K)$ (left) and the corresponding B clones distribution $P(b)$ (right) for different B-B interaction strength $k = 0.1, 0.5, 0.9$ for a regular graph with $K = 4, L = 2$ and $T = 5$ with ferromagnetic interactions.

show its behaviour in temperature and in k respectively. At low temperature, the B clone size distribution is peaked around non-zero values, meaning that B clones are expanding. Increasing k, the probability of having strong clonal expansions (even at high noise levels) increases.

B clone sizes, as anticipated, are often experimentally measured as concentrations and in some cases are thought to follow a Zipf's law [12]. Using our definition of clonal sizes as (relative) log-concentrations $b = \log c/c_0$ we can get concentration distributions as $\mathcal{P}(c) = \int P(b)\delta(c - \mathrm{e}^b)\mathrm{d}b = \frac{1}{c}P(\log(c))$. In fig. 4.16 we show the log-log plot of $\mathcal{P}(c)$ for different values of T (left) and k (right). In particular, we see that increasing k and decreasing T, increases the probability of having clones in high concentrations. More generally, our model enables us to determine the single most important parameters that affect the tail behaviour of these distributions and might be useful to infer the network connectivity and operational noise in health and disease situations.

4.6 Disordered interactions

In order to discuss more general cases of interactions, it is useful to parametrise $P_{\mu\hat{\mu}}(\sigma_j)$, the *message* from factor μ and $\hat{\mu}$ to node i, by an effective field $\psi_{\mu\hat{\mu}\to j}$ and $P_{\backslash\nu\hat{\nu}}(\sigma_j)$, the *message* from node j to factors $\mu, \hat{\mu}$, by the effective field $\phi_{j\to\nu\hat{\nu}}$

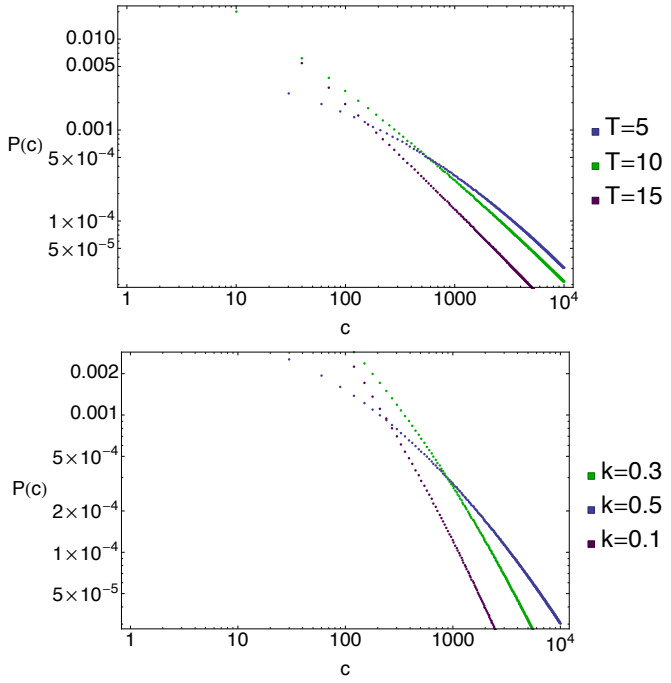

Figure 4.16: Log-Log plot of $\mathcal{P}(c)$ for a regular graph with $K = 4, L = 2$ with ferromagnetic interactions. Top: $\mathcal{P}(c)$ varying the temperature fixing $k = 0.5$. Bottom: $\mathcal{P}(c)$ varying k for $T = 5$.

as

$$P_{\mu\hat{\mu}}(\sigma_j) \propto \mathrm{e}^{\beta\sigma_j\psi_{\mu\hat{\mu}\to j}} , \tag{4.6.91}$$

$$P_{\backslash\nu\hat{\nu}}(\sigma_j) \propto \mathrm{e}^{\beta\sigma_j\phi_{j\to\nu\hat{\nu}}} . \tag{4.6.92}$$

A schematic representation of these messages in a factor graph can be seen in fig. 4.17. Using the relation $\langle\sigma_j\rangle = \frac{\sum_{\sigma_j} P_{\mu\hat{\mu}}(\sigma_j)\sigma_j}{\sum_{\sigma_j} P_{\mu\hat{\mu}}(\sigma_j)} = \tanh(\beta\psi_{\mu\hat{\mu}\to j})$ we can derive an expression for $\psi_{\mu\hat{\mu}\to j}$. From (4.3.53) and the definition of $f_{\mu\hat{\mu}}(\{\sigma_{k\in\partial\mu}\},\{\sigma_{\ell\in\partial\hat{\mu}}\})$ in (4.2.41), we have

$$P_{\mu\hat{\mu}}(\sigma_j) = \frac{1}{\mathcal{Z}_{\mu\hat{\mu}}} \sum_{\{\sigma_{k\in\partial\mu}\},\{\sigma_{\ell\in\partial\hat{\mu}}\}} \int \mathcal{D}\mathbf{y} \mathrm{e}^{\sqrt{\beta}(y_1\sum_{k\in\partial\mu}\xi_k^{\mu}\sigma_k + y_2\sum_{\ell\in\partial\hat{\mu}}\xi_\ell^{\hat{\mu}}\sigma_\ell)} \left[\prod_{k\in\partial\mu\hat{\mu}\backslash j} P_{\backslash\mu\hat{\mu}}(\sigma_k)\right] , \tag{4.6.93}$$

where $\mathcal{Z}_{\mu\hat{\mu}}$ is the normalization. Manipulating (4.6.93) by using the parametrisation (4.6.92), yields

$$P_{\mu\hat{\mu}}(\sigma_j) = \frac{1}{\mathcal{Z}_{\mu\hat{\mu}}} \sum_{\{\sigma_{k\in\partial\mu}\},\{\sigma_{\ell\in\partial\hat{\mu}}\}} \int \mathcal{D}\mathbf{y} \mathrm{e}^{\sqrt{\beta}y_1\xi_j^\mu\sigma_j} \left[\prod_{k\in\partial\mu\setminus j} \mathrm{e}^{(\sqrt{\beta}y_1\xi_k^\mu+\beta\phi_{k\to\mu\hat{\mu}})\sigma_k}\right] \times \\ \times \left[\prod_{\ell\in\partial\hat{\mu}} \mathrm{e}^{(\sqrt{\beta}y_2\xi_\ell^{\hat{\mu}}+\beta\phi_{\ell\to\mu\hat{\mu}})\sigma_\ell}\right] . \tag{4.6.94}$$

Note that due to the sparsity of the links, we assume that each spin is connected to either μ or its complementary $\hat{\mu}$ factor. Summing over $\{\sigma_k\}, \{\sigma_\ell\}$, we obtain

$$P_{\mu\hat{\mu}}(\sigma_j) = \frac{1}{\mathcal{Z}_{\mu\hat{\mu}}} \int \mathcal{D}\mathbf{y} \mathrm{e}^{\sqrt{\beta}y_1\xi_j^\mu\sigma_j} \left[\prod_{k\in\partial\mu\setminus j} 2\cosh\left(\sqrt{\beta}y_1\xi_k^\mu + \beta\phi_{k\to\mu\hat{\mu}}\right)\right] \times \\ \times \left[\prod_{\ell\in\partial\hat{\mu}} 2\cosh\left(\sqrt{\beta}y_2\xi_\ell^{\hat{\mu}} + \beta\phi_{\ell\to\mu\hat{\mu}}\right)\right] . \tag{4.6.95}$$

Hence, the cavity fields in (4.6.91) are $\psi_{\mu\hat{\mu}\to j} =$

$$\operatorname{atanh}\left(\frac{\langle \sinh\sqrt{\beta}(y_1\xi_j^\mu)\prod_{k\in\partial\mu\setminus j}\cosh\left(\sqrt{\beta}y_1\xi_k^\mu+\beta\phi_{k\to\mu\hat{\mu}}\right)\prod_{\ell\in\partial\hat{\mu}}\cosh\left(\sqrt{\beta}y_2\xi_\ell^{\hat{\mu}}+\beta\phi_{\ell\to\mu\hat{\mu}}\right))\rangle_{\mathbf{y}}}{\langle \cosh\sqrt{\beta}(y_1\xi_j^\mu)\prod_{k\in\partial\mu\setminus j}\cosh\left(\sqrt{\beta}y_1\xi_k^\mu+\beta\phi_{k\to\mu\hat{\mu}}\right)\prod_{\ell\in\partial\hat{\mu}}\cosh\left(\sqrt{\beta}y_2\xi_\ell^{\hat{\mu}}+\beta\phi_{\ell\to\mu\hat{\mu}}\right))\rangle_{\mathbf{y}}}\right) . \tag{4.6.96}$$

Finally, using the recursive equation (4.3.54) and the parametrisation (4.6.91, 4.6.92) it follows also that

$$\phi_{j\to\nu\hat{\nu}} = \sum_{\mu\hat{\mu}} \psi_{\mu\hat{\mu}\to j} . \tag{4.6.97}$$

The cavity fields equations (4.6.96, 4.6.97) can be iterated until convergence [71]. We can show analytically and check numerically that $\psi_{\mu\hat{\mu}\to j} = 0, \forall\mu,\hat{\mu},j$ is a fixed point of (4.6.96) for any value of k, simply using symmetries of the integral. In the following we will study the transition from zero to non-zero cavity fields and we will discuss how it affects the retrieval properties and the functioning of the system.

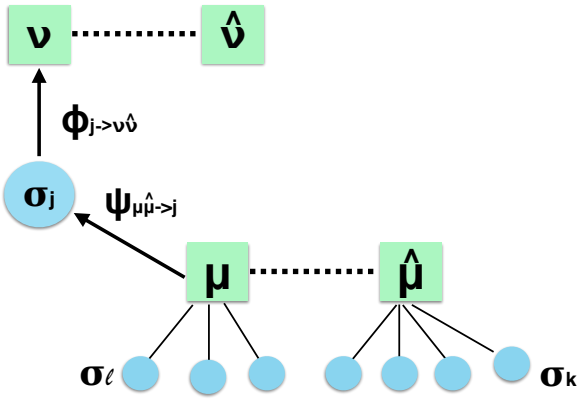

Figure 4.17: Schematic representation of the messages $\phi_{j\to\nu\hat{\nu}}$ from node j to factors $\nu\hat{\nu}$, and $\psi_{\mu\hat{\mu}\to j}$, from the factors $\mu\hat{\mu}$ to node j in the factor tree used to derive equations (4.6.96) and (4.6.97).

4.6.1 Distributions of cavity fields

In the large N limit, the solution of the cavity equations (4.6.97, 4.6.96) can be characterised via the distribution of messages or fields, $W_\psi(\psi)$ and $W_\phi(\phi)$. The field distributions can be computed as follows, denoting by $\Psi(\{\xi_i^\mu\},\{\xi_i^{\hat{\mu}}\},\{\phi_{k\to\mu\hat{\mu}}\},\{\phi_{\ell\to\mu\hat{\mu}}\})$ the rhs of (4.6.96),

$$W_\psi(\psi)=\sum_e \frac{Q(e)e}{\langle e\rangle}\langle\langle\delta(\psi-\Psi(\phi_1,...,\phi_{e-1},\{\xi^1,\dots,\xi^e\}))\rangle\rangle_{\boldsymbol{\xi},\boldsymbol{\phi}}\ , \tag{4.6.98}$$

and taking the average over i.i.d. values of the (non-zero) $\{\xi^e\}$ and over i.i.d. fields $\phi_1,...,\phi_{e-1}$ drawn from $W_\phi(\phi)$. In the above expression, $Q(e)e/\langle e\rangle$ is the probability of picking an edge connected to a cluster of degree e [73]. Since the cluster is composed by the union of the nodes signalling to factor μ and $\hat{\mu}$, its degree distribution $Q(e)$ follows from the degree distribution of the disjoint factors $\mu,\hat{\mu}$ as

$$Q(e)=\sum_{q,\hat{q}} P_q(q)P_q(\hat{q})\delta(e-q-\hat{q})\ , \tag{4.6.99}$$

and

$$\langle e\rangle = \sum_e Q(e)e = \sum_e \sum_{q,\hat{q}} eP_q(q)P_q(\hat{q})\delta(e-q-\hat{q}) = \sum_{q,\hat{q}} P_q(q)P_q(\hat{q})(q+\hat{q}) = 2\langle q\rangle \ . \tag{4.6.100}$$

Similarly we have

$$W_\phi(\phi) = \sum_d \frac{dP_d(d)}{\langle d\rangle} \left\langle\left\langle \delta\left(\phi - \sum_{\mu=1}^{d-1}\psi_\mu\right)\right\rangle\right\rangle_{\boldsymbol{\xi},\boldsymbol{\psi}} \ , \tag{4.6.101}$$

where the average is over i.i.d. values of the (non-zero) $\{\xi^e\}$ and i.i.d. fields $\psi_1, ..., \psi_{d-1}$ drawn from $W_\psi(\psi)$. $P_d(d)$ represents the probability of picking a node of degree d. Field distributions can then be obtained numerically by a population dynamics (PD) algorithm [74], details of which are provided in chapter A.2.

4.6.2 Small fields expansion and bifurcation lines

Transitions from zero to non-zero cavity fields can be located by monitoring bifurcations away from zero of the moments of the field distribution. Depending on the model's parameters, either the first or the second moment will first bifurcate away from zero [39]. To this end, we Taylor expand for small fields the rhs

of equation (4.6.96) to first order around $\boldsymbol{\phi} = \mathbf{0}$, which gives

$$\Psi(\xi_j^\mu, \{\xi_k^\mu\}, \{\xi_\ell^{\hat\mu}\}, \{\phi_{k\to\mu\hat\mu}\}, \{\phi_{\ell\to\mu\hat\mu}\}) \simeq$$

$$\simeq \sum_{k\in\partial\mu\setminus j} \phi_{k\to\mu\hat\mu} \frac{\left\langle \sinh(\sqrt{\beta}y_1\xi_j^\mu)\sinh(\sqrt{\beta}y_1\xi_k^\mu)\prod_{r\in\partial\mu\setminus\{k,j\}}\cosh(\sqrt{\beta}y_1\xi_r^\mu)\prod_{\ell\in\partial\hat\mu}\cosh(\sqrt{\beta}y_2\xi_\ell^{\hat\mu})\right\rangle_{\mathbf{y}}}{\left\langle \prod_{k\in\partial\mu}\cosh(\sqrt{\beta}y_1\xi_k^\mu)\prod_{\ell\in\partial\hat\mu}\cosh(\sqrt{\beta}y_2\xi_\ell^{\hat\mu})\right\rangle_{\mathbf{y}}} +$$

$$+ \sum_{\ell\in\partial\hat\mu} \phi_{\ell\to\mu\hat\mu} \frac{\left\langle \sinh(\sqrt{\beta}y_1\xi_j^\mu)\sinh(\sqrt{\beta}y_2\xi_\ell^{\hat\mu})\prod_{r\in\partial\mu\setminus j}\cosh(\sqrt{\beta}y_1\xi_r^\mu)\prod_{p\in\partial\hat\mu\setminus\ell}\cosh(\sqrt{\beta}y_2\xi_p^{\hat\mu})\right\rangle_{\mathbf{y}}}{\left\langle \prod_{k\in\partial\mu}\cosh(\sqrt{\beta}y_1\xi_k^\mu)\prod_{\ell\in\partial\hat\mu}\cosh(\sqrt{\beta}y_2\xi_\ell^{\hat\mu})\right\rangle_{\mathbf{y}}} =$$

$$= \sum_{k\in\partial\mu\setminus j} \phi_{k\to\mu\hat\mu}\Omega^a(\xi_j^\mu, \xi_k^\mu, \{\xi_r^\mu\}, \{\xi_\ell^{\hat\mu}\}) + \sum_{\ell\in\partial\hat\mu} \phi_{\ell\to\mu\hat\mu}\Omega^b(\xi_j^\mu, \xi_\ell^{\hat\mu}, \{\xi_r^\mu\}, \{\xi_p^{\hat\mu}\}) \ , \tag{4.6.102}$$

where

$$\Omega^a(\xi^1,\dots,\xi^e) = \frac{\langle \sinh(\sqrt{\beta}y_1\xi^1)\sinh(\sqrt{\beta}y_1\xi^2)\prod_{r=3}^{q-1}\cosh(\sqrt{\beta}y_1\xi^r)\prod_{\ell=q}^{e}\cosh(\sqrt{\beta}y_2\xi^\ell)\rangle_{\mathbf{y}}}{\langle\prod_{r=1}^{q-1}\cosh(\sqrt{\beta}y_1\xi^r)\prod_{\ell=q}^{e}\cosh(\sqrt{\beta}y_2\xi^\ell)\rangle_{\mathbf{y}}} \ , \tag{4.6.103}$$

$$\Omega^b(\xi^1,\dots,\xi^e) = \frac{\langle \sinh(\sqrt{\beta}y_1\xi^1)\sinh(\sqrt{\beta}y_2\xi^2)\prod_{r=3}^{q}\cosh(\sqrt{\beta}y_1\xi^r)\prod_{\ell=q+1}^{e}\cosh(\sqrt{\beta}y_2\xi^\ell)\rangle_{\mathbf{y}}}{\langle\prod_{r=1}^{q-1}\cosh(\sqrt{\beta}y_1\xi^r)\prod_{\ell=q}^{e}\cosh(\sqrt{\beta}y_2\xi^\ell)\rangle_{\mathbf{y}}} \ . \tag{4.6.104}$$

Moments can be obtained using the fields distribution (4.6.98, 4.6.101) and averaging the small field expansion (4.6.102) over ϕ, ψ, ξ . For the mean bifurcation we get

$$\langle\psi\rangle_\psi = \langle\phi\rangle_\phi \sum_{e,q} \frac{P_q(q)P_q(e-q)e}{2\langle q\rangle}\left[(q-1)\langle\Omega^a(\xi^1,\dots,\xi^e)\rangle_{\boldsymbol{\xi}} + (e-q)\langle\Omega^b(\xi^1,\dots,\xi^e)\rangle_{\boldsymbol{\xi}}\right] \ , \tag{4.6.105}$$

$$\langle\phi\rangle_\phi = \langle\psi\rangle_\psi \sum_{d} \frac{P_d(d)d(d-1)}{\langle d\rangle} \ . \tag{4.6.106}$$

Combining them we obtain

$$\langle\psi\rangle_\psi = \langle\psi\rangle_\psi \sum_{d,e,q} \frac{P_d(d)d(d-1)}{\langle d\rangle} \frac{P_q(q)P_q(e-q)e}{2\langle q\rangle} \times \\ \times \left[(q-1)\langle\Omega^a(\xi^1,\dots,\xi^e)\rangle_{\boldsymbol{\xi}} + (e-q)\langle\Omega^b(\xi^1,\dots,\xi^e)\rangle_{\boldsymbol{\xi}}\right] , \tag{4.6.107}$$

with solutions $\langle\psi\rangle_\psi = 0$ or

$$1 = \sum_{d} \frac{P_d(d)d(d-1)}{\langle d\rangle} \sum_{e,q} \frac{P_q(q)P_q(e-q)e}{2\langle q\rangle} \times \\ \times \left[(q-1)\langle\Omega^a(\xi^1,\dots,\xi^e)\rangle_{\boldsymbol{\xi}} + (e-q)\langle\Omega^b(\xi^1,\dots,\xi^e)\rangle_{\boldsymbol{\xi}}\right] . \tag{4.6.108}$$

If the means are zero, the transition can be detected monitoring the variances. The equation for the variances read

$$\langle\psi^2\rangle_\psi = \langle\phi^2\rangle_\phi \sum_{e,q} \frac{P_q(q)P_q(e-q)e}{2\langle q\rangle} \left[(q-1)\langle(\Omega^a(\xi^1,\dots,\xi^e)^2\rangle_{\boldsymbol{\xi}} + (e-q)\langle(\Omega^b(\xi^1,\dots,\xi^e))^2\rangle_{\boldsymbol{\xi}}\right] , \tag{4.6.109}$$

$$\langle\phi^2\rangle_\phi = \langle\psi^2\rangle_\psi \sum_{d} \frac{P_d(d)d(d-1)}{\langle d\rangle} . \tag{4.6.110}$$

Combining equations (4.6.109) and (4.6.110) we have

$$\langle\psi^2\rangle_\psi = \langle\psi^2\rangle_\psi \sum_{d,e,q} \frac{P_d(d)d(d-1)}{\langle d\rangle} \frac{P_q(q)P_q(e-q)e}{2\langle q\rangle} \times \\ \times \left[(q-1)\langle(\Omega^a(\xi^1,\dots,\xi^e))^2\rangle_{\boldsymbol{\xi}} + (e-q)\langle(\Omega^b(\xi^1,\dots,\xi^e))^2\rangle_{\boldsymbol{\xi}}\right] , \tag{4.6.111}$$

with solutions $\langle\psi^2\rangle_\psi = 0$ or

$$1 = \sum_{d} \frac{P_d(d)d(d-1)}{\langle d\rangle} \sum_{e,q} \frac{P_q(q)P_q(e-q)e}{2\langle q\rangle} \times \\ \times \left[(q-1)\langle(\Omega^a(\xi^1,\dots,\xi^e))^2\rangle_{\boldsymbol{\xi}} + (e-q)\langle(\Omega^b(\xi^1,\dots,\xi^e))^2\rangle_{\boldsymbol{\xi}}\right] . \tag{4.6.112}$$

Both eq. (4.6.107) and (4.6.111) have a trivial solution with zero moments or a more complicated one, which gives us the line where moments become different from zero. This will depend on the system's temperature, the B-B interaction strength, the graph topology encoded in the distributions $P_d(d), Q(e)$ and the distribution of the disordered interactions ξ's. In the next subsection, we will study the bifurcations for different choices of the disorder. We will focus on the regular graph topology with vertex degree L and factor degree K; hence, we choose $P_d(d) = \delta_{d,L}$ and $P_q(q) = \delta_{q,K}$. We will obtain the critical line numerically via population dynamics simulations and analytically from (4.6.107, 4.6.111).

4.6.2.1 Symmetric pattern distributions

We first consider the case of symmetrically distributed ξ's, i.e. $\mathbb{P}(\xi) = \frac{1}{2}\delta_{\xi,+1} + \frac{1}{2}\delta_{\xi,-1}$. In this case there is no instability from growing means as the field distribution is always symmetric and, indeed, the rhs of (4.6.107) averages to zero. The bifurcation is, therefore, detectable from the instability of growing variances, while the mean remains zero. Specialising (4.6.108) to the regular graph case with $P_d(d) = \delta_{d,L}$ and $P_q(q) = \delta_{q,K}$ we get

$$1 = (L-1)\Bigg[(K-1)\left(\frac{\langle \sinh^2(\sqrt{\beta}y_1)\cosh^{K-2}(\sqrt{\beta}y_1)\cosh^{K}(\sqrt{\beta}y_2)\rangle_{\mathbf{y}}}{\langle \cosh^{K}(\sqrt{\beta}y_1)\cosh^{K}(\sqrt{\beta}y_2)\rangle_{\mathbf{y}}}\right)^2 \\ +K\left(\frac{\langle \sinh(\sqrt{\beta}y_1)\sinh(\sqrt{\beta}y_2)\cosh^{K-1}(\sqrt{\beta}y_1)\cosh^{K-1}(\sqrt{\beta}y_2)\rangle_{\mathbf{y}}}{\langle \cosh^{K}(\sqrt{\beta}y_1)\cosh^{K}(\sqrt{\beta}y_2)\rangle_{\mathbf{y}}}\right)^2\Bigg]. \tag{4.6.113}$$

We can easily compute the large T-limit ($\beta \to 0$) in (4.6.113): the rhs clearly tends to zero, hence no solutions with non-zero fields exist at high temperature. For $\beta \to \infty$ averages are dominated by large values of y_1, y_2 yielding $\sinh^2(\sqrt{\beta}y_1) \sim \cosh^2(\sqrt{\beta}y_1)$ and $1 = (L-1)(2K-1)$. Hence, for a regular graph there is no bifurcation as long as $L < \frac{2K}{2K-1}$. For the general β-dependence, we compute the

Gaussian averages over $\mathbf{y}$ explicitly, obtaining

$$1 = \frac{L-1}{(\omega(\beta,k,K))^2}\left[(K-1)\left(\theta_1(\beta,k,K)\right)^2 + K(\theta_2(\beta,k,K))^2\right] \ , \qquad (4.6.114)$$

where

$$\begin{aligned}
\omega(\beta,k,K) &= \sum_{\ell,f=0}^{K}\binom{K}{\ell}\binom{K}{f}\mathrm{e}^{\frac{\beta}{2(1-k^2)}((K-2\ell)^2+(K-2f)^2+2k(K-2\ell)(K-2f))} \ , \\
\theta_1(\beta,k,K) &= \sum_{p=0}^{2}\sum_{r=0}^{K-2}\sum_{s=0}^{K}(-1)^p\binom{2}{p}\binom{K-2}{r}\binom{K}{s}\mathrm{e}^{\frac{\beta}{2(1-k^2)}((K-2r-2p)^2+(K-2s)^2+2k(K-2r-2p)(K-2s))} \ , \\
\theta_2(\beta,k,K) &= \sum_{n,h=0}^{1}\sum_{j,v=0}^{K-1}(-1)^{n+h}\binom{1}{n}\binom{K-1}{j}\binom{K-1}{v}\binom{1}{h}\mathrm{e}^{\frac{\beta}{2(1-k^2)}((K-2j-2h)^2+(K-2n-2v)^2+2k(K-2j-2h)(K-2n-2v))} \ .
\end{aligned} \qquad (4.6.115)$$

The latter condition gives the critical line for the variances bifurcation in $T = 1/\beta$ as a function of L. We plot this line in fig. 4.18, in the presence and in the absence of idiotypic interactions. We also compare the theoretical bifurcation line in (4.6.114) with the data obtained from population dynamics simulations (markers). For $k = 0$ the line is in agreement with results obtained in [39]. Increasing the B-B interactions strength widens the interference region as one of the effects is to merge clusters (in this case of equal sizes) together.

As in the case with ferromagnetic interactions, the role of the critical line consists in separating the region of clonal cross-talk from the region with no interference. However, with disordered interactions the interference acts on the system in the form of inhomogeneous fields and will therefore represent an additional source of noise in the system (besides the thermal noise). For $T > T_c(K,L)$, the cavity fields are zero and each clique is signalling to a particular B clone, without feeling the interference of the others. In the neural networks jargon, the system works, in this region, as a parallel processor able to retrieve multiple cytokine patterns $\boldsymbol{\xi}^\mu$ simultaneously. For $T < T_c(K,L)$ the system is in the *clonal cross-talk region* or so-called *spin glass phase* (SG) [39], where random fields act on each clique, reducing the parallel processing capabilities and making the signalling process to B clones less effective.

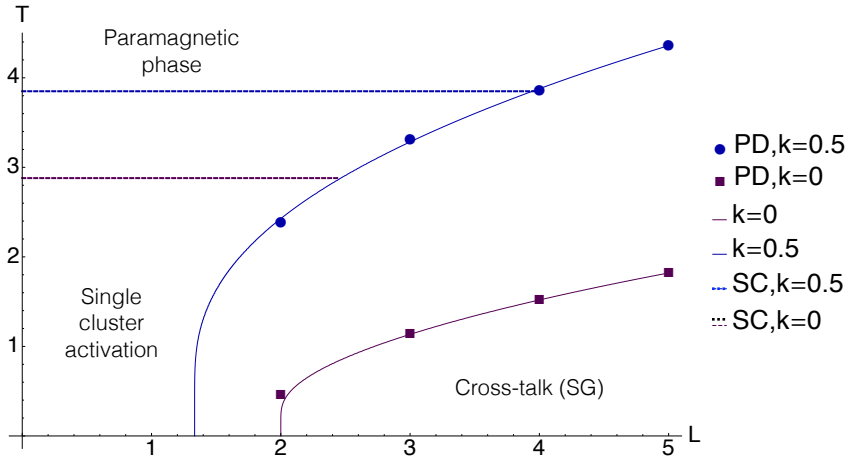

Figure 4.18: Critical line for symmetric patterns distribution (bifurcation in variance) in the plane (T, L) for different values of B-B interaction strength $k = 0, 0.5$. We consider a regular graph with factor degree $K = 2$. At high temperature the system is in the *paramagnetic phase*, where clusters are independent of each other. Crossing the (solid) lines, cavity fields become non-zero and the clonal interference increases entering the *cross-talk* or spin glass (SG) region [74]. Markers (PD) represent numerical results obtained via population dynamics simulations (population size $M = 10^4$, see A.2 for details). Dashed lines represent the single cluster (SC) activation temperatures, derived in sec. 4.4.1.

In addition, the dashed lines highlight the temperature T_{clust} at which single clusters become active (see sec. 4.4.1) for different k. Hence, lowering the temperature the system is subjected to both an increased clonal interference (crossing the solid lines) and to a unimodal-bimodal transition in the overlap distributions for single clusters. Increasing k, the temperature at which the cross-over transition happens increases, meaning that the system increases its tolerance to high noise levels.

4.6.2.2 Non-Symmetric pattern distributions

In this section, we add a degree of asymmetry $a \in [-1, +1]$ in the system, to study the bifurcations in the first moment. Hence, we consider the ξ's entries to be distributed according to $\mathbb{P}(\xi) = \frac{1+a}{2}\delta_{\xi,+1} + \frac{1-a}{2}\delta_{\xi,-1}$. Computing the ξ-averages in (4.6.106) and specialising the equations for the regular graph case

with $P_d(d) = \delta_{d,L}$ and $P_q(q) = \delta_{q,K}$ we get

$$a^{-2} = (L-1)\Bigg[(K-1)\frac{\langle \sinh^2(\sqrt{\beta}y_1)\cosh^{K-2}(\sqrt{\beta}y_1)\cosh^{K}(\sqrt{\beta}y_2)\rangle_{\mathbf{y}}}{\langle \cosh^{K}(\sqrt{\beta}y_1)\cosh^{K}(\sqrt{\beta}y_2)\rangle_{\mathbf{y}}} + K\frac{\langle \sinh(\sqrt{\beta}y_1)\sinh(\sqrt{\beta}y_2)\cosh^{K-1}(\sqrt{\beta}y_1)\cosh^{K-1}(\sqrt{\beta}y_2)\rangle_{\mathbf{y}}}{\langle \cosh^{K}(\sqrt{\beta}y_1)\cosh^{K}(\sqrt{\beta}y_2)\rangle_{\mathbf{y}}}\Bigg] . \quad (4.6.116)$$

Note that when $a \to 0$ the transition point diverges and we recover the symmetric case. First, we analyse the limits for high and zero temperature: for $\beta \to 0$ eq. (4.6.116) does not admit any solution, hence the first moment is zero. At zero temperature, i.e. $\beta \to \infty$, we have a transition to non-zero mean for $a^{-2} = (L-1)(2K-1)$. For a regular graph with vertex degree L and factor degree K, the general β-dependence can be obtained from

$$a^{-2} = \frac{L-1}{\omega(\beta,k,K)}\left[(K-1)\theta_1(\beta,k,K) + K\theta_2(\beta,k,K)\right] , \quad (4.6.117)$$

with

$$\begin{aligned}
\omega(\beta,k,K) &= \sum_{\ell,f=0}^{K}\binom{K}{\ell}\binom{K}{f}\mathrm{e}^{\frac{\beta}{2(1-k^2)}((K-2\ell)^2+(K-2f)^2+2k(K-2\ell)(K-2f))} ,\\
\theta_1(\beta,k,K) &= \sum_{p=0}^{2}\sum_{r=0}^{K-2}\sum_{s=0}^{K}(-1)^p\binom{2}{p}\binom{K-2}{r}\binom{K}{s}\mathrm{e}^{\frac{\beta}{2(1-k^2)}((K-2r-2p)^2+(K-2s)^2+2k(K-2r-2p)(K-2s))} ,\\
\theta_2(\beta,k,K) &= \sum_{n,h=0}^{1}\sum_{j,v=0}^{K-1}(-1)^{n+h}\binom{1}{n}\binom{K-1}{j}\binom{K-1}{v}\binom{1}{h}\mathrm{e}^{\frac{\beta}{2(1-k^2)}((K-2j-2h)^2+(K-2n-2v)^2+2k(K-2j-2h)(K-2n-2v))} .
\end{aligned} \quad (4.6.118)$$

It is worth stressing that, even in presence of a non-symmetric patterns distribution, there may be a bifurcation in the field variances at zero means, whose critical line will still be given by eq. (4.6.114). In this situation, the physical bifurcation is the first one taking place when lowering the temperature.

In fig. 4.19 we plot the critical line with $k = 0.5$ and $k = 0$ for $a = 1$, which retrieves the case of ferromagnetic interactions $\xi_i^\mu = +1 \; \forall \, i, \mu$ analysed in sec. 4.5. Here the critical line is given by bifurcation of the means. The cross-talk region is enlarged by the presence of B-B interactions. In this case, however, interference is constructive, due to the ferromagnetic nature of interactions. Finally, we also checked numerically that in the ferromagnetic case ($a = 1$) with no idiotypic

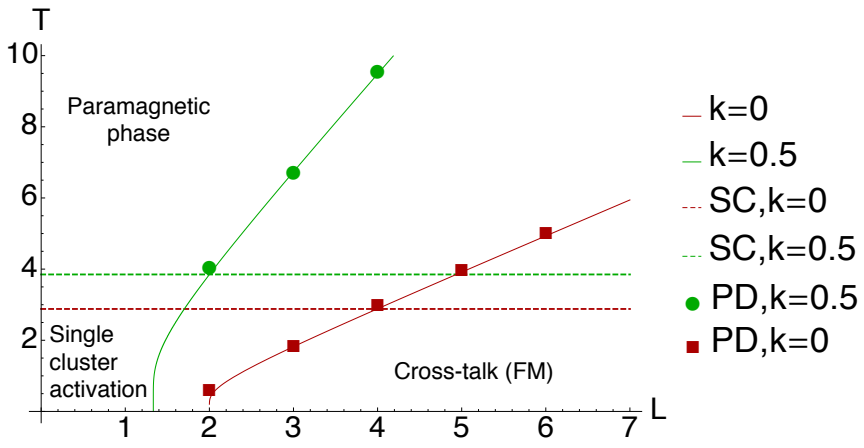

Figure 4.19: Critical line for the bifurcation of means for a regular graph with ferromagnetic interactions, i.e. $a = 1.$ and factor degree $K = 2$. At high temperature the system is in the *paramagnetic phase*, where each cluster within the graph acts independently from the others. Crossing the (solid) lines for different k's, cavity fields become non-zero and the clonal interference increases entering the *cross-talk* or ferromagnetic (FM) region. Markers (PD) represent numerical results obtained via population dynamics simulations (population size $M = 10^4$, see A.2 for details). Dashed lines represent the single cluster (SC) activation temperatures, derived in sec. 4.4.1.

interactions, $k = 0$, for the regular graph topology, we recover the critical line for the ferromagnetic transition for the Bethe lattice [71, 76].

4.7 Summary

In this chapter, we analysed the interacting system of B and T clones, main constituents of the adaptive immune system. In particular, we investigated the effect of idiotypic interactions among B clones, using belief propagation techniques and extending a previous model studied in the sub-extensive regime [61]. We derived cavity equations for the factor graph associated to the system and we discussed preliminary simple cases, which can be solved analytically. In particular, we considered the paramagnetic phase and ferromagnetic interactions and monitored the behaviour of the overlap distributions. The latter is useful to cal-

culate the B clone size distributions varying the temperature and the strength of the idiotypic interactions. We also derived the activation line where the overlap distribution shows a crossover behaviour from a phase of inactive to a phase of active B clones. We find, in particular, that B-B interactions increase the activation temperature making the system more resilient to noise. Indeed, one of the effects of the idiotypic network is to merge groups of T clones signalling to complementary B clones together, producing more stable signals and prolonging memory in the system. Having a model that predicts the most important parameters affecting the B clonal distribution is particularly welcome as this observable has recently become available.

She said that there was comfort to be found in the permanence of mathematical truths, in the lack of arbitrariness and the absence of ambiguity. In knowing that the answers may be elusive, but they could be found. They were there, waiting, chalk scribbles away. "Nothing like life, in other words," he said. "There, its questions with either no answers or messy ones."

K. Hosseini, *And the mountains echoed.*

5

Conclusions and Outlooks

In this thesis we have considered a statistical mechanics approach to modelling B-T lymphocytes interactions in the adaptive immune system. The motivation behind this research project is twofold. On one hand, the ever-growing data in experimental immunology that have been made lately available calls for a unifying and predictive framework to interpret and understand the results and suggest new possible experiments [77]. On the other hand, the analysis of the properties of associative memories with diluted patterns is interesting for possible applications in artificial intelligence.

In chapter 2, 3 we focused on the dynamical analysis of the retrieval capabilities of the diluted associative memories. While lacking so far for this type of memories, the dynamical approach generally provides richer information than the static analysis, which can only describe the behaviour of the system at the steady state, and does not require that the system is in equilibrium. Moreover, it may facilitate the comparison between theoretical predictions and experiments in potential applications, such as immunology, where data are normally available

for time-dependent quantities, e.g. responses in time to vaccination or infections.

In particular, in chapter 2 we show that the sparsity of the B-T interactions makes the system able to activate multiple B clones in parallel. This multitasking capability is one of the core features of the immune system, which in normal conditions, can control and block several simultaneous antigenic invasions. In addition, we find that the parallel activation of B clones may occur in a *symmetric* fashion, where all infections are fought with the same strength, or in a *hierarchical* way, where the system prioritises immune responses against specific pathogens. We are able to identify the system's parameters, such as noise, number of links, number of different infections, which induce the switch from the symmetric to the hierarchical operational mode. This might be potentially useful to investigate the causes of dramatic failures in the functioning of the immune system. The switch to a hierarchical immune response is in fact mostly related to the presence of strong infections, such as hepatitis and autoimmune disorders: while the immune system invests the highest amount of resources tackling the main disease, the progression of minor infections may become lethal [78, 79]. Hence, one of the open question in theoretical immunology is understanding how the immune system prioritises immune responses against different pathogens, executed in parallel.

In chapter 3 we extended our model to incorporate important biological mechanisms and features of real immune systems. We introduced the effect of cell receptor promiscuity, idiotypic interactions as well as new players, the antigens. Our aim was to investigate their effects on the clonal activation, in terms of critical temperature of activation and response strength. We showed that clones with smaller promiscuity are less likely to be activated (sec. 3.2) and idiotypic interactions contribute to the overall stability of the network, in terms of signals propagation (sec. 3.3). Furthermore, we investigated how the immune system responds to antigens, showing that multiple antigens create an interference that leads to less effective response to individual antigens and a reduction in the *basal activity* of the non-infected B clones (sec. 3.4). For higher noise level, the system

tends to simultaneously fight *all* the antigens, at the price of a weaker response strength, whereas for lower noise level it will prioritise some infections over others. Finally, we discussed the immune system short-term memory, which emerges as a hysteresis effect to the antigenic field.

Finally, in chapter 4 we focused on the study of the statics of finitely connected systems of interacting B and T lymphocytes in the presence of idiotypic interactions. Using belief-propagation techniques [71], we analysed the activation properties of single B clones and the clonal interference between many of them within the network. On the experimental side, our study suggests that oscillations of single clone populations may be observed on typical timescales. The availability of experimental data on clonal expansion and contraction patterns in time, in the absence of antigens, may thus feed important information into the model.

Moreover, we could monitor the B clone size distribution, which is attracting interest lately, thanks to the development of experimental techniques needed to extract it [67]. In particular, we were able to study the B clone size distribution behaviour in different regions of the parameters, e.g. noise, B-B interaction strength and graph topology. The so-called high-throughput sequencing of B-cell repertoires and large-scale genetic profiling studies are, indeed, increasingly used to extract information on the immune response in different pathological or healthy situations. So far, these techniques have been exploited to gain a better understanding of autoimmune diseases, cancer and of the ageing of the immune systems [80–82].

Pathways for future research may include the modelling of T-T interactions, which are known to play an important role in the self/non-self discrimination process [5, 16], and T clones receptor promiscuity effects [54]. Moreover, the assumption that each T clone is able to secrete both excitatory and inhibitory cytokines may be replaced with the more realistic scenario of two populations

of T clones, each responsible for sending one type of signal only. Indeed, there are lymphocytes sub-families and other cells, e.g. as naive B-cells, plasma cells, cytotoxic T-cells, helper T-cells, etc., or other cells, e.g. dendritic cells [1], which have not been included in our model so far and whose role is fundamental in specific processes in healthy/unhealthy organisms [1]. The assumption that B clones and T clones evolve in the same thermal noise is also quite strong and biologically debated. This could shed light on the effect of cellular environment on the immune response as shown recently in [83].

The ultimate goal of the research in this interdisciplinary field is to create valid theoretical immune system models, which might be used in the future to design *quantitative immunotherapies*, tailored to specific diseases and individual patients. Indeed, this challenge is still far from a concrete realisation but at least important efforts are put in place from both sides (the theoretical and experimental ones), to meet halfway. It will not be an easy task, however quoting the ancients, *per aspera ad astra*.

A

Simulation codes

In this chapter we list the codes and pseudo-codes for the simulations used in chapters 2, 3, 4. In particular, in sec. A.1 we will discuss the general steps for the Monte Carlo simulations with Glauber dynamics and the algorithm for the case with two patterns and fixed dilution (analysed in ch. 3). All other cases discussed in chapters 2, 3 follow from simple modifications of the codes presented. In sec. A.2 we will present the population dynamics code to compute the critical line for the field bifurcations (see sec. 4.6.2) and the calculation of the update equation for the cavity field.

A.1 Monte Carlo simulations

Here we list the pseudo-code to simulate spins Glauber dynamics.

(a) Define the parameters of the system: the temperature β, the dilution by tuning c, γ, the number of patterns P and the size of the system N. In our simulations we work with $N = \mathcal{O}(10^4)$ spins.

(b) Generate the patterns ξ_i^μ according to their distribution $\mathbb{P}(\{\boldsymbol{\xi}\})$ and create the interaction matrix J_{ij}.

(c) Select possible initial states of the system $\boldsymbol{\sigma}_0$ and calculate their overlap with the patterns.

(d) Start the dynamics, updating the spins with probability given by the Glauber rates $\omega_i = \frac{1}{2}[1 + \sigma_i \tanh(\beta h_i(\boldsymbol{\sigma}))]$ (defined in (2.2.7)).

(e) Monitor and store the magnetisations at each time step.

```
#include "include_lib/general.h"
#include "string.h"

void parameters(int*, int, int**, float*);
int main()
{
  FILE *fopen(),*fp;
  char name[30],filename[30],label[30];
  int i,j,k,l,site,mu,N,flips,h;
  int **J, temp;
  int *s,*ivector(), **pat, **imatrix();
  float *m,**data,**matrix(),*vector(),T,glauber,d,gamma,c;
  double RANDOM(),pow(),tanh();

  SETRANDOM();
  // input file //

  fp=fopen("cycle.in","r");
  fscanf(fp,"%d",&N);              // number of spins//
  fscanf(fp,"%f",&T);              //temperature//
  fscanf(fp,"%f",&gamma);          // less then 1//
  fscanf(fp,"%f",&c);              // dilution//
  fscanf(fp,"%s",name);            // file name//
  fclose(fp);

  fp=fopen("status","w");
```

```
27  fprintf(fp,"N=      %d\n",N);
28  fprintf(fp,"file=  %s\n",name);
29
30  fprintf(fp,"allocations ...\n"); fclose(fp);
31  data=matrix(0,200,1,2);   s=ivector(1,N);
32  J=imatrix(1,N,1,N);       pat=imatrix(1,2,1,N);
33 m=vector(1,2);
34  fp=fopen("status","a"); fprintf(fp,"patterns and interactions ...\
     n");
35  fclose(fp);
36
37  d= c/(pow(N,gamma)); //dilution parameter//
38  fp=fopen("status", "a"); fprintf(fp,"dilution= %f\n", d);
39  fclose(fp);
40
41  //Patterns \xi= +1,-1,0 //
42
43  for(mu=1;mu<=2;mu++) for(i=1;i<=N;i++) {
44       if(RANDOM()> d) pat[mu][i]= 0;
45       else { if(RANDOM()>0.5) pat[mu][i]= 1;
46  else        pat[mu][i]= -1;
47       }
48
49    }
50  fp=fopen("status","a"); for(i=1;i<=N;i++){
51    for(mu=1;mu<=2;mu++){
52       fprintf(fp,"%d\n",pat[mu][i]);}} fclose(fp);
53
54  //Interactions J_ij//
55
56  for(i=1;i<=N;i++){
57    for(j=1;j<=i;j++){
58       temp=0;
59       for(mu=1;mu<=2;mu++){
60          temp += ((pat[mu][i])*(pat[mu][j]));
61       }
```

```
62        J[i][j]=J[j][i]=temp;
63      }
64      J[i][i]=0;
65    }
66    fp=fopen("status","a");
67    for(i=1;i<=N;i++){
68      for(j=1;j<=N;j++) fprintf(fp,"%d",J[i][j]);
69      fprintf(fp,"\n");
70    }
71    fclose(fp);
72
73    //Iterations-choose different initial states//
74
75    fp=fopen("status","a"); fprintf(fp,"\niterations:\n"); fclose(fp);
76    for(i=1;i<=25;i++) {
77        switch(i) {
78      case 1: for(j=1;j<=N;j++) {
79    if(pat[1][j]==pat[2][j]){
80      if(pat[1][j]==0){
81        if(RANDOM()>0.5) s[j]=1;
82        else s[j]=-1;}
83      else {s[j]=pat[1][j];}}
84    else {if(RANDOM()>0.5) s[j]=1;
85      else s[j]=-1;}
86        }
87        break;
88
89  case 2: for(j=1;j<=N;j++) {
90    if(pat[1][j]==pat[2][j]){
91      if(pat[1][j]==0){
92        if(RANDOM()>0.5) s[j]=1;
93        else s[j]=-1;}
94      else {s[j]= -(pat[1][j]);}}
95    else {if(RANDOM()>0.5) s[j]=1;
96      else s[j]=-1;}
97        }
```

```
98        break;
99
100   case 3: for(j=1;j<=N;j++) {
101    if(pat[1][j]==pat[2][j]){
102      if(pat[1][j]==0){
103        if(RANDOM()>0.5) s[j]=-1;
104        else s[j]=1;}
105      else {s[j]= -(pat[1][j]);}}
106    else {if(RANDOM()>0.5) s[j]=1;
107      else s[j]=-1;}
108        }
109        break;
110
111   case 4: for(j=1;j<=N;j++) {
112    if(pat[1][j]==pat[2][j]){
113      if(pat[1][j]==0){
114        if(RANDOM()>0.5) s[j]=-1;
115        else s[j]=1;}
116      else {s[j]= pat[1][j];}}
117    else {if(RANDOM()>0.5) s[j]=-1;
118      else s[j]= 1;}
119        }
120        break;
121
122        default: for(j=1;j<=N;j++) {
123      if(RANDOM()<0.5) s[j]=1;
124      else             s[j]=-1;
125    }
126        }
127
128        parameters(s,N,pat,m);
129        fp=fopen("status","a");
130        for(mu=1;mu<=2;mu++) data[0][mu]=m[mu];
131        fprintf(fp,"(m1,m2)=(%f,%f)\n",m[1],m[2]);
132        fclose(fp);
133
```

```
134       if(i<26) sprintf(label,"%1d",i);
135       else     sprintf(label,"26");
136       strcpy(filename,name); strcat(filename,label);
137
138       flips=(int)(0.5 + 0.1*N);
139
140       for(j=1;j<=200;j++) {
141   for(k=1;k<=flips;k++) {
142     site=(int)(RANDOM()*N)+1;
143
144     h= 0;
145     for(l=1;l<=N;l++) {
146       if(s[l]>0)  {h+=J[site][l];}
147       else        {h-=J[site][l];}
148     }
149
150     glauber = 0.5 * (1.0 + tanh((float)h/(N*T)));
151     if(RANDOM()<glauber) s[site]=1;
152     else                 s[site]=-1;
153   }
154         parameters(s,N,pat,m);
155         data[j][1]= m[1];
156         data[j][2]= m[2];
157       }
158
159       //write on file 'status' time step, m1, m2
160       fp=fopen(filename,"w");
161       for(j=0;j<=200;j++) fprintf(fp,"%f %f %f\n",(float)(0.1*j),
    data[j][1],data[j][2]);
162       fclose(fp);
163   }
164   free_matrix(data,0,200,1,2);
165   free_vector(m,1,2);
166   free_imatrix(J,1,N,1,N);
167   free_imatrix(pat,1,2,1,N);
168   free_ivector(s,1,N);
```

```
169   return(0);
170 }
171
172 //Function-parameters compute magnetisations//
173
174 void  parameters(int *s, int N, int **pat, float *m){
175   int i,mu, temp;
176   for(mu=1;mu<=2;mu++){
177     temp=0;
178     for(i=1;i<=N;i++){temp += ((s[i])*(pat[mu][i]));}
179     m[mu]=(float)temp/((float)N);
180   }
181 }
```

Listing A.1: Monte Carlo simulations

A.2 Population dynamics algorithm

In this section we provide fomulae to compute the cavity fields for the population dynamics algorithm. Starting form the parametrisation of $P_{\backslash\mu\hat{\mu}}(\sigma_i)$ introduced in sec. 4.6

$$P_{\backslash\mu\hat{\mu}}(\sigma_i) \propto \mathrm{e}^{\beta\phi_{i\to\mu\hat{\mu}}\sigma_i} \ , \tag{1.2.1}$$

we can express the cavity field $\phi_{i\to\mu\hat{\mu}}$ in terms of the cavity marginals

$$\phi_{i\to\mu\hat{\mu}} = \frac{1}{2\beta}\log\left[\frac{P_{\backslash\mu\hat{\mu}}(+1)}{P_{\backslash\mu\hat{\mu}}(-1)}\right] \ , \tag{1.2.2}$$

which satisfy the recursion

$$P_{\backslash\mu\hat{\mu}}(\sigma_i) = \prod_{\nu\in\partial i\backslash\mu\hat{\mu}} \sum_{\{\sigma_{k\in\partial\nu\backslash i}\},\{\sigma_{\ell\in\partial\hat{\nu}}\}} f_{\nu\hat{\nu}}(\sigma_i,\{\sigma_{k\in\partial\nu\backslash i}\},\{\sigma_{\ell\in\partial\hat{\nu}}\}) \prod_{k\in\partial\nu\hat{\nu}\backslash i} P_{\backslash\nu\hat{\nu}}(\sigma_k) \ . \tag{1.2.3}$$

obtained by combining (4.3.53) and (4.3.54). Inserting the explicit expression for the factors

$$f_{\nu\hat{\nu}}(\sigma_i,\{\sigma_{k\in\partial\nu\setminus i}\},\{\sigma_{\ell\in\partial\hat{\nu}}\}) = \left\langle \exp\left[\sqrt{\beta}\left(y_1\xi_i^\nu\sigma_i + y_1\sum_{k\in\partial\nu\setminus i}\xi_k^\nu\sigma_k + y_2\sum_{\ell\in\partial\hat{\nu}}\xi_\ell^{\hat{\nu}}\sigma_\ell\right)\right]\right\rangle_{\mathbf{y}} , \tag{1.2.4}$$

we obtain

$$P_{\setminus\mu\hat{\mu}}(\sigma_i) = \prod_{\nu\in\partial i\setminus\mu\hat{\mu}}\sum_{\{\sigma_{k\in\partial\nu\setminus i}\},\{\sigma_{\ell\in\hat{\nu}}\}}\int\mathcal{D}\mathbf{y}\exp\left[\sqrt{\beta}\left(y_1\xi_i^\nu\sigma_i + y_1\sum_{k\in\partial\nu\setminus i}\xi_k^\nu\sigma_k + y_2\sum_{\ell\in\partial\hat{\nu}}\xi_\ell^{\hat{\nu}}\sigma_\ell\right)\right] \times\prod_{k\in\partial\nu\hat{\nu}\setminus i}P_{\setminus\nu\hat{\nu}}(\sigma_k) , \tag{1.2.5}$$

with $\mathcal{D}\mathbf{y} = \sqrt{1-k^2}\frac{\mathrm{d}y_1\mathrm{d}y_2}{2\pi}\mathrm{e}^{-\frac{1}{2}\mathbf{y}^T\mathbf{C}^{-1}\mathbf{y}}$. Expressing everything in terms of the fields $\phi_{k\to\nu\hat{\nu}}$ and introducing the variables $X_\nu = \sum_{k\in\partial\nu\setminus i}\xi_k^\nu\sigma_k$ and $\hat{X}_\nu = \sum_{\ell\in\partial\hat{\nu}}\xi_\ell^{\hat{\nu}}\sigma_\ell$, yields

$$P_{\setminus\mu\hat{\mu}}(\sigma_i) = \prod_{\nu\hat{\nu}\in\partial i\setminus\mu\hat{\mu}}\sum_{X_\nu,\hat{X}_\nu}\int\mathcal{D}\mathbf{y}\mathrm{e}^{\sqrt{\beta}(y_1\xi_i^\nu\sigma_i+y_1X_\nu+y_2\hat{X}_\nu)}\sum_{\{\sigma_k\}}\prod_{k\in\partial\nu\setminus i}\mathrm{e}^{\beta\phi_{k\to\nu\hat{\nu}}\sigma_k}\delta_{X_\nu,\sum_{k\in\partial\nu\setminus i}\xi_k^\nu\sigma_k} \times\sum_{\{\sigma_\ell\}}\prod_{\ell\in\partial\hat{\nu}}\mathrm{e}^{\beta\phi_{\ell\to\nu\hat{\nu}}\sigma_\ell}\delta_{\hat{X}_\nu,\sum_{\ell\in\partial\hat{\nu}}\xi_\ell^{\hat{\nu}}\sigma_\ell} . \tag{1.2.6}$$

Finally, we can integrate over $\mathbf{y}$. The integral in (1.2.6) can be written in the form

$$\sqrt{1-k^2}\int\frac{\mathrm{d}y_1\mathrm{d}y_2}{2\pi}\mathrm{e}^{-\frac{1}{2}\mathbf{y}^T\mathbf{C}^{-1}\mathbf{y}+\mathbf{J}^T\mathbf{y}} , \tag{1.2.7}$$

with $\mathbf{J}^T = (\sqrt{\beta}(\xi_i^\nu\sigma_i + X_\nu), \sqrt{\beta}\hat{X}_\nu)$ and $\mathbf{y}^T = (y_1, y_2)$ and considering

$$\mathbf{C}^{-1} = \begin{pmatrix} 1 & -k \\ -k & 1 \end{pmatrix} . \tag{1.2.8}$$

Solving it by Gaussian integration gives

$$\sqrt{1-k^2}\int\frac{\mathrm{d}y_1\mathrm{d}y_2}{2\pi}\mathrm{e}^{-\frac{1}{2}\mathbf{y}^T\mathbf{C}^{-1}\mathbf{y}+\mathbf{J}^T\mathbf{y}}=\mathrm{e}^{\frac{1}{2}\mathbf{J}^T\mathbf{C}\mathbf{J}}\ , \tag{1.2.9}$$

with

$$\mathbf{C}=\begin{pmatrix}\frac{1}{1-k^2} & \frac{k}{1-k^2}\\ \frac{k}{1-k^2} & \frac{1}{1-k^2}\end{pmatrix}\ . \tag{1.2.10}$$

Hence, we get

$$\mathrm{e}^{\frac{1}{2}\mathbf{J}^T\mathbf{C}\mathbf{J}}=\exp\left(\frac{\beta}{2(1-k^2)}\left((\xi_i^\nu\sigma_i+X_\nu)^2+2k\hat{X}_\nu(\xi_i^\nu\sigma_i+X_\nu)+\hat{X}_\nu^2\right)\right)\ . \tag{1.2.11}$$

We insert this result in (1.2.6), obtaining

$$P_{\setminus\mu\hat{\mu}}(\sigma_i)=\prod_{\nu\hat{\nu}\in\partial i\setminus\mu\hat{\mu}}\sum_{X_\nu,\hat{X}_\nu}\exp\left(\frac{\beta}{2(1-k^2)}\left((\xi_i^\nu\sigma_i+X_\nu)^2+2k\hat{X}_\nu(\xi_i^\nu\sigma_i+X_\nu)+\hat{X}_\nu^2\right)\right)\times$$
$$\times\sum_{\{\sigma_k\}}\prod_{k\in\partial\nu\setminus i}\mathrm{e}^{\beta\phi_{k\to\nu\hat{\nu}}\sigma_k}\delta_{X_\nu,\sum_{k\in\partial\nu\setminus i}\xi_k^\nu\sigma_k}\sum_{\{\sigma_\ell\}}\prod_{\ell\in\partial\hat{\nu}}\mathrm{e}^{\beta\phi_{\ell\to\nu\hat{\nu}}\sigma_\ell}\delta_{\hat{X}_\nu,\sum_{\ell\in\partial\hat{\nu}}\xi_\ell^{\hat{\nu}}\sigma_\ell}\ . \tag{1.2.12}$$

Lastly, we plug (1.2.12) in (1.2.2) in order to get the expression for the update

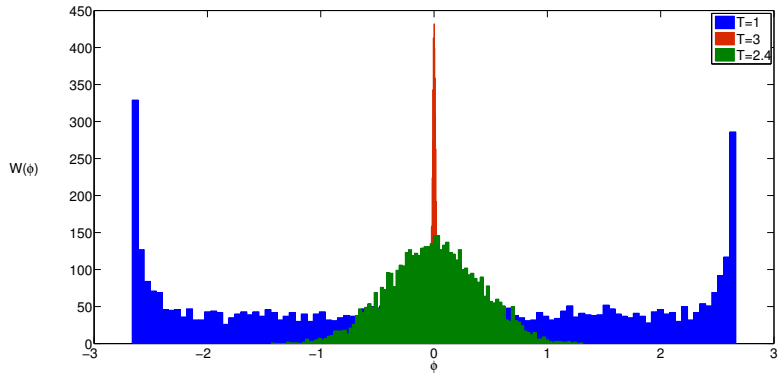

Figure A.1: Plot of $W(\phi)$ fields distribution for a regular random graph with $L=2=K$, $k=0.5$ and bias $a=0$ (symmetric entries) for $T=1,2.4,3$. Crossing the critical line of interference (4.6.114) the variance of the distribution increases.

of the cavity field

$$\phi_{i\to\mu\hat{\mu}} = \frac{1}{2\beta}\sum_{\nu\hat{\nu}\in\partial i\setminus\mu\hat{\mu}} \log\left\{\left[\sum_{X_\nu,\hat{X}_\nu} \mathrm{e}^{\left(\frac{\beta}{2(1-k^2)}\left((\xi_i^\nu+X_\nu)^2+2k\hat{X}_\nu(\xi_i^\nu+X_\nu)+\hat{X}_\nu^2\right)\right)}\times\right.\right.$$
$$\times\sum_{\{\sigma_k\}}\prod_{k\in\partial\nu\setminus i}\mathrm{e}^{\beta\phi_{k\to\nu\hat{\nu}}\sigma_k}\delta_{X_\nu,\sum_{k\in\partial\nu\setminus i}\xi_k^\nu\sigma_k}\sum_{\{\sigma_\ell\}}\prod_{\ell\in\partial\hat{\nu}}\mathrm{e}^{\beta\phi_{\ell\to\nu\hat{\nu}}\sigma_\ell}\delta_{\hat{X}_\nu,\sum_{\ell\in\partial\hat{\nu}}\xi_\ell^{\hat{\nu}}\sigma_\ell}\Bigg]\Bigg/$$
$$\left[\sum_{X_\nu,\hat{X}_\nu}\exp\left(\frac{\beta}{2(1-k^2)}\left((-\xi_i^\nu+X_\nu)^2+2k\hat{X}_\nu(-\xi_i^\nu+X_\nu)+\hat{X}_\nu^2\right)\right)\times\right.$$
$$\left.\left.\times\sum_{\{\sigma_k\}}\prod_{k\in\partial\nu\setminus i}\mathrm{e}^{\beta\phi_{k\to\nu\hat{\nu}}\sigma_k}\delta_{X_\nu,\sum_{k\in\partial\nu\setminus i}\xi_k^\nu\sigma_k}\sum_{\{\sigma_\ell\}}\prod_{\ell\in\partial\hat{\nu}}\mathrm{e}^{\beta\phi_{\ell\to\nu\hat{\nu}}\sigma_\ell}\delta_{\hat{X}_\nu,\sum_{\ell\in\partial\hat{\nu}}\xi_\ell^{\hat{\nu}}\sigma_\ell}\Bigg]\right\}. \tag{1.2.13}$$

This expression for the fields update will be used in the population dynamics algorithm. We summarise here the main steps of the algorithm:

(a) Define the degree distributions $Q(e)$ and $P_d(d)$ - in our algorithm we used a regular graph with vertex degree L and factor degree K.

(b) Generate the links $\boldsymbol{\xi} = (\xi^1, \ldots, \xi^e)$ as e i.i.d. random variables with probability distribution $P(\xi) = \frac{1+a}{2}\delta_{\xi,1} + \frac{1-a}{2}\delta_{\xi,-1}$, dependent on the parameter a.

(c) Extract a population composed of M fields ϕ_i, $i = 1, \ldots M$ uniformly in the interval $[-f_{max}, f_{max}]$: their histogram defines the zero-step approximation of the field distribution $W_\psi^0(\Psi)$.

(d) Start the iteration: choose e, d and generate ξ's.

(e) Choose $e-1$ fields randomly: compute the updated field ϕ_{new} using (1.2.13).

(f) Choose randomly one field ϕ_i and replace it with the field just computed.

In figure A.1 we plot the field distribution $W(\phi)$ in different regions of the phase diagram fixing $k = 0.5$ and varying the temperature. Patterns are drawn from a

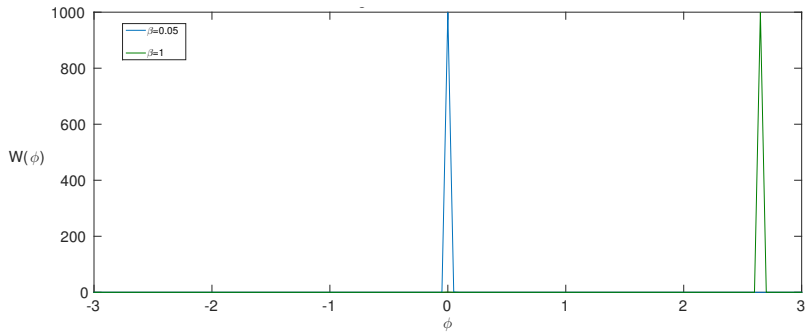

Figure A.2: Ferromagnetic case ($a = 1$) on a regular graph $L = 2, K = 2$ with $k = 0.5$ for $\beta = 1$ (blue), 0.05 (green). When $\beta = 1$ the distribution is peaked around $\phi = 2.66$ as predicted analytically from (4.5.89). Crossing the critical line of interference (4.6.117) the distribution is peaked around a non-zero value of the field ϕ.

symmetric distribution, i.e. $a = 0$. Starting from the high temperature regime, where the distribution is delta-peaked in $\phi = 0$, the variance increases decreasing the temperature, when crossing the critical line. For ordered interactions, i.e. $a = 1$, the distribution, plotted in figure A.2, shows a transition from a delta peak in $\phi = 0$ at high temperature, to a peak at non-zero values, when the critical line is crossed.

```
#include <stdio.h>
#include <math.h>
#include <stdlib.h>
#include "include_lib/general.h"
#include "string.h"
#define M 10000     /*fields population size*/
#define SUPPH 0.05       /*support of W(psi) - initial */
#define SUPPHMAX 5.0    /*max support for the final histogram of psi
    */
#define DH 0.05          /* W(psi) bin */
#define CAMPO "campo.dat"
#define MOMENT "moments.dat"
#define PM "pm.dat"
#define MOMENTIT "momentstime.dat"
#define Q 2 /*vertex  degree */
#define R 0.0 /* xi's bias */
```

```
16 #define K 0.0 /* BB interaction strength */
17
18 /*FUNCTIONS*/
19 int UniformRandomNumber (int); /* Generate Uniform Random Number */
20 int PoissonRandomNumber(double); /* Generate Uniform Random Number
       */
21 void MakeIsto (double *,double *); /* Produce histogram */
22 int Delta (int, int); /*delta function*/
23 void fast_d2b(unsigned long int, int *,int ); /* function needed to
       create the string sigma*/
24 double Compute_Var(double *psi,int m); /*Compute the variance of the
        field distribution*/
25 double Compute_Mean(double *psi,int m); /*Compute the mean of the
       field distribution*/
26 double Compute_Mom2(double *,int); /*compute second moment*/
27 void  product(int **, double *, int ,int ,int , double *,double); /*
       compute the product over k/l of e^sigma*psitest*/
28 void  magnetization(int **, int **, int **, int , int ,int ,int); /*
        Compute magnetization to insert in the delta function*/
29 void vectorxxh(int **, int , int ); /*create vectors x[a][i] and xh[
       a][i]*/
30
31 static double *psitest;
32 int main (void){
33   double newpsi, *prodk, *prodl, *fieldpart;
34   int Nbit, j,size,e,d,q,qh,i,r,l, **pat,t, *bink, *binl;
35   int **sigmak, **sigmal, **mk,**ml,***x,***xh, **imatrix(), *ivector
       ();
36   double **dmatrix(), *dvector(),psi[M],mean,var;
37
38   Nbit=2*((int)(SUPPHMAX/DH))+2;
39   double istopop[Nbit];
40
41   FILE *Campofile;
42   Campofile=fopen(CAMPO,"w"); // store  cavity field distribution
43   FILE *moments;
```

```
44  moments=fopen(MOMENT,"w"); //store moments of the cavity field
      distribution
45  FILE *momentstime;
46  momentstime=fopen(MOMENTIT,"w"); //store moments at each step of
      the pop dyn to check convergence
47  q=Q; // factor degree factor /mu - regular topology
48  qh=Q; //degree of factor /hat/mu  - regular topology
49  double B; //beta
50
51  /*start cycle in temperature and in factor degree d */
52
53  for(d=1;d<=7;d++){
54    for(B=0.7;B<=2;B+=0.02){
55    /* field initialization */
56  {int i;
57    for(i=0;i<M;i++){
58      //psi[i]=0; (check that psi=0 is  a Fixed Point)
59      psi[i] = (2*SUPPH*rand()/(RAND_MAX + 1.0))-SUPPH;
60    }
61  }
62
63    /* POPULATION DYNAMICS */
64  {for(t=0;t<100;t++){
65      mean=0.0;
66      var=0.0;
67      for(r=0;r<1000;r++){
68  e=q+qh; /* neighbourhood of mu,muh factor */
69  if(e>2){psitest=(double *)malloc(e*sizeof(double));}
70  {int s; for(s=0;s<(e-1);s++){psitest[s]=psi[UniformRandomNumber(M)
      ];}}
71  // choose e-1 field from the population at random
72
73
74  pat=imatrix(1,d-1,1,e);
75        //Generate xi's
76        int k;
```

```
77          if(d>1){for(k=1; k<=d-1;k++) {
78              for(i=1;i<=e;i++){
79          if(RANDOM()> (1+R)/2){pat[k][i]= +1;}
80          else{pat[k][i]= -1;}
81  }
82      }}
83
84          //Generate matrix sigmak
85          int m; m=(int)pow(2,q-1); //size of sigmak
86          int bink[q];
87          sigmak=imatrix(0,m-1,1,q-1);
88          for(i=0;i<=m-1;i++)
89      {fast_d2b(i,bink,q);
90        for(j=1;j<=q-1;j++)
91          { if(bink[j-1]==1) sigmak[i][j]=1;
92      else if (bink[j-1]==0) sigmak[i][j]=-1;
93          }
94      }
95
96          //Generate matrix sigmal
97          int n; n=(int)pow(2,qh); //size of sigmal
98          int binl[qh];
99          sigmal=imatrix(0,n-1,1,qh);
100         for(i=0;i<=(n-1);i++)
101     {fast_d2b(i,binl,qh);
102             for(j=1;j<=(qh);j++)
103         { if(binl[j-1]==1) sigmal[i][j]=1;
104                 else if (binl[j-1]==0) sigmal[i][j]=-1;
105         }}
106
107         //Product function prodk
108         {int gg=-1;
109     prodk=dvector(0,m-1);
110     product(sigmak,psitest,q-1,gg,m,prodk,B);
111    }
112
```

```
113         //Product function prodl
114         {int ll=0;
115    prodl=dvector(0,n-1);
116    product(sigmal,psitest,qh,q-2,n,prodl,B);
117   }
118
119  //Compute magnetization of cluster
120         mk=imatrix(1,d-1,0,m-1);
121         ml=imatrix(1,d-1,0,n-1);
122         magnetization(sigmak, pat, mk, d-1,m, q-1,0);
123         magnetization(sigmal, pat, ml, d-1,n, qh,q-1);
124
125
126  //Creat x[a],xh[a] vectors (to sum over them)
127         x=imatrix(1,d-1,-(q-1),q-1);
128         xh=imatrix(1,d-1,-(qh),qh);
129         vectorxxh(x, q-1, d-1);
130         vectorxxh(xh, qh, d-1);
131
132  //Compute new field
133  fieldpart=dvector(1,d-1);
134         double kk;
135         kk=pow(K,2);
136         {int r,s,a;
137    double temp,norm;
138    temp=0;
139    norm=0;
140    for (a=1; a<=d-1; a++) {for(r=-(q-1);r<= q-1;r+=2){for(s=-qh;s<=
     qh;s+=2){
141    for (i=0; i<=m-1; i++) {
142      for (j=0; j<=n-1; j++){
143         temp +=((exp((((B*(pat[a][e]+x[a][r]+xh[a][s])*(pat[a][e]+x[
     a][r]+xh[a][s]))))/(2*(1-kk))-B*((pat[a][e]+x[a][r])*xh[a][s])/(1+
     K)))*prodk[i]*Delta(x[a][r],mk[a][i])*prodl[j]*Delta(xh[a][s],ml[
     a][j])));
```

```
144          norm +=((exp((((B*(-(pat[a][e])+x[a][r]+xh[a][s])*(-(pat[a][
     e])+x[a][r]+xh[a][s]))))/(2*(1-kk))-B*((-(pat[a][e])+x[a][r])*xh[a
     ][s])/(1+K)))*prodk[i]*Delta(x[a][r],mk[a][i])*prodl[j]*Delta(xh[
     a][s],ml[a][j])));}}}}
145        fieldpart[a]=(double)temp/(double)norm;
146 }}
147
148   // Compute Sum_a of log (fieldpart[a])
149   double psitemp; psitemp=0;
150   int a;
151   for (a=1; a<=d-1; a++){
152     psitemp+=(log(fieldpart[a]));}
153
154   // New field psi - random update
155   newpsi=((double)psitemp)/(double)(2*B);
156
157   /*Update a field at random among the available ones*/
158          psi[UniformRandomNumber(M)]=newpsi;
159
160   /*COMPUTE MEAN & VARIANCE */
161   double mm,vv; // mean and variance variables
162
163  mm=Compute_Mean(psi,M);
164          vv=Compute_Var(psi,M);
165          mean+=mm;
166   var+=vv;
167
168          free_dvector(prodk,0,m-1);
169          free_dvector(prodl,0,n-1);
170          free_imatrix(sigmak,0,m-1,1,q-1);
171          free_imatrix(pat,1,d-1,1,e);
172          free_imatrix(sigmal,0,n-1,1,qh);
173          free_imatrix(mk,1,d-1,0,m-1);
174          free_imatrix(ml,1,d-1,0,n-1);
175          free_imatrix(x,1,d-1,-(q-1),q-1);
176          free_imatrix(xh,1,d-1,-qh,qh);
```

```
177   free_dvector(fieldpart,1,d-1);
178
179         if(e>2){free(psitest);}
180       }//close r loop
181       mean=mean/M;
182       var=var/M;
183       fprintf(momentstime,"%d %f %f \n",t, mean,var); // useful to
      monitor convergence in time
184      }
185   }// close t loop
186
187         double vfin,mfin,m2fin;
188         mfin=Compute_Mean(psi,M);
189         vfin=Compute_Var(psi,M);
190         m2fin=Compute_Mom2(psi,M);
191
192   for(i=1;i<=M;i++){fprintf(Campofile, "%f\n",psi[i]);}
193   fprintf(moments,"%d %f %f %f %f\n",d,(1/B), mfin,vfin,m2fin);
194
195    }}
196  fclose(moments);
197  fclose(Campofile);
198  fclose(momentstime);
199   return 0;
200 }
201
202 /*************************FUNCTIONS**********************
203
204 /* Compute mean of W(psi)*/
205
206 double Compute_Mean(double *psi,int m){
207   double mean=0.0;
208   int i;
209   for(i=0;i<m;i++){mean+=psi[i];}
210   return mean/m;
211 }
```

```
212
213 /*Compute variance of W(psi)*/
214
215 double Compute_Var(double *psi,int m){
216   double var=0.0;
217   double mean;
218   int i;
219   mean=Compute_Mean(psi,m);
220   for(i=0;i<m;i++){var+=((psi[i]-mean)*(psi[i]-mean));}
221   return var/(m-1);
222 }
223
224 double Compute_Mom2(double *psi,int m){
225     double var=0.0;
226     double mean;
227     int i;
228     for(i=0;i<m;i++){var+=((psi[i])*(psi[i]));}
229     return var/(m);
230 }
231
232 /* histogram of the population*/
233
234 void MakeIsto (double *psi,double *istopop){
235   int i;
236   int n;
237   for(i=0;i<M;i++){
238     if(psi[i]>=0){
239       if(psi[i]>SUPPHMAX){istopop[-2]=istopop[-2]+1.0/M;}
240       else{
241   n=(int)(psi[i]/DH);
242   istopop[2*n]=istopop[2*n]+1.0/(M*DH);
243       }
244     }
245     else{
246       if(psi[i]<-SUPPHMAX){istopop[-1]=istopop[-1]+1.0/M;}
247       else{
```

```
248   n=(int)(-psi[i]/DH);
249   istopop[2*n+1]=istopop[2*n+1]+1.0/(M*DH);
250       }
251     }
252   }
253   return ;
254 }
255
256 /*int between 0-m with uniform probability*/
257
258 int UniformRandomNumber (int m){
259   return (int)((double)m*rand()/(RAND_MAX+1.0));
260 }
261
262 /*Delta Function */
263
264 int Delta (int n, int m){
265   if(n==m){return 1;}else {return 0;}
266 }
267
268 /* Create all possible combinations of sigmas*/
269
270 void fast_d2b(unsigned long int x, int *bin,int size) {
271   int i;
272   /*
273     Fast decimal-to-binary conversions
274     Author: Giovanni Motta (gim@ieee.org)
275   */
276   for(i=0;i<size;i++)*(bin++) = (x >> i) & 0x1;
277 }
278
279 /* Compute products over k/l of e^(psi*sigma) magnetization */
280
281 void  product(int **s, double *psi, int v,int l,int sizesigma,
      double *prod, double b){
282   int i,k;
```

```
283   double temp, pp;
284   for(i=0;i<=sizesigma-1;i++){
285     temp=0;
286     pp=1;
287     for(k=1;k<=v;k++){temp=((s[i][k])*(psi[k+l]));
288       pp *=exp(b*temp);}
289     prod[i]=(double)pp;
290   }
291 }
292
293 /*Compute magnetization mk[a][i] &  ml[a][i] */
294
295 void  magnetization(int **s, int **pat, int **m, int a, int
      sizesigma,int v, int l){
296   int i,mu,k, temp;
297   for(mu=1;mu<=a;mu++){
298     for (i=0;i<=sizesigma-1;i++) {
299       temp=0;
300       for(k=1;k<=v;k++){temp += ((s[i][k])*(pat[mu][k+l]));}
301       m[mu][i]=(int)temp;
302     }
303   }}
304
305 /*Create matrix x[a][i],xh[a][j]*/
306
307 void vectorxxh(int **x, int size, int sizepatt){
308   int i,a;
309   for (a=1; a<=sizepatt; a++) {
310     for(i=-(size);i<=size;i+=2){
311       x[a][i]= i;
312     }
313   }}
```

Listing A.2: Population dynamics code

Bibliography

[1] Abbas, A. K., Lichtman, A. H. and Pillai, S. (2014). *Cellular and molecular immunology.* Elsevier Health Sciences.

[2] Janeway, C. A., Travers, P. and Walport, M. (2005). *Immunobiology: the immune system in health and disease*. New York: Garland Science.

[3] Burnet, S. F. M. (1959). *The clonal selection theory of acquired immunity (Vol. 3)*. Nashville: Vanderbilt University Press.

[4] Jerne, N. K. (1974). Towards a network theory of the immune system. *Annales d'immunologie*, **125(1-2)**, 373-389.

[5] Uner, A. and Gavalchin, J. (2006). Idiotypes. *eLS*.

[6] Bell, G. I. (1970). Mathematical model of clonal selection and antibody production. *Journal of Theoretical Biology*, **29(2)**, 191-232.

[7] Perelson, A. S. and Weisbuch, G. (1997). Immunology for physicists. *Reviews of Modern Physics*, **69(4)**, 1219.

[8] Perelson, A. S. and Oster, G. F. (1979). Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *Journal of Theoretical Biology*, **81(4)**, 645-670.

[9] Perelson, A. S., Mirmirani, M. and Oster, G. F. (1976). Optimal strategies in immunology. *Journal of Mathematical Biology*, **3(3-4)**, 325-367.

[10] Borghans, J. A., Noest, A. J. and De Boer, R. J. (1999). How specific should immunological memory be?. *The Journal of Immunology*, **163(2)**, 569-575.

[11] Elhanati, Y., Murugan, A., Callan, C. G., Mora, T. and Walczak, A. M. (2014). Quantifying selection in immune receptor repertoires. *Proceedings of the National Academy of Sciences*, **111(27)**, 9875-9880.

[12] Mora, T., Walczak, A. M., Bialek, W. and Callan, C. G. (2010). Maximum entropy models for antibody diversity. *Proceedings of the National Academy of Sciences*, **107(12)**, 5405-5410.

[13] Elhanati, Y., Sethna, Z., Marcou, Q., Callan, C. G., Mora, T. and Walczak, A. M. (2015). Inferring processes underlying B-cell repertoire diversity. *Phil. Trans. R. Soc. B*, **370(1676)**, 20140243.

[14] Perelson, A. S. (1989). Immune network theory. *Immunological Reviews*, **110(1)**, 5-36.

[15] Celli, S., Day, M., Mller, A. J., Molina-Paris, C., Lythe, G. and Bousso, P. (2012). How many dendritic cells are required to initiate a T-cell response?. *Blood*, **120(19)**, 3945-3948.

[16] Butler, T. C., Kardar, M. and Chakraborty, A. K. (2013). Quorum sensing allows T cells to discriminate between self and nonself. *Proceedings of the National Academy of Sciences*, **110(29)**, 11833-11838.

[17] van den Berg, H. A. and Rand, D. A. (2004). Dynamics of T cell activation threshold tuning. *Journal of Theoretical Biology*, **228(3)**, 397-416.

[18] Parisi, G. (1990). A simple model for the immune network. *Proceedings of the National Academy of Sciences*, **87(1)**, 429-433.

[19] Weisbuch, G., De Boer, R. J. and Perelson, A. S. (1990). Localized memories in idiotypic networks. *Journal of Theoretical Biology*, **146(4)**, 483-499.

[20] Barra, A. and Agliari, E. (2010). A statistical mechanics approach to autopoietic immune networks. *Journal of Statistical Mechanics: Theory and Experiment*, **2010(07)**, P07004.

[21] Barra, A. and Agliari, E. (2010). Stochastic dynamics for idiotypic immune networks. *Physica A: Statistical mechanics and its applications*, **389(24)**, 5903-5911.

[22] Agliari, E., Barra, A., Del Ferraro, G., Guerra, F. and Tantari, D. (2015). Anergy in self-directed B lymphocytes: a statistical mechanics perspective. *Journal of Theoretical Biology*, **375**, 21-31.

[23] Agliari, E., Barra, A., Guerra, F. and Moauro, F. (2011). A thermodynamic perspective of immune capabilities. *Journal of Theoretical Biology*, **287**, 48-63.

[24] Agliari, E., Barra, A., Bartolucci, S., Galluzzi, A., Guerra, F. and Moauro, F. (2013). Parallel processing in immune networks. *Physical Review E*, **87(4)**, 042701.

[25] Parisi, G., Mézard, M. and Virasoro, M. A. (1987). *Spin glass theory and beyond*. World Scientific, Singapore.

[26] Agliari, E., Annibale, A., Barra, A., Coolen, A. C. C. and Tantari, D. (2013). Immune networks: multi-tasking capabilities at medium load. *Journal of Physics A: Mathematical and Theoretical*, **46(33)**, 335101.

[27] Agliari, E., Annibale, A., Barra, A., Coolen, A. C. C. and Tantari, D. (2013). Immune networks: multitasking capabilities near saturation. *Journal of Physics A: Mathematical and Theoretical*, **46(41)**, 415003.

[28] Amit, D. J. (1992). *Modeling brain function: The world of attractor neural networks*. Cambridge University Press.

[29] Amit, D. J., Gutfreund, H. and Sompolinsky, H. (1985). Spin-glass models of neural networks. *Physical Review A*, **32(2)**, 1007.

[30] Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, **79(8)**, 2554-2558.

[31] Coolen, A. C., Kühn, R. and Sollich, P. (2005). *Theory of neural information processing systems*. OUP Oxford.

[32] Amit, D. J., Gutfreund, H. and Sompolinsky, H. (1985). Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters*, **55(14)**, 1530.

[33] Derrida, B., Gardner, E. and Zippelius, A. (1987). An exactly solvable asymmetric neural network model. *Europhysics Letters*, **4(2)**, 167.

[34] Watkin, T. L. H. and Sherrington, D. (1991). A neural network with low symmetric connectivity. *Europhysics Letters*, **14(8)**, 791.

[35] Wemmenhove, B. and Coolen, A. C. C. (2003). Finite connectivity attractor neural networks. *Journal of Physics A: Mathematical and General*, **36(37)**, 9617.

[36] Agliari, E., Barra, A., Galluzzi, A., Guerra, F. and Moauro, F. (2012). Multitasking associative networks. *Physical Review Letters*, **109(26)**, 268101.

[37] Hertz, J., Krogh, A. and Palmer, R. G. (1991). *Introduction to the theory of neural computation (Vol. 1)*. Basic Books.

[38] Barra, A., Bernacchia, A., Santucci, E. and Contucci, P. (2012). On the equivalence of Hopfield networks and Boltzmann Machines. *Neural Networks*, **34**, 1-9.

[39] Sollich, P., Tantari, D., Annibale, A. and Barra, A. (2014). Extensive parallel processing on scale-free networks. *Physical Review Letters*, **113(23)**, 238106.

[40] Agliari, E., Barra, A., Galluzzi, A. and Isopi, M. (2014). Multitasking attractor networks with neuronal threshold noise. *Neural Networks*, **49**, 19-29.

[41] Coolen, A. C. C. and Ruijgrok, T. W. (1988). Image evolution in Hopfield networks. *Physical Review A*, **38(8)**, 4253.

[42] Coolen, A. C. C., Laughton, S. N. and Sherrington, D. (1996). Dynamical replica theory for disordered spin systems. *Physical Review B*, **53(13)**, 8184.

[43] Sherrington, D., Coolen, A. C. C. and Laughton, S. N. (1996). Macrodynamics of Disordered and Frustrated Systems. *arXiv preprint* cond-mat/9606098.

[44] Van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry (Vol. 1)*. Elsevier.

[45] Abramowitz, M. and Stegun, I. A. (1966). *Handbook of mathematical functions.* Applied mathematics series, 55, 62.

[46] Dorogovtsev, S. N. and Mendes, J. F. (2013). *Evolution of networks: From biological nets to the Internet and WWW.* OUP Oxford.

[47] Kouskoff, V., Lacaud, G., Pape, K., Retter, M. and Nemazee, D. (2000). B cell receptor expression level determines the fate of developing B lymphocytes: receptor editing versus selection. *Proceedings of the National Academy of Sciences*, **97(13)**, 7435-7439.

[48] Raff, M. (1977). Immunological networks. *Nature*,**265**, 205 - 207.

[49] Menshikov, I., Beduleva, L., Frolov, M., Abisheva, N., Khramova, T., Stolyarova, E. and Fomina, K. (2015). The idiotypic network in the regulation

of autoimmunity: Theoretical and experimental studies. *Journal of Theoretical Biology*, **375**, 32-39.

[50] Pendergraft, W. F., Preston, G. A., Shah, R. R., Tropsha, A., Carter, C. W., Jennette, J. C. and Falk, R. J. (2004). Autoimmunity is triggered by cPR-3 (105201), a protein complementary to human autoantigen proteinase-3. *Nature Medicine*, **10(1)**, 72-79.

[51] Shoenfeld, Y. (2004). The idiotypic network in autoimmunity: antibodies that bind antibodies that bind antibodies. *Nature Medicine*, **10(1)**, 17-18.

[52] Menshikov, I. and Beduleva, L. (2008). Evidence in favor of a role of idiotypic network in autoimmune hemolytic anemia induction: theoretical and experimental studies. *International Immunology*, **20(2)**, 193-198.

[53] Plas, D. R., Rathmell, J. C. and Thompson, C. B. (2002). Homeostatic control of lymphocyte survival: potential origins and implications. *Nature Immunology*, **3(6)**, 515-521.

[54] Viola, A. and Lanzavecchia, A. (1996). T cell activation determined by T cell receptor number and tunable thresholds. *Science*, **273(5271)**, 104-106.

[55] Cosenza, H. and Köhler H. (1972). Specific inhibition of plaque formation to phosphorylcholine by antibody against antibody. *Science*, **176(4038)**, 1027-1029.

[56] Eichmann, K. and Rajewsky, K. (1975). Induction of T and B cell immunity by anti-idiotypic antibody. *European Journal of Immunology*, **5(10)**, 661-666.

[57] Urbain, J., Wikler, M., Franssen, J. D. and Collignon, C. (1977). Idiotypic regulation of the immune system by the induction of antibodies against anti-idiotypic antibodies. *Proceedings of the National Academy of Sciences*, **74(11)**, 5126-5130.

[58] Hoffmann, G. (1975). A theory of regulation and self-nonself discrimination in an immune network. *European Journal of Immunology*, **5(9)**, 638-647.

[59] Hiernaux, J. (1977). Some remarks on the stability of the idiotypic network. *Immunochemistry*, **14(11)**, 733-739.

[60] Geha, R. S. (1985). Idiotypic-anti-idiotypic interactions in man. *American Journal of Diseases of Children*, **139(4)**, 417-420.

[61] Bartolucci, S. and Annibale, A. (2015). A dynamical model of the adaptive immune system: effects of cells promiscuity, antigens and BB interactions. *Journal of Statistical Mechanics: Theory and Experiment*, **2015(8)**, P08017.

[62] Laughton, S. N. and Coolen, A. C. C. (1995). Macroscopic Lyapunov functions for separable stochastic neural networks with detailed balance. *Journal of Statistical Physics*, **80(1-2)**, 375-387.

[63] Kochmański, M., Paszkiewicz, T. and Wolski, S. (2013). Curie-Weiss magnet - a simple model of phase transition. *European Journal of Physics*, **34(6)**, 1555.

[64] De Monvel, J. H. B. and Martin, O. C. (1995). Memory capacity in large idiotypic networks. *Bulletin of Mathematical Biology*, **57(1)**, 109-136.

[65] Krapivsky, P. L., Redner, S. and Ben-Naim, E. (2010). *A kinetic view of statistical physics*. Cambridge University Press.

[66] Zouali, M. (2007). Immunological tolerance: mechanisms. *eLS*.

[67] Yaari, G. and Kleinstein, S. H. (2015). Practical guidelines for B-cell receptor repertoire sequencing analysis. *Genome Medicine*, **7(1)**, 1-14.

[68] Georgiou, G., Ippolito, G. C., Beausang, J., Busse, C. E., Wardemann, H. and Quake, S. R. (2014). The promise and challenge of high-throughput

sequencing of the antibody repertoire. *Nature Biotechnology*, **32(2)**, 158-168.

[69] Weinstein, J. A., Jiang, N., White, R. A., Fisher, D. S. and Quake, S. R. (2009). High-throughput sequencing of the zebrafish antibody repertoire. *Science*, **324(5928)**, 807-810.

[70] Desponds, J., Mora, T. and Walczak, A. M. (2016). Fluctuating fitness shapes the clone-size distribution of immune repertoires. *Proceedings of the National Academy of Sciences*, **113(2)**, 274-279.

[71] Mézard, M. and Montanari, A. (2009). *Information, physics, and computation*. Oxford University Press.

[72] Bartolucci, S. and Annibale, A. (2014). Associative networks with diluted patterns: dynamical analysis at low and medium load. *Journal of Physics A: Mathematical and Theoretical*, **47(41)**, 415001.

[73] Newman, M. (2010). *Networks: an introduction*. OUP Oxford.

[74] Mézard, M. and Parisi, G. (2001). The Bethe lattice spin glass revisited. *The European Physical Journal B -Condensed Matter and Complex Systems*, **20(2)**, 217-233.

[75] Moore, C. and Mertens, S. (2011). *The nature of computation*. OUP Oxford.

[76] Baxter, R. J. (1982). *Exactly solvable models in statistical mechanics*. Academic Press.

[77] Bialek, W. (2015). Perspectives on theory at the interface of physics and biology. *arXiv preprint* arXiv:1512.08954.

[78] Gao, L., Zhou, F., Li, X. and Jin, Q. (2010). HIV/TB co-infection in mainland China: a meta-analysis. *PloS one*, **5(5)**, e10736.

[79] Zonana-Nacach, A., Camargo-Coronel, A., Yanez, P., Sánchez, L., Jimenez-Balderas, F. J. and Fraga, A. (2001). Infections in outpatients with systemic lupus erythematosus: a prospective study. *Lupus*, **10(7)**, 505-510.

[80] Hershberg, U. and Prak, E. T. L. (2015). The analysis of clonal expansions in normal and autoimmune B cell repertoires. *Phil. Trans. R. Soc. B*, **370(1676)**, 20140239.

[81] Dunn-Walters, D. K. and Ademokun, A. A. (2010). B cell repertoire and ageing. *Current Opinion in Immunology*, **22(4)**, 514-520.

[82] Kurtz, D. M. et al. (2015). Noninvasive monitoring of diffuse large B-cell lymphoma by immunoglobulin high-throughput sequencing. *Blood*, **125(24)**, 3679-3687.

[83] Mozeika, A. and Coolen, A. C. (2016). Statistical mechanics of clonal expansion in lymphocyte networks modelled with slow and fast variables. *arXiv preprint* arXiv:1603.01328.