Immune networks: multitasking capabilities near saturation

2013 J. Phys. A: Math. Theor. 46 415003

(http://iopscience.iop.org/1751-8121/46/41/415003)

J. Phys. A: Math. Theor. **46** (2013) 415003 (48pp) doi:10.1088/1751-8113/46/41/415003

Immune networks: multitasking capabilities near saturation

E Agliari[1,2], A Annibale[3,4], A Barra[5], A C C Coolen[4,6] and D Tantari[7]

[1] Dipartimento di Fisica, Università degli Studi di Parma, Viale GP Usberti 7/A, I-43124 Parma, Italy
[2] INFN, Gruppo Collegato di Parma, Viale Parco Area delle Scienze 7/A, I-43100 Parma, Italy
[3] Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK
[4] Institute for Mathematical and Molecular Biomedicine, King's College London, Hodgkin Building, London SE1 1UL, UK
[5] Dipartimento di Fisica, Sapienza Università di Roma, Ple Aldo Moro 2, I-00185 Roma, Italy
[6] London Institute for Mathematical Sciences, 35a South St, Mayfair, London W1K 2XF, UK
[7] Dipartimento di Matematica, Sapienza Università di Roma, Ple Aldo Moro 2, I-00185 Roma, Italy

E-mail: agliari@fis.unipr.it, alessia.annibale@kcl.ac.uk, adriano.barra@roma1.infn.it, ton.coolen@kcl.ac.uk and tantari@mat.uniroma1.it

Received 25 May 2013, in final form 9 September 2013
Published 27 September 2013
Online at stacks.iop.org/JPhysA/46/415003

Abstract

Pattern-diluted associative networks were recently introduced as models for the immune system, with nodes representing T-lymphocytes and stored patterns representing signalling protocols between T- and B-lymphocytes. It was shown earlier that in the regime of extreme pattern dilution, a system with N_T T-lymphocytes can manage a number $N_B = \mathcal{O}(N_T^\delta)$ of B-lymphocytes simultaneously, with $\delta < 1$. Here we study this model in the extensive load regime $N_B = \alpha N_T$, with a high degree of pattern dilution, in agreement with immunological findings. We use graph theory and statistical mechanical analysis based on replica methods to show that in the finite-connectivity regime, where each T-lymphocyte interacts with a finite number of B-lymphocytes as $N_T \to \infty$, the T-lymphocytes can coordinate effective immune responses to an extensive number of distinct antigen invasions in parallel. As α increases, the system eventually undergoes a second order transition to a phase with clonal cross-talk interference, where the system's performance degrades gracefully. Mathematically, the model is equivalent to a spin system on a finitely connected graph with many short loops, so one would expect the available analytical methods, which all assume locally tree-like graphs, to fail. Yet it turns out to be solvable. Our results are supported by numerical simulations.

PACS numbers: 87.18.−h, 82.39.Rt, 75.10.Nr

(Some figures may appear in colour only in the online journal)

Contents

1. Introduction

After a long period of dormancy since the pioneering paper [1], we have in recent years seen a renewed interest in statistical mechanical models of the immune system [2–10]. These complement the standard approaches to immune system modelling, which are formulated in terms of dynamical systems [11–14]. However, to make further progress, we need quantitative tools that are able to handle the complexity of the immune system's intricate signalling patterns. Fortunately, over the past few decades a powerful arsenal of statistical mechanical techniques has been developed in the disordered system community to deal with heterogeneous many-variable systems on complex topologies [15–19]. In the present paper we exploit these new techniques to model the multitasking capabilities of the (adaptive) immune network, where effector branches (B-cells) and coordinator branches (T-cells) interact via (eliciting and suppressive) signalling proteins called cytokines. From a theoretical physics perspective, a network of interacting B- and T-cells resembles a bi-partite spin-glass. It was recently shown that such a bi-partite spin-glass is thermodynamically equivalent to a Hopfield-like neural network with effective Hebbian interactions [20, 21].

The analogy between immune and neural networks was already noted decades ago: both networks are able to learn (e.g. how to fight new antigens), memorize (e.g. previously encountered antigens) and 'think' (e.g. select the best strategy to cope with pathogens). However, their architectures are very different. Models with fully connected topology, mathematically convenient simplifications of biological reality, are tolerable for neural networks, where each neuron is known to have a huge number of connections with others [22]. In immune networks, however, interactions among lymphocytes are much more specific and signalled via chemical messengers, leading to network topologies that display finite connectivity. This difference is not purely formal, but also plays a crucial operational role. Neural networks are designed to perform high-resolution serial information processing, with neurons interacting with many others to collectively retrieve a *single* pattern at a time. The immune system, in contrast, must simultaneously recall multiple patterns (i.e. defense strategies), since many antigens will normally attack the host at the same time. Remarkably, diluting interactions in the underlying bi-partite spin-glass causes a switch from serial to parallel processing (i.e. to simultaneous pattern recall) of the thermodynamically equivalent Hopfield network[8] [23, 24].

The inextricable link between retrieval and the topological features of such systems requires a combination of techniques from statistical mechanics and graph theory, which will be the focus of the present paper. The paper is organized as follows. In section 2 we describe a minimal biological scenario for the immune system, based on the analogy with neural networks, and define our model. Section 3 gives a comprehensive analysis of the topological properties of the network in the finite-connectivity and high load regime, which is the one assumed throughout our paper. Section 4 is dedicated to the statistical mechanical analysis of the system, focusing on simultaneous pattern recall of the network. In section 5 we use a population dynamics algorithm to numerically inspect different regions of the phase diagram. We finish with a summary of our main findings.

2. Statistical mechanical modelling of the adaptive immune system

2.1. The underlying biology

All mammals have an innate (broad range) immunity, managed by macrophages, neutrophils, etc, and an adaptive immune response. We refer the reader to the excellent books [29, 30] for comprehensive reviews of the immune system, and to a selection of papers [2–4, 23, 24, 31] for theoretical modelling inspired by biological reality. Our prime interest is in B-cells and in T-cells; in particular, among T-cells, in the subgroups of so-called helpers and suppressors. B-cells produce antibodies which are able to recognize and bind pathogens, and those that produce the same antibody are said to form a clone. The human immune repertoire consists of $\mathcal{O}(10^8$–$10^9)$ clones. The size of a clone, i.e. the number of identical B-cells, may vary strongly. A clone at rest may contain some $\mathcal{O}(10^3$–$10^4)$ cells, but when it undergoes clonal expansion its size may increase by several orders of magnitude, up to $\mathcal{O}(10^6$–$10^7)$. Beyond this size the state of the immune system would be pathological, and is referred to as lymphocytosis.

When an antigen enters the body, several antibodies produced by different clones may be able to bind to it, making it chemically inert and biologically inoffensive. In this case, conditional on authorization by T-helpers (mediated via cytokines), the binding clones undergo clonal expansion and start releasing high quantities of soluble antibodies to inhibit the enemy. After the antigen has been deleted, B-cells are instructed by T-suppressors, again

[8] In contrast, diluting the bonds in a Hopfield network does not affect pattern retrieval qualitatively [17, 18, 25–28]: the system would still recall only one pattern at a time, but simply have a lower storage capacity.

via cytokines, to stop producing antibodies and undergo apoptosis. In this way the clones reduce their sizes, and order is restored. Thus, two signals are required within a small time interval for B-cells to start clonal expansion: the first is binding to antigen, the second is a 'consensus' signal, in the form of an eliciting cytokine [32, 33] secreted by T-helpers. This mechanism, known as the 'two-signal model' [34–37], prevents abnormal reactions, such as autoimmune manifestations[9]. The focus of this study is to understand, from a statistical mechanics perspective, the ability of helpers and suppressors to coordinate and manage an extensive ensemble of B-clones *simultaneously*.

2.2. A minimal model

We consider an immune repertoire of N_B different clones, labelled $\mu \in \{1, \dots, N_B\}$. The size of clone μ is b_μ. In the absence of interactions with helpers, we take the clone sizes to be Gaussian distributed; without loss of generality we may take the mean to be zero and unit width, so $b_\mu \sim \mathcal{N}(0, 1)$. A value $b_\mu \gg 1$ now implies that clone μ has expanded (relative to the typical clonal size), while $b_\mu \ll 1$ implies inhibition. The Gaussian clone size distribution is supported both by experiments and by theoretical arguments [4]. Similarly, we imagine having N_T helper cells, labelled $i \in \{1, \dots, N_T\}$. The state of helper cell i is denoted by σ_i. For simplicity, helpers are assumed to be in only two possible states: secreting cytokines ($\sigma_i = +1$) or quiescent ($\sigma_i = -1$). Both the clone sizes b_μ and the helper states σ_i are dynamical variables. We will abbreviate $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{N_T}) \in \{-1, 1\}^{N_T}$, and $\mathbf{b} = (b_1, \dots, b_{N_B}) \in \mathbb{R}^{N_B}$.

The interaction between the helpers and the B-clones is implemented by cytokines. These are taken to be frozen (quenched) discrete variables. The effect of a cytokine secreted by helper i and detected by clone μ can be nonexistent ($\xi_i^\mu = 0$), excitatory ($\xi_i^\mu = 1$), or inhibitory ($\xi_i^\mu = -1$). To achieve a Hamiltonian formulation of the system, and thereby enable equilibrium statistical mechanical analysis, we have to impose symmetry of the cytokine interactions. So, in addition to the B-clones being influenced by cytokine signals from helpers, the helpers will similarly feel a signal from the B-clones. This symmetry assumption can be viewed as a necessary first step, to be relaxed in future investigations, similar in spirit to the early formulation of symmetric spin-glass models for neural networks [41, 42]. We are then led to a Hamiltonian $\hat{\mathcal{H}}(\mathbf{b}, \boldsymbol{\sigma}|\xi)$ for the combined system of the following form (modulo trivial multiplicative factors):

$$\hat{\mathcal{H}}(\mathbf{b}, \boldsymbol{\sigma}|\xi) = -\sum_{i=1}^{N_T}\sum_{\mu=1}^{N_B} \xi_i^\mu \sigma_i b_\mu + \frac{1}{2\sqrt{\beta}} \sum_{\mu=1}^{N_B} b_\mu^2. \tag{1}$$

In the language of disordered systems, this is a bi-partite spin-glass. We can integrate out the variables b_μ, and map our system to a model with helper–helper interactions only. The partition function $Z_{N_T}(\beta, \xi)$, at inverse clone size noise level $\sqrt{\beta}$ (which is the level consistent with our assumption $b_\mu \sim \mathcal{N}(0, 1)$) follows straightforwardly, and reveals the mathematical equivalence with an associative attractor network:

$$\begin{aligned} Z_{N_T}(\beta, \xi) &= \sum_{\boldsymbol{\sigma}} \int \mathrm{d}b_1 \dots \mathrm{d}b_{N_B} \exp[-\sqrt{\beta}\hat{\mathcal{H}}(\mathbf{b}, \boldsymbol{\sigma}|\xi)] \\ &= \sum_{\boldsymbol{\sigma}} \exp[-\beta \mathcal{H}(\boldsymbol{\sigma}|\xi)], \end{aligned} \tag{2}$$

[9] Through a phenomenon called 'cross-linking', a B-cell can also have the ability to bind a self-peptide, and may accidentally start duplication and antibody release, which is a dangerous unwanted outcome.

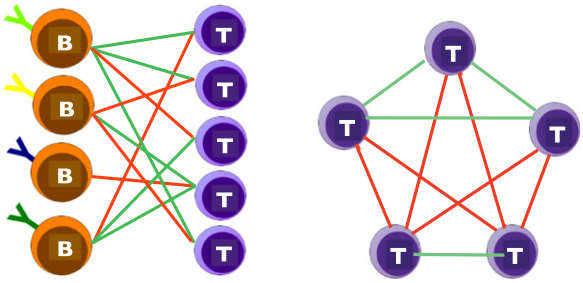

Figure 1. Left: the bi-partite spin-glass which models the interaction between B- and T-cells through cytokines. Green (red) links represent stimulatory (inhibitory) cytokines. Note that the network is diluted. Right: the equivalent associative multitasking network consisting of T-cells only, obtained by integrating out the B-cells. This network is also diluted, with links given by the Hebbian prescription.

in which, apart from an irrelevant additive constant,

$$\mathcal{H}(\boldsymbol{\sigma}|\xi) = -\frac{1}{2}\sum_{ij=1}^{N_T}\sigma_i J_{ij}\sigma_j, \qquad J_{ij} = \sum_{\mu=1}^{N_B}\xi_i^\mu \xi_j^\mu. \tag{3}$$

Thus, the system with Hamiltonian $\hat{\mathcal{H}}(\mathbf{b}, \boldsymbol{\sigma}|\xi)$, where helpers and B-clones interact stochastically through cytokines, is thermodynamically equivalent to a Hopfield-type associative network represented by $\mathcal{H}(\boldsymbol{\sigma}|\xi)$, in which helpers mutually interact through an effective Hebbian coupling (see figure 1). Learning a pattern in this model means adding a new B-clone with an associated string of new cytokine variables.

If all $\{\xi_i^\mu\}$ are nonzero, the system characterized by (3) is well-known in the information processing systems community. It is able to retrieve each of the N_B 'patterns' $(\xi_1^\mu, \ldots, \xi_{N_T}^\mu)$, provided these are sufficiently uncorrelated, and both the ratio $\alpha = N_B/N_T$ and the noise level $1/\beta$ are sufficiently small [4, 26, 43, 46]. Retrieval quality can be quantified by introducing N_B suitable order parameters, the so-called Mattis magnetizations $m_\mu(\boldsymbol{\sigma}) = N_T^{-1}\sum_i \xi_i^\mu \sigma_i$, in terms of which we can write (3) as

$$\mathcal{H}(\boldsymbol{\sigma}|\xi) = -\frac{N_T}{2}\sum_{\mu=1}^{N_B} m_\mu^2(\boldsymbol{\sigma}). \tag{4}$$

If α is sufficiently small, the minimum energy configurations of the system are those where $m_\mu(\boldsymbol{\sigma}) = 1$ for some μ ('pure states'), which implies that $\boldsymbol{\sigma} = (\xi_1^\mu, \ldots, \xi_{N_T}^\mu)$ and pattern μ is said to be retrieved perfectly. In our immunological context this means the following. If $m_\mu(\boldsymbol{\sigma}) = 1$, all the helpers are 'aligned' with their coupled cytokines: those i that inhibit clone μ (i.e. secrete $\xi_i^\mu = -1$) will be quiescent ($\sigma_i = -1$), and those i that excite clone μ (i.e. secrete $\xi_i^\mu = 1$) will be active ($\sigma_i = 1$) and release the eliciting cytokine. As a result the B-clone μ receives the strongest possible positive signal (i.e. the random environment becomes a 'staggered magnetic field'), hence it is forced to expand (see figure 2). Conversely, for $m_\mu(\boldsymbol{\sigma}) = -1$, clone μ receives the strongest possible negative signal and it is forced to contract. However, in this scenario of $\xi_i^\mu \in \{-1, 1\}$ for all (i, μ) (where the bi-partite network is fully connected) only one B-clone at a time can expand (apart from minor spurious mixture states). This would be a disaster for the immune system.

We need the dilution in the bi-partite B–H network that is caused by also having $\xi_i^\mu = 0$ (i.e. no signalling between helper i and clone μ), to enable multiple clonal expansions. In this

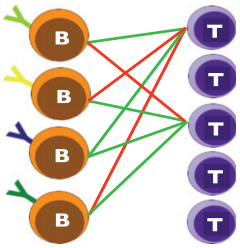

Figure 2. The specific T-cell configuration that would give the strongest possible positive signal to the first clone. Upward arrows indicate cytokine-secreting T-cells; downward arrows indicate quiescent ones. Eliciting and suppressive cytokines are represented by green and red links, respectively.

case, the network (3) stores patterns that also have blank entries, and retrieving a pattern no longer employs all spins σ_i: those corresponding to null entries can be used to recall other patterns. This is energetically favourable since the energy is quadratic in the magnetizations $m_\mu(\boldsymbol{\sigma})$. Conceptually, this is only a redefinition of the network's recall task: no theoretical bound for information content is violated, and global retrieval is still performed through N_B bits. However, the perspective is shifted: the system no longer requires a sharp resolution in information exchange between a helper clone and a B-clone.[10] It suffices that a B-clone receives an attack signal, which could be encoded even by a single bit. In a diluted bi-partite B–H system the associative capabilities of the helper network are distributed, in order to simultaneously manage the whole ensemble of B-cells. The analysis of these immunologically relevant pattern-diluted versions of associative networks has so far been carried out in the low storage case $N_B \sim \log N_T$ [23, 24] and the medium storage case $N_B \sim N_T^\delta$, $0 < \delta < 1$, where the system indeed performs as a multitasking associative memory [38]. The focus of this paper is to analyse the ability of the network to simultaneously retrieve an extensive number of patterns, i.e. $N_B = \alpha N_T$ with $\alpha > 0$ fixed and $N_T \to \infty$, while in addition implementing a higher degree of dilution for the B–H system, in agreement with immunological findings [29, 30].

3. Topological properties of the emergent network

3.1. Definitions

The system composed of N_T T-lymphocytes, that interact with N_B B-lymphocytes via cytokines, can be described as a bi-partite graph $\mathcal{B}$, in which the nodes, belonging to the sets $\mathcal{V}_T$ and $\mathcal{V}_B$, of cardinality $|\mathcal{V}_T| = N_T$ and $|\mathcal{V}_B| = N_B$, respectively, are pairwise connected via undirected links. We assign to the link between T-lymphocyte i and B-lymphocyte μ a variable ξ_i^μ, which takes values 1 if the cytokines produced by T-lymphocyte i triggers expansion of B-clone μ, -1 if it triggers contraction and 0 if i and μ do not interact. We assume that the $\{\xi_i^\mu\}$ are identically and independently distributed random variables, drawn from

$$P(\xi_i^\mu|d) = \frac{1-d}{2}\delta_{\xi_i^\mu-1,0} + \frac{1-d}{2}\delta_{\xi_i^\mu+1,0} + d\,\delta_{\xi_i^\mu,0}, \tag{5}$$

where $\delta_{x,0}$ is the Kronecker delta symbol. $P(\xi_i^\mu|d)$ implicitly accounts for bond dilution within the graph $\mathcal{B}$.

[10] In fact, the high-resolution analysis is performed in the antigenic recognition on the B-cell surface, which is based on a sharp key-and-lock mechanism [2].

It is experimentally well established that although helpers are much more numerous than B-cells, their relative sizes are still comparable in a statistical mechanical sense, hence we will assume

$$N_B = \alpha N_T, \quad 0 < \alpha < 1. \tag{6}$$

We have shown in the previous section how the signalling process of the B- and T-cells on this bi-partite graph can be mapped to a thermodynamically equivalent process on a new graph, $\mathcal{G}$, built only of the N_T nodes in $\mathcal{V}_T$, and occupied by spins σ_i that interact pairwise through the coupling matrix

$$J_{ij} = \sum_{\mu=1}^{N_B} \xi_i^\mu \xi_j^\mu. \tag{7}$$

The topology of the (weighted, monopartite) graph $\mathcal{G}$ can range from fully connected to sparse, as d is tuned [47, 48]. Our interest is in the ability of this system to perform as a multitasking associative memory such that the maximum number of pathogens can be fought simultaneously. A recent study [38] suggested that in order to bypass the spin-glass structure of phase space at the load level (6), a finite-connectivity topology is required:

$$d = 1 - c/N_T. \tag{8}$$

Remarkably, the finite-connectivity topology is also in agreement with the biological picture of highly-selective touch-interactions among B- and T-cells.

3.2. Simple characteristics of graph $\mathcal{B}$

Let us now describe in more detail the topology of the graph $\mathcal{B}$ under condition (8). We denote with k_i the degree of node $i \in \mathcal{V}_T$ (the number of links stemming from i), and with κ_μ the degree of node $\mu \in \mathcal{V}_B$ (the number of links stemming from μ):

$$k_i = \sum_{\mu \in \mathcal{V}_B} |\xi_i^\mu| \in [0, N_B], \qquad \kappa_\mu = \sum_{i \in \mathcal{V}_T} |\xi_i^\mu| \in [0, N_T]. \tag{9}$$

Since links in $\mathcal{B}$ are independently and identically drawn, k and κ both follow a binomial distribution

$$P_T(k|d, N_B) = \binom{N_B}{k} (1-d)^k d^{N_B - k}, \qquad P_B(\kappa|d, N_T) = \binom{N_T}{\kappa} (1-d)^\kappa d^{N_T - \kappa}, \tag{10}$$

hence we have $\mathbb{E}_T(k) \equiv \sum_k P_T(k|d, N_B)k = (1-d)N_B = c\alpha$, $\mathbb{E}_T(k^2) - [\mathbb{E}_T(k)]^2 = (1-d)dN_B = c\alpha(1 - c/N_T)$, and $\mathbb{E}_B(\kappa) \equiv \sum_\kappa P_B(\kappa|d, N_T)\kappa = (1-d)N_T = c$, $\mathbb{E}_B(\kappa^2) - [\mathbb{E}_B(\kappa)]^2 = (1-d)dN_T = c(1 - c/N_T)$.

Due to the finite connectivity, we expect an extensive fraction of nodes $i \in \mathcal{V}_T$ and $\mu \in \mathcal{V}_B$ to be isolated. In the thermodynamic limit, the fraction of isolated nodes in $\mathcal{V}_T$ and $\mathcal{V}_B$ is $d^{N_B} = (1 - c/N_T)^{\alpha N_T} \to \mathrm{e}^{-c\alpha}$ and $d^{N_T} = (1 - c/N_T)^{N_T} \to \mathrm{e}^{-c}$, respectively. In order to have a low number of non-signalled B-cells, one should therefore choose a relatively large value of c. Moreover, as will be shown below, by reducing α and/or c we can break $\mathcal{B}$ into small components, each yielding, upon marginalization (2), a distinct component within $\mathcal{G}$ (see figure 3). This fragmentation is crucial to allow parallel pattern retrieval. In general, a macroscopic component emerges when the link probability $(1-d)$ exceeds the percolation threshold $1/\sqrt{N_T N_B}$ [38, 45], which, recalling equations (6) and (8), can be translated into

$$c > 1/\sqrt{\alpha}. \tag{11}$$

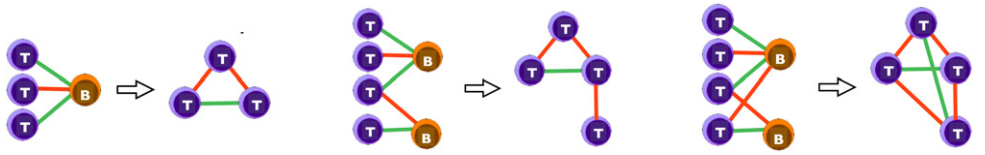

Figure 3. Examples showing how different components within $\mathcal{B}$ are mapped into $\mathcal{G}$ upon marginalization. Left: any star S_n in $\mathcal{B}$ with a node in $\mathcal{V}_B$ at its centre and n leaves $i_1, i_2, \ldots, i_n \in \mathcal{V}_T$ corresponds in $\mathcal{G}$ to a complete graph K_n, where each link J_{ij} has unit magnitude. Middle: two stars S_n and S_m in $\mathcal{B}$ that share a leaf correspond to two connected complete graphs in $\mathcal{G}$, K_n and K_m respectively, that have a common node. Again, each link J_{ij} has unit weight. Right: when a loop of length 4 is present in the bi-partite graph $\mathcal{B}$, the corresponding nodes in $\mathcal{G}$ may be connected by a link with weight larger (in modulus) than 1.

3.3. Analysis of graph $\mathcal{G}$

Due to the finite connectivity of $\mathcal{B}$, we expect also that $\mathcal{G}$ will have a macroscopic fraction $f(\alpha, c)$ of isolated nodes, which will be larger than $\mathrm{e}^{-c\alpha}$ (the fraction of nodes that are isolated in $\mathcal{B}$). In fact, a node $i \in \mathcal{V}_T$, which in $\mathcal{B}$ has a number of neighbours $\mu \in \mathcal{V}_B$, but does not share any of these with any other node $j \in \mathcal{V}_T$, remains isolated upon marginalization. Put differently, whenever i is the centre of a star S_n in $\mathcal{B}$, with $n = 0, 1, \ldots, N_B$, it will be isolated in $\mathcal{G}$. We recall that a star S_n is a tree with a central node and n leaves; this includes isolated nodes ($n = 0$), dimers ($n = 1$), etc. The larger n, the less likely the occurrence of the component S_n in $\mathcal{B}$ and the smaller the related contribution $f_n(\alpha, c)$ to $f(\alpha, c)$. On average, one will have

$$f_n(\alpha, c) = (1-d)^n d^{N_B - n} \binom{N_B}{n}, \tag{12}$$

so that, overall, the fraction of isolated nodes in $\mathcal{G}$ is roughly $\mathrm{e}^{-c\alpha} + c\alpha\, \mathrm{e}^{-c\alpha}$. In the following subsections we will inspect the architecture of $\mathcal{G}$ in more detail.

3.3.1. Coupling distribution. Let us introduce the probability distribution $P(J|N_B, N_T, c)$ that an arbitrary link J_{ij}, as given by (7), has weight J. The average link probability is then $1 - P(J = 0|N_B, N_T, c) \equiv 1 - p$. This distribution $P(J|N_B, N_T, c)$ can be viewed as the probability distribution for the end-to-end distance of a one-dimensional random walk endowed with a waiting probability p_w, which corresponds to the probability that a term $\xi_i^\mu \xi_j^\mu$ is null, and equal probabilities $p_l = p_r$ of moving left or right, respectively:

$$p_w = d(2-d) = 1 - \left(\frac{c}{N_T}\right)^2, \tag{13}$$

$$p_l = p_r = \frac{1-p_w}{2} = \frac{1}{2}(1-d)^2 = \frac{1}{2}\left(\frac{c}{N_T}\right)^2. \tag{14}$$

Therefore, we can write

$$P(J|N_B, N_T, c) = \sum_{S=0}^{N_B - J}{}' \frac{N_B!}{S!\left(\frac{N_B-S-J}{2}\right)!\left(\frac{N_B-S+J}{2}\right)!} p_w^S p_r^{(N_B-S+J)/2} p_l^{(N_B-S-J)/2}, \tag{15}$$

where the primed sum means that only values of S with the same parity as $(N_B \pm J)$ are taken into account. The distribution (15) can easily be generalized to the case of a biased random walk, i.e. biased distribution for weights [44]. The couplings among links have (in the limit of large N_T) the following average values [38]

$$\langle J \rangle = 0 \tag{16}$$

$$\langle J^2\rangle = \alpha c^2/N_T, \tag{17}$$

and for $J = 0$ one has

$$P(0|N_B, N_T, c) = \sum_{S=0}^{N_B}{}' \frac{N_B!}{S!\left[\left(\frac{N_B-S}{2}\right)\right]^2} p_w^S\, p_r^{(N_B-S)} = \sum_{S=0}^{N_B}{}' \binom{N_B}{S}\binom{N_B-S}{\frac{N_B-S}{2}} p_w^S p_r^{N_B-S}, \tag{18}$$

where now S must have the same parity as N_B. Assuming N_B even, we can write

$$P(0|N_B, N_T, c) = p_r^{N_B}\binom{N_B}{N_B/2}\, {}_2F_1\left(-\frac{N_B}{2}, -\frac{N_B}{2}, \frac{1}{2}, \frac{p_w^2}{4p_r^2}\right) \tag{19}$$

$$\approx \left(\frac{c\alpha}{N_B}\right)^{2N_B}\sqrt{\frac{2}{\pi N_B}}\, {}_2F_1\left(-\frac{N_B}{2}, -\frac{N_B}{2}, \frac{1}{2}, \left[\left(\frac{N_B}{c\alpha}\right)^2 - 1\right]^2\right), \tag{20}$$

where in the last step we used $N_B \gg 1$. Hence, upon expanding the hypergeometric function we get

$$P(0|N_B, N_T, c) \approx \left[1 - \left(\frac{c}{N_T}\right)^2\right]^{\alpha N_T}\left[1 + \mathcal{O}\left(N_T^{-2}\right)\right] \sim \mathrm{e}^{-c^2\alpha/N_T}. \tag{21}$$

Following a mean-field approach, we can estimate the degree distribution $P_{\mathrm{MF}}(z|N_B, N_T, c)$ in $\mathcal{G}$ by means of a binomial, in which the link probability is simply $p \equiv 1 - P(J = 0|N_B, N_T, c)$, namely

$$P_{\mathrm{MF}}(z|N_B, N_T, c) = \binom{N_T}{z}(1-p)^z p^{N_T - z}, \tag{22}$$

in which the average degree and the variance are

$$\langle z\rangle_{\mathrm{MF}} = (1-p)N_T \sim (1 - \mathrm{e}^{-c^2\alpha/N_T})N_T \sim c^2\alpha, \tag{23}$$

$$\langle z^2\rangle_{\mathrm{MF}} - \langle z\rangle_{\mathrm{MF}}^2 = (1-p)pN_T \sim c^2\alpha. \tag{24}$$

Due to the homogeneity assumption intrinsic to the mean-field approach, we expect our estimate to be accurate only for the first moment, while fluctuations are underestimated [38]. In order to account for the topological inhomogeneity characteristic of $\mathcal{G}$ we need to return to analysis of $\mathcal{B}$.

In the bi-partite graph, given two nodes $i, j \in \mathcal{V}_T$, with k_i and k_j the nearest neighbours respectively, the number ℓ of shared nearest neighbours corresponds to the number of non-null matchings between the related strings, $(\xi_i^1, \ldots, \xi_i^{N_B})$ and $(\xi_j^1, \ldots, \xi_j^{N_B})$, which is distributed according to

$$P(\ell|k_i, k_j, N_B) = \frac{N_B!}{(N_B + \ell - k_i - k_j)!(k_i - \ell)!(k_j - \ell)!\ell!}\left[\binom{N_B}{k_i}\binom{N_B}{k_j}\right]^{-1}. \tag{25}$$

Note that the number ℓ also provides an upper bound for J_{ij}. From (25) the average $\langle\ell\rangle_{k_i,k_j}$ is found to be

$$\langle\ell\rangle_{k_i,k_j} = k_i k_j/N_B. \tag{26}$$

We evaluate the typical environment for node i by averaging $P(\ell|k_i, k_j, N_B)$ over $P_T(k_j, c, N_B)$, as given by (10), and get

$$P(\ell|k_i, c, N_B) = \sum_{k_j=0}^{N_B}\binom{N_B}{k_j}\left(\frac{c}{N_T}\right)^{k_j}\left(1 - \frac{c}{N_T}\right)^{N_B-k_j} P(\ell|k_i, k_j, N_B) = \frac{d^{k_i-\ell}(1-d)^\ell k_i!}{\ell!(k_i-\ell)!}. \tag{27}$$

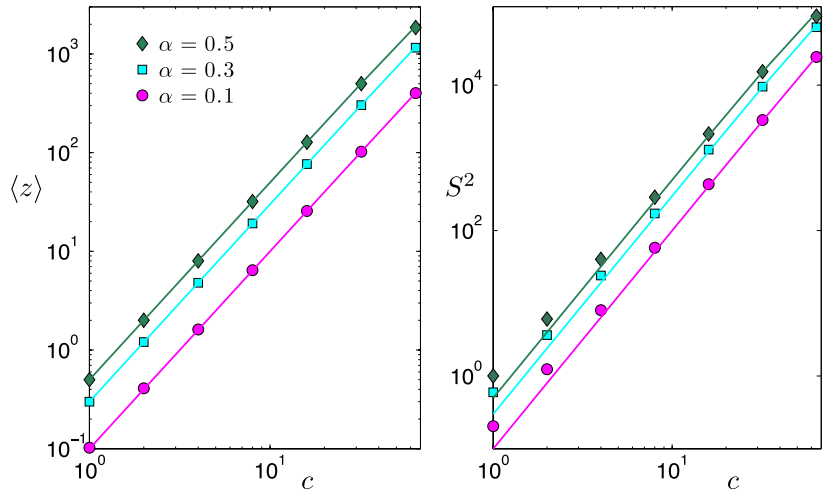

Figure 4. Average degree $\langle z \rangle$ (left) and its fluctuations $S^2 = \langle z^2 \rangle - \langle z^2 \rangle$ (right) in the bi-partite graph $\mathcal{B}$, as a function of c and for different values of α as shown in the legend. Data from numerical simulations (symbols) performed on systems of fixed size $N_T = 1.5 \times 10^3$ are compared with the analytical predictions (solid lines) given by (29) and (30).

In particular, the probability for i to be connected to an arbitrary node j can be estimated as $p_{k_i} \equiv 1 - P(\ell = 0|k_i, c, N_B)$, with which the average degree of node i can be written as

$$\langle z \rangle_{k_i} = p_{k_i} N_T = (1 - d^{k_i}) N_T. \tag{28}$$

Upon averaging $\langle z \rangle_{k_i}$ and $\langle z \rangle_{k_i}^2$ over $P_T(k_i|c, N_B)$, we get estimates for the average degree and its variance:

$$\langle z \rangle = \{1 - [d(2-d)]^{N_B}\} N_T \sim (1 - \mathrm{e}^{\alpha c^2/N_T}) N_T \sim \alpha c^2, \tag{29}$$

$$\langle z^2 \rangle - \langle z \rangle^2 = \{1 - 2[d(2-d)]^{N_B} + d^{N_B}(1 + d - d^2)^{N_B} - [1 - (d(2-d))^{N_B}]^2\} N_T^2 \sim \alpha c^3, \tag{30}$$

where the last approximation holds when c/N_T is small. As expected, we indeed recover the average degree predicted by the mean-field approach (23), while the fluctuations display an additional factor c (see (24)). The analytical results (29) and (30) are compared with numerical data in figure 4. The agreement is very good, especially for large c where the number of bonds is larger and hence the statistics are more sound.

3.3.2. Growth and robustness. As anticipated, the point where $\alpha c^2 = 1$ defines the percolation threshold for the bi-partite graph $\mathcal{B}$: when $c < 1/\sqrt{\alpha}$, the graph is fragmented into a number of components with sub-extensive size, while for $c > 1/\sqrt{\alpha}$ a giant (i.e. extensive) component emerges. This phenomenology is mirrored in the monopartite graph $\mathcal{G}$. In particular, we will show that for $c < 1/\sqrt{\alpha}$ there is a large number of disconnected components in $\mathcal{G}$ with finite size and a high degree of modularity, while for $c > 1/\sqrt{\alpha}$ bridges appear between these components, modularity progressively decays, and again a giant component emerges (see figure 5). The transition across the percolation threshold is rather smooth, as it stems from a main component which encompasses, as αc^2 is increased, more and more isolated nodes and small-sized components. This contrasts sharply with the situation in explosive percolation

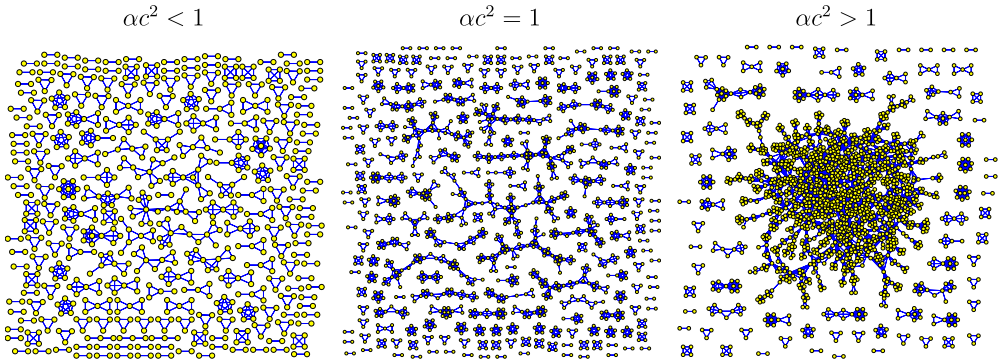

Figure 5. Examples of typical graphs $\mathcal{G}$ obtained for different values of c, while $\gamma = 1, \delta = 1$, $N_T = 5.10^3$ and $\alpha = 0.1$ are kept fixed. Left: the under-percolated regime. Middle: the percolation threshold. Right: the over-percolated regime. Isolated nodes, amounting to 4229, 3664 and 3243, respectively, are not shown here. As expected, although many short loops are already present for low connectivity, non-trivial (longer) loops start to occur at the percolation threshold $\alpha c^2 = 1$.

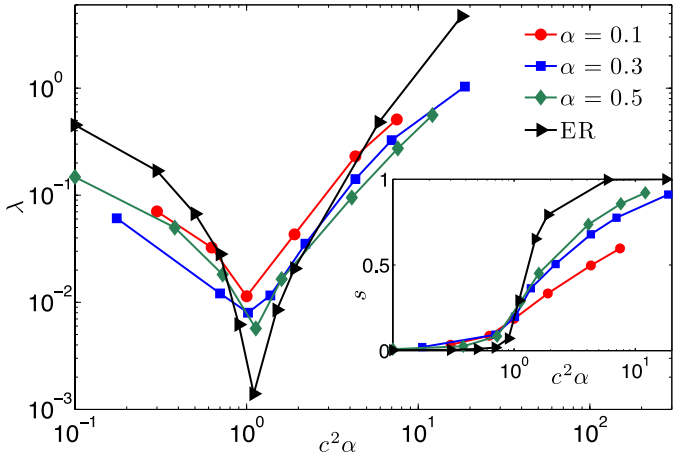

Figure 6. Main plot: algebraic connectivity λ versus $c^2\alpha$, measured on the largest components of graphs of size $N_T = 5000$, with different values of α (see legend). Similar results for ER graphs are shown for comparison; here λ is plotted versus the link probability times N_T, which represents the mean coordination number over the whole network. Inset: size s of the largest component versus αc^2.

processes [39], where a number of components develop and their merging at the percolation threshold gives rise to a steep growth in the size s of the largest component. Here s grows smoothly and, even at relatively large values of $c^2\alpha$, a significant fraction of nodes remain isolated or form small-size components (see figure 6, inset).

Moreover, the largest component exhibits high levels of modularity and clustering (see [38] for more details). This can be understood. For $\alpha < 1$, any set $\mathcal{C}_\mu$ such that all nodes $i \in \mathcal{C}_\mu \subseteq \mathcal{V}_T$ share at least one neighbour $\mu \in \mathcal{V}_B$ will, upon marginalization, result in a clique in $\mathcal{G}$. Hence, $\mathcal{G}$ is relatively compact and redundant and, due to its smooth growth, will remain so even around $\alpha c^2 = 1$. One can check this by measuring the algebraic connectivity, i.e. the

spectral gap λ of the largest component; results are shown in figure 6 (main plot)[11]. A graph with a small λ has a relatively clean bisection, while large λ values characterize non-structured networks, in which a simple clear-cut separation into subgraphs is not possible. As shown in figure 6, the minimum of λ provides a consistent signature of percolation, since the possible coalescence of different components is likely to yield the formation of bridges. Moreover, by comparing data for $\mathcal{G}$ and for an Erdös–Rényi (ER) graph we see that when $\alpha c^2 \approx 1$, where the related largest-size components are comparable, the former displays a larger λ and is hence more structured.

3.3.3. Component size distribution and retrieval. We next analyse the structure of the small components in $\mathcal{G}$, as they are strongly related to retrieval properties, starting with the *under-percolated regime*. Here the typical components in $\mathcal{B}$ are stars S_n centred in a node $\mu \in \mathcal{V}_B$ (because $N_T > N_B$), possibly with arms of length larger than 1, or a combination of stars. In all these cases, two nodes $i, j \in \mathcal{V}_T$ share at most one neighbour $\mu \in \mathcal{V}_B$, so the spins σ_i and σ_j can communicate non-conflicting signals to μ. More precisely, such components allow for spin configurations with nonzero Mattis magnetizations for *all* the patterns involved in the component (see figure 3). This scenario is therefore compatible with parallel retrieval.

Parallel retrieval can be jeopardized by loops in $\mathcal{B}$, which can create disruptive feedback mechanisms between spins, which prevent the complete and simultaneous retrieval of all patterns within the component (see the image on the right in figure 3). We can estimate the probability that a loop involving two nodes $i, j \in \mathcal{V}_T$ occurs in $\mathcal{B}$: since the graph is bipartite, the minimum length for loops is 4, which requires that i and j share a number $\ell \geqslant 2$ of neighbours in $\mathcal{B}$. We can write $P(\ell < 2|k_i, k_j, N_B) = P(\ell = 0|k_i, k_j, N_B) + P(\ell = 1|k_i, k_j, N_B)$. By replacing $\ell = 0$ and $\ell = 1$ in (25), we get, respectively,

$$P(\ell = 0|k_i, k_j, N_B) = \binom{N_B - k_i}{k_j} \binom{N_B}{k_j}^{-1}, \tag{31}$$

$$P(\ell = 1|k_i, k_j, N_B) = k_j \binom{N_B - k_i}{k_j - 1} \binom{N_B}{k_j}^{-1}. \tag{32}$$

By averaging over the distribution $P(k_j|d, N_B)$ (10) of k_j, we obtain the typical behaviour for an arbitrary node i

$$P(\ell < 2|k_i, d, N_B) = d^{k_i} + d^{k_i - 1}(1 - d)k_i = \left(1 - \frac{c}{N_T}\right)^{k_i} \left(1 + \frac{c}{N_T} k_i\right) \sim 1 - \left(\frac{ck_i}{N_T}\right)^2, \tag{33}$$

where in the last step we used $k_i \ll N_T$. In particular, when c is relatively large ($c\alpha > 1$), the approximation $k_i \approx \langle k \rangle$ is valid, and we see that the number of node pairs sharing at least two neighbours in $\mathcal{B}$ scales as $(\alpha c^2)^2$. Hence, in the under-percolated regime $\alpha c^2 < 1$, the graph $\mathcal{B}$ is devoid of loops, which is a necessary condition for straightforward error-free parallel retrieval. As mentioned before, in this regime the typical components in $\mathcal{B}$ are stars S_n centred in a node $\mu \in \mathcal{V}_B$, possibly presenting arms of length larger than 1, or a combination of stars. Upon marginalization, these arrangements give rise to complete graphs K_n, with nodes possibly linked to small trees, or combinations of complete graphs, respectively (see figure 5,

[11] The algebraic connectivity λ, or the 'spectral gap', i.e. the second smallest eigenvalue of the Laplacian matrix of a graph, is regarded as a useful quantifier in the analysis of various robustness-related problems. For instance, λ is a lower bound on both the node and the link connectivity. More precisely, a small algebraic connectivity means that it is relatively easy to disconnect the graph, i.e. to cut it into independent components. This means that there exist 'bottle-necks', i.e. one can identify subgraphs that are connected only via a small number of 'bridges'. A small algebraic connectivity is also known to influence transport processes on the graph itself and to favour instability of synchronized states (synchronizability) [40].

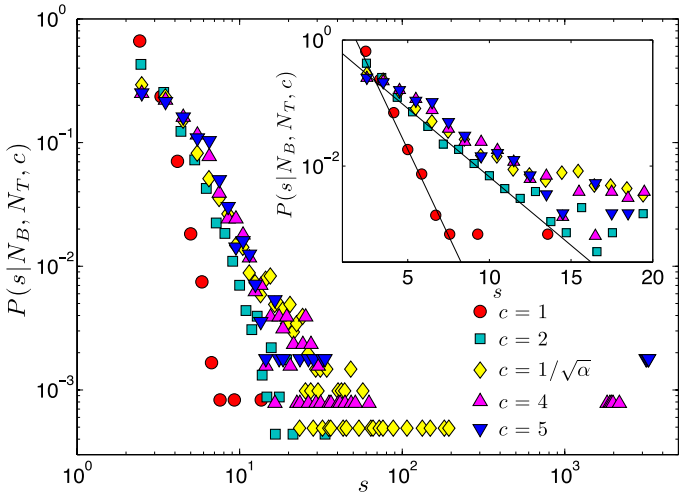

Figure 7. Size distribution of the connected components in $\mathcal{G}$ for $\alpha = 0.1$ and different values of c (see legend). Data are obtained from numerical simulations of systems with $N_T = \frac{3}{2}.10^3$. For $c < 1/\sqrt{\alpha}$ (in the under-percolated regime, $\circ$, $\square$) the decay is exponential with a finite cut-off, for $c = 1/\sqrt{\alpha}$ ($\Diamond$) the exponential decay broadens, while for $c > 1/\sqrt{\alpha}$ (in the over-percolated regime, $\triangle$,$\triangledown$) large components appear.

left panel). Hence, in the under-percolated regime the typical components in $\mathcal{G}$ are of finite size, and form cliques. The typical size of these cliques decays exponentially with s, as shown in figure 7 (see also [38]).

At the percolation threshold, larger loops start to appear in $\mathcal{B}$. For the graphs $\mathcal{G}$ this implies that two cliques can share not only nodes, but even links, and that two nodes i, j can display a coupling $|J_{ij}| \geqslant 2$ (see figure 5, middle panel). As a result, the simultaneous retrieval of all patterns within the same component is no longer ensured, and the distribution of component sizes will broaden.

In the over-percolated regime, a giant component of size $\mathcal{O}(N_T)$ emerges, while many isolated nodes and finite-size components will still remain. Now the average coordination number in the whole graph $\mathcal{G}$ is approximately $c^2\alpha$ (see (29)), but will be larger on the giant component. It is worth focusing on the macroscopic component to find out how it is organized, and how it compares to a random structure such as the ER graph. We note that even for the giant component the distribution $P(J|N_B, N_T, c)$ has a finite variance, and is concentrated on small values of J. To see this we calculate from (15) $P(J = 1|N_B, N_T, c)$, $P(J = 2|N_B, N_T, c)$ and infer the asymptotic behaviour for $P(J|N_B, N_T, c)$. From (15) we get

$$\begin{aligned} P(J=1|N_B,N_T,c) &= {\sum_{S=1}^{N_B-1}}' \frac{N_B!}{S!\left(\frac{N_B-S+1}{2}\right)!\left(\frac{N_B-S-1}{2}\right)!} p_w^S p_r^{N_B-S} \\ &= \frac{(N_B-\frac{1}{2})!}{\sqrt{\pi}\,(N_B-1)!}\left(\frac{c\alpha}{N}\right)^{2N_B-2} \\ &\quad \times \left[1-\left(\frac{c\alpha}{N}\right)^2\right] {}_2F_1\left(-\frac{N_B}{2}, 1-\frac{N_B}{2}, \frac{3}{2}, \left(\frac{c\alpha}{N_B}\right)^4 \left[1-\left(\frac{c\alpha}{N_B}\right)^2\right]^2\right) \\ &\approx N_B\left(\frac{c\alpha}{N_B}\right)^2 \left[1-\left(\frac{c\alpha}{N_B}\right)^2\right]^{N_B-1} \sim \left(\frac{c^2\alpha^2}{N_B}\right) \mathrm{e}^{-c^2\alpha^2/N_B}, \end{aligned} \tag{34}$$

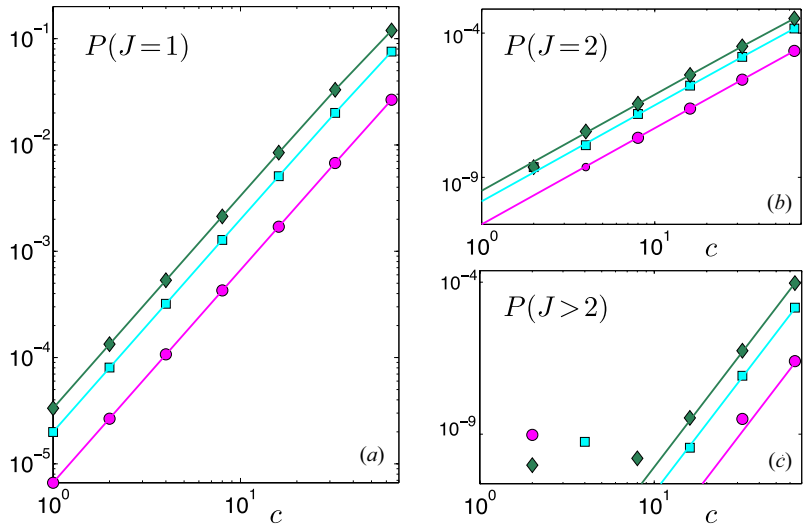

Figure 8. These plots show the average probability for an arbitrary link in $\mathcal{G}$ to have weight $J = 1$, $J = 2$ or $J > 2$, respectively, as a function of c, for several values of α (being $\alpha = 0.1$, $\alpha = 0.3$ and $\alpha = 0.5$, with markers as in the legend of figure 4). Numerical simulations (symbols) were carried out for systems with $N_T = \frac{3}{2}.10^3$ nodes, and are compared with the analytical estimates provided by (34), (35) and for arbitrary J (see formula below (35)).

where the prime restricts the sum to values of S with different parity from N_B (here assumed even), and where we used the isotropy ($p_r = p_l$) of the random walk. The asymptotic form obtained in the last step applies to $N_B \gg 1$. Similarly, for the case $J = 2$ we find

$$
\begin{aligned}
P(J=2|N_B,N_T,c) &= \sum_{S=0}^{N_B-2}{}' \frac{N_B!}{S!\left(\frac{N_B-S+2}{2}\right)!\left(\frac{N_B-S-2}{2}\right)!} p_w^S p_r^{N_B-S} \\
&= \frac{\left(\frac{c\alpha}{N_B}\right)N_B!}{\left(\frac{N_B}{2}+1\right)!\left(\frac{N_B}{2}-1\right)!2^{N_B}} \\
&\quad \times {}_2F_1\left(-1-\frac{N_B}{2}, 1-\frac{N_B}{2}, \frac{1}{2}, \left(\frac{N_B}{c\alpha}\right)^4\left[1-\left(\frac{c\alpha}{N_B}\right)^2\right]^2\right) \\
&\approx \left(\frac{c^2\alpha^2}{N_B}\right)^2\left[1-\left(\frac{c\alpha}{N_B}\right)^2\right]^{N_B-2}\sqrt{1-\frac{2}{N}\frac{1}{8}} \sim \left(\frac{c^2\alpha^2}{N_B}\right)^2 \mathrm{e}^{-c^2\alpha^2/N_B}. \qquad (35)
\end{aligned}
$$

Hence we expect the leading terms to scale as $P(J) \approx (c^2\alpha^2/N_B)^J[1-(c\alpha/N_B)^2]^{N_B-J}$. These results are confirmed by numerical simulations, with different choices for the parameters c and α; see figure 8.

4. Equilibrium analysis in the regime of high storage and finite connectivity

4.1. Definitions

We now turn to a statistical mechanics analysis and consider the effective network consisting solely of T-cells, with effective interactions described by the following Hamiltonian (rescaled

by a factor c relative to (3)):

$$\mathcal{H}(\boldsymbol{\sigma}|\xi) = -\frac{1}{2c}\sum_{i,j}^{N_T}\sum_{\mu}^{N_B}\xi_i^\mu\xi_j^\mu\sigma_i\sigma_j. \tag{36}$$

It is not *a priori* obvious that solving this model analytically will be possible. Most methods for spins systems on finitely connected heterogeneous graphs rely (explicitly or implicitly) on these being locally tree-like; due to the pattern dilution, the underlying topology of the system (36) is a heterogeneous graph with many short loops. From now on we will write N_T simply as N, with $N_B = \alpha N$ and $N \to \infty$. The cytokine components $\xi_i^\mu \in \{-1, 0, 1\}$ are quenched random variables, identically and independently distributed according to

$$P(\xi_i^\mu = 1) = P(\xi_i^\mu = -1) = \frac{c}{2N}, \qquad P(\xi_i^\mu = 0) = 1 - \frac{c}{N}, \tag{37}$$

with c finite. The Hamiltonian is normalized correctly: since the term $\sum_{i=1}^{N}\xi_i^\mu\sigma_i$ is $\mathcal{O}(1)$ both for condensed and non-condensed patterns [38], (36) is indeed extensive in N. The aim of this section is to compute the disorder-averaged free energy $\overline{f}$, at inverse temperature $\beta = T^{-1}$, where $\overline{\cdots}$ denotes averaging over the αN^2 variables $\{\xi_i^\mu\}$ and

$$f = -\lim_{N\to\infty}\frac{1}{\beta N}\log Z_N(\beta, \xi), \tag{38}$$

where $Z_N(\beta, \xi)$ is the partition function

$$Z_N(\beta, \xi) = \sum_{\boldsymbol{\sigma}\in\{-1,1\}^N} \mathrm{e}^{\frac{\beta}{2c}\sum_{\mu=1}^{\alpha N}(\sum_{i=1}^{N}\xi_i^\mu\sigma_i)^2}. \tag{39}$$

The state of the system can be characterized in terms of the αN (non-normalized) Mattis magnetizations, i.e. the overlaps between the system configuration and each cytokine pattern

$$M_\mu(\boldsymbol{\sigma}) = \sum_{i=1}^{N}\xi_i^\mu\sigma_i. \tag{40}$$

However, since in the high load regime the number of overlaps is extensive, it is more convenient to work with the overlap distribution

$$P(M|\boldsymbol{\sigma}) = \frac{1}{\alpha N}\sum_{\mu=1}^{\alpha N}\delta_{M_\mu(\boldsymbol{\sigma}),M}. \tag{41}$$

Although $M_\mu(\boldsymbol{\sigma})$ can take (discrete) values in the whole range $\{-N, -N+1, \ldots, N\}$, we expect that, due to dilution, the number of values that the $M_\mu(\boldsymbol{\sigma})$ assume remains effectively finite for large N, so that (41) represents an effective finite number of order parameters. In order to probe responses of the system to selected perturbations or triggering of clones we introduce external fields $\{\psi_\mu\}$ coupled to the overlaps $\{M_\mu(\boldsymbol{\sigma})\}$, so we consider the extended Hamiltonian

$$\mathcal{H}(\sigma, \xi) = -\frac{1}{2c}\sum_{i,j}^{N}\sum_{\mu}^{\alpha N}\xi_i^\mu\xi_j^\mu\sigma_i\sigma_j - \sum_{\mu=1}^{\alpha N}\psi_\mu M_\mu(\boldsymbol{\sigma}). \tag{42}$$

We also define the field distribution $P(\psi)$ and the joint distribution $P(M, \psi|\boldsymbol{\sigma})$ of magnetizations and fields, which is the most informative observable from a biological point of view (and of which $P(\psi)$ is a marginal):

$$P(\psi) = \frac{1}{\alpha N}\sum_{\mu=1}^{\alpha N}\delta(\psi - \psi_\mu), \qquad P(M, \psi|\boldsymbol{\sigma}) = \frac{1}{\alpha N}\sum_{\mu=1}^{\alpha N}\delta_{M,M_\mu(\boldsymbol{\sigma})}\delta(\psi - \psi_\mu). \tag{43}$$

4.2. The free energy

The free energy per spin (38) for the Hamiltonian (42) can be written as

$$f = -\lim_{N\to\infty}\frac{1}{\beta N}\log\sum_{\boldsymbol{\sigma}} \mathrm{e}^{\frac{\beta}{2c}\sum_{\mu=1}^{\alpha N} M_\mu^2(\boldsymbol{\sigma})+\beta\sum_{\mu=1}^{\alpha N}\psi_\mu M_\mu(\boldsymbol{\sigma})}. \tag{44}$$

We insert the following integrals of delta-functions written in Fourier representation

$$\begin{aligned} 1 &= \prod_M\prod_\psi\int \mathrm{d}P(M,\psi)\delta\left[P(M,\boldsymbol{\psi}) - \frac{1}{\alpha N}\sum_{\mu=1}^{\alpha N}\delta_{M,M_\mu(\boldsymbol{\sigma})}\delta(\psi-\psi_\mu)\right] \\ &= \prod_M\prod_\psi\int \frac{\mathrm{d}P(M,\psi)\mathrm{d}\hat{P}(M,\psi)}{2\pi/\Delta N}\mathrm{e}^{\mathrm{i}N\Delta\hat{P}(M,\psi)[P(M,\psi)-\frac{1}{\alpha N}\sum_{\mu=1}^{\alpha N}\delta_{M,M_\mu(\boldsymbol{\sigma})}\delta(\psi-\psi_\mu)]}. \end{aligned} \tag{45}$$

In the limit $\Delta\to 0$ we use $\Delta\sum_\psi\ldots\to\int\mathrm{d}\psi\ldots$, and we define the path integral measure $\{\mathrm{d}P\,\mathrm{d}\hat{P}\} = \lim_{\Delta\to 0}\,\mathrm{d}P(M,\psi)\,\mathrm{d}\hat{P}(M,\psi)\Delta N/2\pi$. This gives us

$$1 = \int\{\mathrm{d}P\,\mathrm{d}\hat{P}\}\,\mathrm{e}^{\mathrm{i}N\int\mathrm{d}\psi\sum_M\hat{P}(M,\psi)P(M,\psi)-\frac{\mathrm{i}}{\alpha}\sum_{\mu=1}^{\alpha N}\hat{P}(M_\mu(\boldsymbol{\sigma}),\psi_\mu)}. \tag{46}$$

Insertion into (44) leads us to an expression for f involving the density of states $\Omega[\hat{P}]$:

$$f = -\lim_{N\to\infty}\frac{1}{\beta N}\log\int\{\mathrm{d}P\,\mathrm{d}\hat{P}\}\,\mathrm{e}^{N\{\mathrm{i}\int\mathrm{d}\psi\sum_M P(M,\psi)\hat{P}(M,\psi)+\beta\alpha\int\mathrm{d}\psi\sum_M P(M,\psi)(\frac{M^2}{2c}+M\psi)+\Omega[\hat{P}]\}} \tag{47}$$

$$\Omega[\hat{P}] = \lim_{N\to\infty}\frac{1}{N}\log\sum_{\boldsymbol{\sigma}}\mathrm{e}^{-\frac{\mathrm{i}}{\alpha}\sum_\mu\hat{P}(M_\mu(\boldsymbol{\sigma}),\psi_\mu)}. \tag{48}$$

Hence via steepest descent integration for $N\to\infty$, and after avering the result over the disorder, we obtain:

$$\begin{aligned} \overline{f} = -\frac{1}{\beta}\mathrm{extr}_{\{P,\hat{P}\}}\Bigg\{&\mathrm{i}\int\mathrm{d}\psi\sum_M P(M,\psi)\hat{P}(M,\psi) \\ &+\beta\alpha\int\mathrm{d}\psi\sum_M P(M,\psi)\left(\frac{M^2}{2c}+M\psi\right)+\overline{\Omega}[\hat{P}]\Bigg\}, \end{aligned} \tag{49}$$

with

$$\overline{\Omega}[\hat{P}] = \lim_{N\to\infty}\frac{1}{N}\overline{\log\sum_{\boldsymbol{\sigma}}\mathrm{e}^{-\frac{\mathrm{i}}{\alpha}\sum_\mu\hat{P}(M_\mu(\boldsymbol{\sigma}),\psi_\mu)}}. \tag{50}$$

Working out the functional saddle-point equations that define the extremum in (49) gives

$$\hat{P}(M,\psi) = \mathrm{i}\alpha\beta\left(\frac{M^2}{2c}+M\psi\right), \qquad P(M,\psi) = \mathrm{i}\frac{\delta\overline{\Omega}[\hat{P}]}{\delta\hat{P}(M,\psi)}, \tag{51}$$

and inserting the first of these equations into (49) leads us to

$$\overline{f} = -\frac{1}{\beta}\overline{\Omega}[\hat{P}]\Big|_{\hat{P}(M,\psi)=\mathrm{i}\alpha\beta(\frac{M^2}{2c}+M\psi)}. \tag{52}$$

Hence, calculating the disorder-averaged free energy boils down to calculating (50). This can be done using the replica method, which is based on the identity $\overline{\log Z} = \lim_{n\to 0} n^{-1}\log\overline{Z^n}$, yielding

$$\overline{\Omega}[\hat{P}] = \lim_{N\to\infty}\lim_{n\to 0}\frac{1}{Nn}\log\sum_{\boldsymbol{\sigma}^1\ldots\boldsymbol{\sigma}^n}\overline{\mathrm{e}^{-\frac{\mathrm{i}}{\alpha}\sum_{\alpha=1}^n\sum_{\mu=1}^{\alpha N}\hat{P}(M_\mu(\boldsymbol{\sigma}^\alpha),\psi_\mu)}}. \tag{53}$$

The free energy (52) could also have been calculated directly from (44), by taking the average over disorder and using the replica identity. The advantage of working with the log-density of states is that, working out $\overline{\Omega}[\hat{P}]$ first for arbitrary functions $\hat{P}$ gives us via (51) a formula for the distribution $P(M,\psi)$, from which we can obtain useful information on the system retrieval phases and response to external perturbations. Finally we set $\hat{P}(M,\psi) = \mathrm{i}\alpha\beta\chi(M,\psi)$ with a real-valued function χ, to compactify our equations, with which we can write our problem as follows

$$\overline{f} = f[\chi]\Big|_{\chi(M,\psi)=\frac{M^2}{2c}+M\psi} \qquad f[\chi] = -\lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta N n}\log\sum_{\boldsymbol{\sigma}^1\ldots\boldsymbol{\sigma}^n}\overline{\mathrm{e}^{\beta\sum_{\alpha=1}^{n}\sum_{\mu=1}^{\alpha N}\chi(M_\mu(\boldsymbol{\sigma}^\alpha),\psi_\mu)}}, \tag{54}$$

$$P(M,\psi) = -\frac{1}{\alpha}\frac{\delta f[\chi]}{\delta\chi}\Big|_{\chi(M,\psi)=\frac{M^2}{2c}+M\psi}. \tag{55}$$

For simple tests of (54) and (55) in special limits see appendix A.

4.3. Derivation of saddle-point equations

From now on, unless indicated otherwise, all summations and products over α, i, and μ will be understood to imply $\alpha = 1\ldots n$, $i = 1\ldots N$, and $\mu = 1\ldots\alpha N$, respectively. We next need to introduce order parameters that allow us to carry out the disorder average in (54). The simplest choice is to isolate the overlaps themselves by inserting

$$1 = \prod_{\alpha\mu}\left[\sum_{M_{\alpha\mu}=-N}^{N}\delta_{M_{\alpha\mu},\sum_i\xi_i^\mu\sigma_i^\alpha}\right] = \prod_{\alpha\mu}\left[\sum_{M_{\alpha\mu}=-N}^{N}\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega_{\alpha\mu}}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega_{\alpha\mu}(M_{\alpha\mu}-\sum_i\xi_i^\mu\sigma_i^\alpha)}\right]. \tag{56}$$

This gives

$$f[\chi] = -\lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta N n}\log\left\{\prod_{\alpha\mu}\left[\sum_{M_{\alpha\mu}=-\infty}^{\infty}\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega_{\alpha\mu}}{2\pi}\right]\mathrm{e}^{\mathrm{i}\sum_{\alpha\mu}\omega_{\alpha\mu}M_{\alpha\mu}+\sum_{\alpha\mu}\beta\chi(M_\mu^\alpha,\psi_\mu)}\right.$$
$$\left.\times\sum_{\boldsymbol{\sigma}^1\ldots\boldsymbol{\sigma}^n}\overline{\mathrm{e}^{-\mathrm{i}\sum_i\sum_{\alpha\mu}\omega_{\alpha\mu}\xi_i^\mu\sigma_i^\alpha}}\right\}. \tag{57}$$

We can carry out the disorder average

$$\overline{\mathrm{e}^{-\mathrm{i}\sum_i\sum_{\alpha\mu}\omega_{\alpha\mu}\xi_i^\mu\sigma_i^\alpha}} = \prod_{i\mu}\left\{1-\frac{c}{N}+\frac{c}{2N}(\mathrm{e}^{\mathrm{i}\sum_\alpha\omega_{\alpha\mu}\sigma_i^\alpha}+\mathrm{e}^{-\mathrm{i}\sum_\alpha\omega_{\alpha\mu}\sigma_i^\alpha})\right\}$$
$$= \mathrm{e}^{\frac{c}{N}\sum_{i\mu}[\cos(\sum_\alpha\omega_{\alpha\mu}\sigma_i^\alpha)-1]+\mathcal{O}(N^0)}, \tag{58}$$

which leads us to

$$f[\chi] = -\lim_{n\to 0}\lim_{N\to\infty}\frac{1}{\beta N n}\log\left\{\prod_{\alpha\mu}\left[\sum_{M_{\alpha\mu}}\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega_{\alpha\mu}}{2\pi}\right]\cdot\mathrm{e}^{\mathrm{i}\sum_{\alpha\mu}\omega_{\alpha\mu}M_{\alpha\mu}+\sum_{\alpha\mu}\beta\chi(M_\mu^\alpha,\psi_\mu)}\right.$$
$$\left.\times\left[\sum_{\sigma_1\ldots\sigma_n}\mathrm{e}^{\frac{c}{N}\sum_\mu[\cos(\sum_\alpha\omega_{\alpha\mu}\sigma_\alpha)-1]}\right]^N\right\} \tag{59}$$

where we have also interchanged the limits $n\to 0$ and $N\to\infty$, as is usually done to progress in the calculation by using the saddle-point method. We next introduce n-dimensional vectors: $\boldsymbol{\sigma} = (\sigma_1,\ldots,\sigma_n)\in\{-1,1\}^n$, $\mathbf{M}^\mu = (M_{1\mu},\ldots,M_{n\mu})\in\mathbb{Z}^n$ and

$\boldsymbol{\omega}^\mu = (\omega_{1\mu}, \ldots, \omega_{n\mu}) \in [-\pi, \pi]^n$. This allows us to write (59) as

$$f[\chi] = -\lim_{n\to 0}\lim_{N\to\infty}\frac{1}{\beta N n}\log\left\{\prod_\mu\left[\sum_{\mathbf{M}^\mu}\int_{-\pi}^{\pi}\frac{\mathrm{d}\boldsymbol{\omega}^\mu}{(2\pi)^n}\right]\cdot \mathrm{e}^{\mathrm{i}\sum_\mu \boldsymbol{\omega}^\mu\cdot\mathbf{M}^\mu+\sum_\mu \beta\chi(\mathbf{M}^\mu,\psi_\mu)}\right.$$
$$\left.\times\left[\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\frac{c}{N}\sum_\mu[\cos(\boldsymbol{\omega}^\mu\cdot\boldsymbol{\sigma})-1]}\right]^N\right\}. \tag{60}$$

This last expression invites us to introduce the distribution $P(\boldsymbol{\omega}) = (\alpha N)^{-1}\sum_\mu \delta(\boldsymbol{\omega}-\boldsymbol{\omega}^\mu)$, for $\boldsymbol{\omega}\in[-\pi,\pi]^n$, via path integrals. We therefore insert

$$1 = \prod_{\boldsymbol{\omega}}\int \mathrm{d}P(\boldsymbol{\omega})\delta\left[P(\boldsymbol{\omega})-\frac{1}{\alpha N}\sum_\mu\delta(\boldsymbol{\omega}-\boldsymbol{\omega}^\mu)\right]$$
$$= \prod_{\boldsymbol{\omega}}\int\frac{\mathrm{d}P(\boldsymbol{\omega})\mathrm{d}\hat{P}(\boldsymbol{\omega})}{2\pi/\Delta N}\mathrm{e}^{\mathrm{i}N\Delta\hat{P}(\boldsymbol{\omega})[P(\boldsymbol{\omega})-\frac{1}{\alpha N}\sum_\mu\delta(\boldsymbol{\omega}-\boldsymbol{\omega}^\mu)]}. \tag{61}$$

In the limit $\Delta\to 0$ we use $\Delta\sum_{\boldsymbol{\omega}}\ldots\to\int\mathrm{d}\boldsymbol{\omega}\ldots$, and we define the usual path integral measure $\{\mathrm{d}P\mathrm{d}\hat{P}\} = \lim_{\Delta\to 0}\mathrm{d}P(\boldsymbol{\omega})\,\mathrm{d}\hat{P}(\boldsymbol{\omega})\Delta N/2\pi$. This converts the above to

$$1 = \int\{\mathrm{d}P\,\mathrm{d}\hat{P}\}\,\mathrm{e}^{\mathrm{i}N\int\mathrm{d}\boldsymbol{\omega}\hat{P}(\boldsymbol{\omega})P(\boldsymbol{\omega})-(i/\alpha)\sum_\mu\hat{P}(\boldsymbol{\omega}^\mu)}. \tag{62}$$

and upon insertion into (60) we get

$$f[\chi] = -\lim_{n\to 0}\lim_{N\to\infty}\frac{1}{\beta N n}\log\int\{\mathrm{d}P\mathrm{d}\hat{P}\}\mathrm{e}^{\mathrm{i}N\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}\hat{P}(\boldsymbol{\omega})P(\boldsymbol{\omega})}\left[\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\alpha c\int\mathrm{d}\boldsymbol{\omega}P(\boldsymbol{\omega})[\cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})-1]}\right]^N$$
$$\times\prod_{\mu=1}^{\alpha N}\left(\sum_{\mathbf{M}}\int_{-\pi}^{\pi}\frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n}\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\sum_\alpha\beta\chi(M^\alpha,\psi_\mu)-\frac{\mathrm{i}}{\alpha}\hat{P}(\boldsymbol{\omega})}\right). \tag{63}$$

In the limit $N\to\infty$, evaluation of the integrals by steepest descent leads to

$$f[\chi] = -\lim_{n\to 0}\frac{1}{\beta n}\mathrm{extr}_{\{P,\hat{P}\}}\Psi_n[\{P,\hat{P}\}], \tag{64}$$

$$\Psi_n[\{P,\hat{P}\}] = \mathrm{i}\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}\hat{P}(\boldsymbol{\omega})P(\boldsymbol{\omega})+\alpha\left\langle\log\left(\sum_{\mathbf{M}}\int_{-\pi}^{\pi}\frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n}\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\sum_\alpha\beta\chi(M^\alpha,\psi)-\frac{\mathrm{i}}{\alpha}\hat{P}(\boldsymbol{\omega})}\right)\right\rangle_\psi$$
$$+\log\left(\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\alpha c\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}P(\boldsymbol{\omega})[\cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})-1]}\right), \tag{65}$$

in which $\langle\ldots\rangle_\psi = \int\mathrm{d}\psi\,P(\psi)\ldots$. We mostly write $\langle\ldots\rangle$ in what follows, when there is no risk of ambiguities. The saddle-point equations are found by functional variation of Ψ_n with respect to P and $\hat{P}$, leading to

$$\hat{P}(\boldsymbol{\omega}) = \mathrm{i}c\alpha\frac{\sum_{\boldsymbol{\sigma}}[\cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})-1]\,\mathrm{e}^{\alpha c\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}'P(\boldsymbol{\omega}')[\cos(\boldsymbol{\omega}'\cdot\boldsymbol{\sigma})-1]}}{\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\alpha c\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}'P(\boldsymbol{\omega}')[\cos(\boldsymbol{\omega}'\cdot\boldsymbol{\sigma})-1]}}, \tag{66}$$

$$P(\boldsymbol{\omega}) = \left\langle\frac{\sum_{\mathbf{M}}\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\sum_\alpha\beta\chi(M^\alpha,\psi)-\frac{\mathrm{i}}{\alpha}\hat{P}(\boldsymbol{\omega})}}{\sum_{\mathbf{M}}\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}'\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}'\cdot\mathbf{M}+\sum_\alpha\beta\chi(M^\alpha,\psi)-\frac{\mathrm{i}}{\alpha}\hat{P}(\boldsymbol{\omega}')}}\right\rangle. \tag{67}$$

The joint distribution of fields and magnetizations now follows directly from (55) and (64), (65), and is seen to require only knowledge of the conjugate order parameters $\hat{P}(\boldsymbol{\omega})$:

$$\frac{P(M,\psi)}{P(\psi)} = \lim_{n\to 0}\left.\frac{\sum_{\mathbf{M}}\left(\frac{1}{n}\sum_\gamma\delta_{M,M_\gamma}\right)\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}\,\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\beta\sum_\alpha\chi(M^\alpha,\psi)-\frac{\mathrm{i}}{\alpha}\hat{P}(\boldsymbol{\omega})}}{\sum_{\mathbf{M}}\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}\,\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\beta\sum_\alpha\chi(M^\alpha,\psi)-\frac{\mathrm{i}}{\alpha}\hat{P}(\boldsymbol{\omega})}}\right|_{\chi=\frac{M^2}{2c}+\psi M}. \tag{68}$$

Thus the right-hand side (RHS) is an expression for $P(M|\psi)$. A last simple transformation $F(\boldsymbol{\omega}) = -\frac{\mathrm{i}}{c\alpha}\hat{P}(\boldsymbol{\omega}) + 1$ converts the saddle-point equations into

$$F(\boldsymbol{\omega}) = \frac{\sum_{\boldsymbol{\sigma}} \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma}) \mathrm{e}^{\alpha c \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}' P(\boldsymbol{\omega}') \cos(\boldsymbol{\omega}'\cdot\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}} \mathrm{e}^{\alpha c \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}' P(\boldsymbol{\omega}') \cos(\boldsymbol{\omega}'\cdot\boldsymbol{\sigma})}}, \tag{69}$$

$$P(\boldsymbol{\omega}) = \left\langle \frac{\mathrm{e}^{cF(\boldsymbol{\omega})} \prod_{\alpha} D_{\psi}(\omega_{\alpha}|\beta)}{\int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}' \mathrm{e}^{cF(\boldsymbol{\omega}')} \prod_{\alpha} D_{\psi}(\omega_{\alpha}|\beta)} \right\rangle, \tag{70}$$

where we have introduced

$$D_{\psi}(\omega|\beta) = \frac{1}{2\pi} \sum_{M\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\omega M + \beta\chi(M,\psi)}. \tag{71}$$

Similarly, (68) and (64) can now be expressed, respectively, as

$$P(M|\psi) = \lim_{n\to 0} \frac{\sum_{\mathbf{M}} \left(\frac{1}{n}\sum_{\gamma} \delta_{M,M_{\gamma}}\right) \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}\, \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M} + \beta\sum_{\alpha}\chi(M^{\alpha},\psi) + cF(\boldsymbol{\omega})}}{\sum_{\mathbf{M}} \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}\, \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M} + \beta\sum_{\alpha}\chi(M^{\alpha},\psi) + cF(\boldsymbol{\omega})}} \Bigg|_{\chi = M^2/2c + M\psi}, \tag{72}$$

and

$$f[\chi] = -\lim_{n\to 0} \frac{1}{\beta n} \Bigg\{ -c\alpha \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega} F(\boldsymbol{\omega}) P(\boldsymbol{\omega}) + \log\left(\sum_{\boldsymbol{\sigma}} \mathrm{e}^{\alpha c \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega} P(\boldsymbol{\omega})[\cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})-1]} \right) + \alpha \left\langle \log\left(\sum_{\mathbf{M}} \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M} + \sum_{\alpha}\beta\chi(M^{\alpha},\psi) + cF(\boldsymbol{\omega})} \right) \right\rangle \Bigg\}. \tag{73}$$

We note that the saddle-point equations guarantee that $P(\boldsymbol{\omega})$ is normalized correctly on $[-\pi,\pi]^n$, while for $F(\boldsymbol{\omega})$ we have (see appendix B)

$$\int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega} F(\boldsymbol{\omega}) = 0. \tag{74}$$

We observe that in the absence of external fields, i.e. for $\psi = 0$, the function (71) is real and symmetric:

$$D_0(\omega|\beta) = \frac{1}{2\pi} \sum_{M\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\omega M + \frac{\beta}{2c}M^2} \in \mathbb{R}, \qquad \forall\, \omega \in [-\pi,\pi] : D_0(-\omega|\beta) = D_0(\omega|\beta). \tag{75}$$

The introduction of external fields breaks the symmetry of $D_{\psi}(\omega|\beta)$ under the transformation $\omega \to -\omega$.

4.4. The RS ansatz—route I

To solve the saddle-point equations for $n \to 0$ we need to make an ansatz on the form of the order parameter functions $P(\boldsymbol{\omega})$ and $F(\boldsymbol{\omega})$. Since the conditioned overlap distribution (72) depends on $F(\boldsymbol{\omega})$ only, a first route to proceed is by eliminating the order function $P(\boldsymbol{\omega})$ from our equations and making a replica-symmetric (RS) ansatz for $F(\boldsymbol{\omega})$. Since $\boldsymbol{\omega} \in [-\pi,\pi]^n$ is continuous, the RS ansatz for $F(\boldsymbol{\omega})$ reads:

$$F(\boldsymbol{\omega}) = \int \{\mathrm{d}\pi\} W[\{\pi\}] \prod_{\alpha=1}^{n} \pi(\omega_{\alpha}), \tag{76}$$

where $W[\ldots]$ is a measure over functions, normalized according to $\int \{\mathrm{d}\pi\} W[\{\pi\}] = 1$ and nonzero (in view of (74)) only for functions $\pi(\ldots)$ that are real and obey $\int_{-\pi}^{\pi} \mathrm{d}\omega \pi(\omega) = 0$.

The RS ansatz (76) is to be inserted into the saddle-point equations. Insertion into (70) gives, with a normalization factor $C_n(\psi)$,

$$\begin{aligned} P(\boldsymbol{\omega}) &= \left\langle C_n^{-1}(\psi) \prod_\alpha D_\psi(\omega_\alpha|\beta)\, \mathrm{e}^{c\int\{\mathrm{d}\pi\}W[\{\pi\}]\prod_\alpha \pi(\omega_\alpha)} \right\rangle \\ &= \left\langle C_n^{-1}(\psi) \prod_\alpha D_\psi(\omega_\alpha|\beta) \sum_{k\geqslant 0} \frac{c^k}{k!} \left[\int \{\mathrm{d}\pi\} W[\{\pi\}] \prod_\alpha \pi(\omega_\alpha) \right]^k \right\rangle \\ &= \left\langle C_n^{-1}(\psi) \sum_{k\geqslant 0} \frac{c^k}{k!} \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W[\{\pi_\ell\}]] \prod_\alpha R_k(\omega_\alpha) \right\rangle, \end{aligned} \tag{77}$$

with

$$R_k(\omega) = D_\psi(\omega|\beta) \prod_{\ell=1}^k \pi_\ell(\omega). \tag{78}$$

Next we turn to (69). We first work out for $\boldsymbol{\sigma} \in \{-1, 1\}^n$ the quantity

$$\begin{aligned} L(\boldsymbol{\sigma}) &= \alpha c \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega} P(\boldsymbol{\omega}) \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma}) \\ &= \alpha c \Bigg\langle C_n^{-1}(\psi) \sum_{k\geqslant 0} \frac{c^k}{k!} \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W[\{\pi_\ell\}]] \\ &\quad \times \left[\frac{1}{2} \prod_\alpha \int_{-\pi}^{\pi} \mathrm{d}\omega_\alpha R_k(\omega_\alpha)\, \mathrm{e}^{\mathrm{i}\omega_\alpha\sigma^\alpha} + \frac{1}{2} \prod_\alpha \int_{-\pi}^{\pi} \mathrm{d}\omega_\alpha R_k(\omega_\alpha)\, \mathrm{e}^{-\mathrm{i}\omega_\alpha\sigma^\alpha} \right] \Bigg\rangle, \end{aligned} \tag{79}$$

with $\int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}\, P(\boldsymbol{\omega}) = 1$ requiring $L(\mathbf{0}) = \alpha c$. For Ising spins one can use the general identity

$$\tilde{R}_k(\sigma) = \int_{-\pi}^{\pi} \mathrm{d}\omega R_k(\omega)\, \mathrm{e}^{\mathrm{i}\omega\sigma} = B(\{R_k\})\, \mathrm{e}^{\mathrm{i}A(\{R_k\})\sigma}, \tag{80}$$

where B and A are, respectively, the absolute value and the argument of the complex function $\tilde{R}_k$ evaluated at the point 1, $\tilde{R}_k(1) = |\tilde{R}_k(1)|\, \mathrm{e}^{\mathrm{i}\phi_{\tilde{R}(1)}}$, i.e.

$$B(\{R_k\}) = |\tilde{R}_k(1)|, \qquad A(\{R_k\}) = \phi_{\tilde{R}(1)} = \arctan\left(\frac{\mathrm{Im}[\tilde{R}_k(1)]}{\mathrm{Re}[\tilde{R}_k(1)]} \right). \tag{81}$$

This simplifies (79) to

$$L(\boldsymbol{\sigma}) = \alpha c \left\langle C_n^{-1}(\psi) \sum_{k\geqslant 0} \frac{c^k}{k!} \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W(\{\pi_\ell\})] B^n(\{R_k\}) \cos\left[A(\{R_k\}) \sum_\alpha \sigma^\alpha \right] \right\rangle. \tag{82}$$

In order to have $L(\mathbf{0}) = \alpha c$ in the limit $n \to 0$, one must have $C_0(\psi) = e^c \forall \psi$. Inserting $L(\boldsymbol{\sigma})$ into (69) gives

$$K_n F(\boldsymbol{\omega}) = \sum_{\boldsymbol{\sigma}} \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})\, \mathrm{e}^{c\alpha\langle C_n^{-1}(\psi)\sum_{k\geqslant 0}\frac{c^k}{k!}\int\prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\}W[\{\pi_\ell\}]]B^n(\{R_k\})\cos[A(\{R_k\})\sum_\alpha \sigma^\alpha]\rangle}, \tag{83}$$

with

$$K_n = \sum_{\boldsymbol{\sigma}} \mathrm{e}^{c\alpha\langle C_n^{-1}(\psi)\sum_{k\geqslant 0}\frac{c^k}{k!}\int\prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\}W[\{\pi_\ell\}]]B^n(\{R_k\})\cos[A(\{R_k\})\sum_\alpha \sigma^\alpha]\rangle}. \tag{84}$$

Upon isolating the term $\sum_\alpha \sigma^\alpha$ via $\sum_m \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \mathrm{e}^{\mathrm{i}m\theta - \mathrm{i}\theta \sum_\alpha \sigma^\alpha} = 1$ we obtain

$$\begin{aligned} K_n F(\boldsymbol{\omega}) &= \sum_m \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \mathrm{e}^{\mathrm{i}m\theta + c\alpha \langle C_n^{-1}(\psi) \sum_{k\geqslant 0} \frac{c^k}{k!} \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W[\{\pi_\ell\}]] B^n(\{R_k\}) \cos[A(\{R_k\})m]\rangle} \\ &\quad \times \sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\mathrm{i}\theta \sum_\alpha \sigma^\alpha} \left(\frac{1}{2} \mathrm{e}^{\mathrm{i} \sum_\alpha \sigma^\alpha \omega^\alpha} + \frac{1}{2} \mathrm{e}^{-\mathrm{i} \sum_\alpha \sigma^\alpha \omega^\alpha} \right) \\ &= 2^{n-1} \sum_m \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \mathrm{e}^{\mathrm{i}m\theta + c\alpha \langle C_n^{-1}(\psi) \sum_{k\geqslant 0} \frac{c^k}{k!} \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W[\{\pi_\ell\}]] B^n(\{R_k\}) \cos[A(\{R_k\})m]\rangle} \\ &\quad \times \left[\prod_\alpha \cos(\omega^\alpha - \theta) + \prod_\alpha \cos(\omega^\alpha + \theta) \right]. \end{aligned} \tag{85}$$

The two terms inside the square brackets in the last line yield identical contributions to the θ-integral, so

$$\begin{aligned} K_n F(\boldsymbol{\omega}) &= 2^n \sum_m \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \mathrm{e}^{\mathrm{i}m\theta + c\alpha \langle C_n^{-1}(\psi) \sum_{k\geqslant 0} \frac{c^k}{k!} \int \prod_{\ell=1}^k [\{d\pi_\ell\} W[\{\pi_\ell\}]] B^n(\{R_k\}) \cos[A(\{R_k\})m]\rangle} \\ &\quad \times \prod_\alpha \cos(\omega^\alpha - \theta), \end{aligned} \tag{86}$$

with K_0 simply following from the demand $F(\boldsymbol{\omega} = \mathbf{0}) = 1$, as required by (69). Next we insert

$$1 = \int \{\mathrm{d}\pi\} \prod_\omega \delta[\pi(\theta) - \cos(\omega - \theta)], \tag{87}$$

where we have used the symbolic notation $\prod_\omega \delta[\pi(\omega) - f(\omega)]$ for the functional version of the δ-distribution, as defined by the identity $\int \{\mathrm{d}\pi\}\, G[\{\pi\}] \prod_\omega \delta[\pi(\omega) - f(\omega)] = G[\{f\}]$. This leads us to

$$\begin{aligned} K_n F(\boldsymbol{\omega}) &= 2^n \sum_m \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \mathrm{e}^{\mathrm{i}m\theta + c\alpha \langle C_n^{-1}(\psi) \sum_{k\geqslant 0} \frac{c^k}{k!} \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W[\{\pi_\ell\}]] B^n(\{R_k\}) \cos[A(\{R_k\})m]\rangle} \\ &\quad \times \int \{\mathrm{d}\pi\} \prod_\omega \delta[\pi(\theta) - \cos(\omega - \theta)] \prod_\alpha \pi(\omega^\alpha). \end{aligned} \tag{88}$$

Substituting (76) for $F(\boldsymbol{\omega})$ in the left-hand side of this last equation shows that in the replica limit $n \to 0$, our RS ansatz indeed generates a saddle point if

$$W[\{\pi\}] = \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \lambda(\theta|W) \prod_\omega \delta[\pi(\omega) - \cos(\omega - \theta)], \tag{89}$$

with the short-hand

$$\lambda(\theta|W) = K_0^{-1} \sum_{m\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}m\theta + c\alpha \sum_{k\geqslant 0} \frac{c^k \mathrm{e}^{-c}}{k!} \langle \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W[\{\pi_\ell\}]] \cos[A(\{R_k\})m]\rangle}. \tag{90}$$

The constant K_0 follows simply from normalization, which now takes the form $\int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \lambda(\theta|W) = 1$, giving

$$\begin{aligned} K_0 &= \int \frac{\mathrm{d}\theta}{2\pi} \sum_{m\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}m\theta + c\alpha \sum_{k\geqslant 0} \frac{c^k \mathrm{e}^{-c}}{k!} \langle \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W[\{\pi_\ell\}]] \cos[A(\{R_k\})m]\rangle} \\ &= \sum_{m\in\mathbb{Z}} \delta_{m,0}\, \mathrm{e}^{c\alpha \sum_{k\geqslant 0} \frac{c^k \mathrm{e}^{-c}}{k!} \langle \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W[\{\pi_\ell\}]] \cos[A(\{R_k\})m]\rangle} = \mathrm{e}^{c\alpha}. \end{aligned} \tag{91}$$

We then arrive at

$$\lambda(\theta|W) = \sum_{m\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}m\theta + c\alpha \sum_{k\geqslant 0} \frac{c^k \mathrm{e}^{-c}}{k!} \langle \int \prod_{\ell=1}^k [\{\mathrm{d}\pi_\ell\} W[\{\pi_\ell\}]] [\cos[A(\{R_k\})m] - 1]\rangle}. \tag{92}$$

It is convenient to write $D(\omega|\beta) = D'(\omega|\beta) + \mathrm{i}D''(\omega|\beta)$, with $D'(\omega|\beta) = \mathrm{Re}[D(\omega|\beta)]$ and $D''(\omega|\beta) = \mathrm{Im}[D(\omega|\beta)]$. Similarly, we write $R_k(\omega) = R'_k(\omega) + \mathrm{i}R''_k(\omega)$. We note that for $\chi(M,\psi) = M^2/2c + M\psi$ the function $D_\psi(\omega|\beta)$ defined in (71) has several useful properties, e.g.

$$\forall\, \omega \in [-\pi,\pi] : D'_\psi(-\omega|\beta) = D'_\psi(\omega|\beta), \qquad D''_\psi(-\omega|x) = -D''_\psi(\omega|x), \tag{93}$$

$$\int_{-\pi}^{\pi} \mathrm{d}\omega D_\psi(\omega|\beta) = \sum_{M\in\mathbb{Z}} \mathrm{e}^{\beta\chi(M,\psi)} \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi}\, \mathrm{e}^{\mathrm{i}\omega M} = \sum_{M\in\mathbb{Z}} \mathrm{e}^{\beta\chi(M,\psi)}\delta_{M,0} = 1, \tag{94}$$

$$D_\psi(\omega|0) = \frac{1}{2\pi}\sum_{M\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\omega M} = \delta(\omega) \text{ for } \omega \in [-\pi,\pi]. \tag{95}$$

From (81) we have

$$A(\{R_k\}) = \arctan\left[\frac{\mathrm{Im}[\tilde{R}_k(1)]}{\mathrm{Re}[\tilde{R}_k(1)]}\right] = \arctan\left[\frac{\int_{-\pi}^{\pi}\mathrm{d}\omega[R'_k(\omega)\sin\omega + R''_k(\omega)\cos\omega]}{\int_{-\pi}^{\pi}\mathrm{d}\omega[R'_k(\omega)\cos\omega - R''_k(\omega)\sin\omega]}\right], \tag{96}$$

and insertion in (92) gives

$$\lambda(\theta|W) = \sum_{m\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}m\theta + c\alpha\sum_{k\geqslant 0}\frac{c^k \mathrm{e}^{-c}}{k!}\int\prod_{\ell=1}^{k}[\{\mathrm{d}\pi_\ell\}W[\{\pi_\ell\}]]\{\cos[m\arctan f_k(\{\pi_1,\ldots,\pi_k\})]-1\}}, \tag{97}$$

with

$$f_k(\{\pi_1,\ldots,\pi_k\}) = \frac{\int_{-\pi}^{\pi}\mathrm{d}\omega[D'(\omega|\beta)\sin\omega + D''(\omega|\beta)\cos\omega]\prod_{\ell=1}^{k}\pi_\ell(\omega)}{\int_{-\pi}^{\pi}\mathrm{d}\omega[D'(\omega|\beta)\cos\omega - D''(\omega|\beta)\sin\omega]\prod_{\ell=1}^{k}\pi_\ell(\omega)}. \tag{98}$$

For high temperatures $D'(\omega|0) = \delta(\omega)$ and $D''(\omega|0) = 0$, so $f_k(\{\pi_1,\ldots,\pi_k\}) = 0$ and $\lambda(\theta|W) = \delta(\theta)$. Hence

$$\beta = 0: \quad W[\{\pi\}] = \prod_{\omega}\delta[\pi(\omega) - \cos(\omega)]. \tag{99}$$

We note that for any symmetric set of functions $\{\pi_1,\ldots,\pi_k\}$ one has, from (98), $f_k(\{\pi_1,\ldots,\pi_k\}) = 0$ due to the symmetry properties (93) of D_ψ, and thus $\lambda(\theta|W) = \delta(\theta)$. Hence, (99) is a solution of (89) for *all* temperatures, and the only solution at infinite temperature.

4.5. Conditioned distribution of overlaps

In order to give a physical interpretation to the RS solution (76), (99), we consider the conditioned overlap distribution (72). Insertion of (99) into (76) gives

$$F(\boldsymbol{\omega}) = \int\{\mathrm{d}\pi\}W[\{\pi\}]\prod_{\alpha}\pi(\omega_\alpha) = \prod_{\alpha}\cos(\omega_\alpha),$$

and subsequent insertion into (72) leads to, with C_n and $\tilde{C}_n$ representing the normalization constants,

$$\begin{aligned} P(M|\psi) &= \lim_{n\to 0} C_n^{-1}\sum_{\mathbf{M}}\left(\frac{1}{n}\sum_{\gamma=1}^{n}\delta_{M,M_\gamma}\right)\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}\,\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\beta\sum_\alpha\chi(M_\alpha,\psi)}\sum_{k\geqslant 0}\frac{c^k}{k!}\prod_{\alpha}\cos^k(\omega_\alpha) \\ &= \lim_{n\to 0}\frac{\tilde{C}_n^{-1}}{n}\sum_{k\geqslant 0}\frac{c^k}{k!}\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}\prod_{\alpha}\cos^k(\omega_\alpha)\int_{-\pi}^{\pi}\mathrm{d}\lambda\mathrm{e}^{\mathrm{i}\lambda M}\sum_{\gamma=1}^{n}\sum_{M_\gamma\in\mathbb{Z}}\mathrm{e}^{\mathrm{i}(\omega_\gamma-\lambda)M_\gamma+\chi(M_\gamma,\psi)} \end{aligned}$$

$$\times \prod_{\alpha\neq\gamma}\sum_{M_\alpha} \mathrm{e}^{\mathrm{i}\omega_\alpha M_\alpha+\chi(M^\alpha,\psi)}$$
$$= \lim_{n\to 0}\frac{C_n^{-1}}{n}\sum_{k\geqslant 0}\frac{c^k}{k!}\int_{-\pi}^{\pi}\mathrm{d}\lambda\,\mathrm{e}^{\mathrm{i}\lambda M}\sum_{\gamma=1}^{n}\int_{-\pi}^{\pi}\mathrm{d}\omega_\gamma\cos^k(\omega_\gamma)D_\psi(\omega_\gamma-\lambda|\beta)$$
$$\times \prod_{\alpha\neq\gamma}\int_{-\pi}^{\pi}\mathrm{d}\omega_\alpha\cos^k(\omega_\alpha)D_\psi(\omega_\alpha|\beta)$$
$$= \lim_{n\to 0} C_n^{-1}\sum_{k\geqslant 0}\frac{c^k}{k!}\int_{-\pi}^{\pi}\mathrm{d}\lambda\,\mathrm{e}^{\mathrm{i}\lambda M} I_k(\lambda,\beta)I_k^{n-1}(0,\beta), \tag{100}$$

with

$$\begin{aligned} I_k(\lambda,\beta) &= \int_{-\pi}^{\pi}\mathrm{d}\omega\cos^k(\omega)D_\psi(\omega-\lambda|\beta)\\ &= \frac{1}{2^k}\sum_{n=0}^{k}\binom{k}{n}\int_{-\pi}^{\pi}\mathrm{d}\omega\,\mathrm{e}^{-\mathrm{i}\omega(k-2n)}\sum_{m\in\mathbb{Z}}\mathrm{e}^{\mathrm{i}(\omega-\lambda)m+\beta\chi(m,\psi)}\\ &= \frac{1}{2^k}\sum_{n=0}^{k}\binom{k}{n}\mathrm{e}^{-\mathrm{i}\lambda(k-2n)+\beta\chi(k-2n,\psi)}\\ &= \frac{1}{2^k}\sum_{m=-k}^{k}\binom{k}{\frac{k-m}{2}}\mathrm{e}^{-\mathrm{i}\lambda m+\beta\chi(m,\psi)}. \end{aligned} \tag{101}$$

We can now work out

$$\int_{-\pi}^{\pi}\mathrm{d}\lambda\,\mathrm{e}^{\mathrm{i}\lambda M}I_k(\lambda,\beta) = \begin{cases} 2^{-k}\left(\binom{k}{(k-M)/2}\right)\mathrm{e}^{\beta\chi(M,\psi)} & \text{if } |M|\leqslant k\\ 0 & \text{if } |M|>k, \end{cases} \tag{102}$$

and obtain our desired formula for $P(M|\psi)$ corresponding to the saddle point (99), in which the normalization constant comes out as $C_0=\mathrm{e}^c$. The result then is

$$P(M|\psi) = \sum_{k\geqslant|M|}\mathrm{e}^{-c}\frac{c^k}{k!}\frac{\binom{k}{(k-M)/2}\mathrm{e}^{\beta\chi(M,\psi)}}{\sum_{m=-k}^{k}\binom{k}{(k-m)/2}\mathrm{e}^{\beta\chi(m,\psi)}}. \tag{103}$$

We can rewrite this result, with the short-hand $p_c(k)=\mathrm{e}^{-c}c^k/k!$, in the more intuitive form

$$P(M|\psi) = \sum_{k\geqslant 0}p_c(k)P(M|k,\psi), \tag{104}$$

$$P(M|k,\psi) = \theta\left(k-|M|+\frac{1}{2}\right)\frac{\binom{k}{(k-M)/2}\mathrm{e}^{\beta\chi(M,\psi)}}{\sum_{m=-k}^{k}\binom{k}{(k-m)/2}\mathrm{e}^{\beta\chi(m,\psi)}}. \tag{105}$$

We recognize that $p_c(k)$ is the asymptotic probability that any cytokine pattern $\left(\xi_1^\mu,\ldots,\xi_N^\mu\right)$ has k nonzero entries; since each pattern has N independent entries with probability c/N to be nonzero, k will for $N\to\infty$ indeed be a Poissonian random variable with average c. Hence, $P(M|k,\psi)$ is the conditional probability of having an overlap of value M, given that the cytokine pattern concerned has k nonzero entries and is triggered by an external field ψ. We have apparently mapped the neural network with N neurons and $N_B=\alpha N$ diluted stored patterns to a system of k neurons with a single undiluted binary pattern. We will see that this is due to the fact that in the regime where RS theory holds one is always able, as a consequence

of the dilution, to decompose the original system into an extensive number of independent finite-size subsystems, each recalling one particular pattern.

The solution (99), leading to (105), is a saddle point for any temperature. At infinite temperatures it is the only solution, and simplifies further. For $\beta = 0$ expression (105) gives

$$P(M|k,\psi) = 2^{-k}\binom{k}{(k-M)/2}\theta\left(k-|M|+\frac{1}{2}\right), \tag{106}$$

which is the probability that a system of k spins has an overlap M with an undiluted stored pattern, if each spin behaves completely randomly. This describes, as expected, an immune network behaving as a paramagnet, i.e. unable to retrieve the stored strategies. For the distribution of overlaps we find

$$P(M) = \mathrm{e}^{-c}\sum_{k\geqslant|M|}\frac{(\frac{1}{2}c)^k}{k!}\binom{k}{(k-M)/2}. \tag{107}$$

In the limit $\beta\to\infty$, the sum in the denominator of (105) is dominated by the value of m which maximizes $\chi(m,\psi) = m^2/2c + \psi m$, being $m = k\mathrm{sgn}(\psi)$ if $\psi \neq 0$ and $m = \pm k$ for $\psi = 0$. In either case we obtain

$$\sum_{m=-k}^{k}\binom{k}{(k-m)/2}\mathrm{e}^{\beta\chi(m,\psi)} \sim \begin{cases}\mathrm{e}^{\beta(k^2/2c+k|\psi|)} & \psi\neq 0\\ 2\mathrm{e}^{\beta k^2/2c} & \psi = 0, k\neq 0.\end{cases} \tag{108}$$

Substitution into (105) and (104) subsequently gives

$$\begin{aligned}&\lim_{\beta\to\infty}P(M|\psi)\\ &= \lim_{\beta\to\infty}\begin{cases}\mathrm{e}^{-c}\sum_{k\geqslant|M|}\dfrac{c^k}{k!}\left(\dbinom{k}{(k-M)/2}\right)\mathrm{e}^{-\beta(k^2-M^2)/2c-\beta|\psi|(k-\mathrm{sgn}(\psi)M)} & \text{if } \psi\neq 0\\ \dfrac{1}{2}\mathrm{e}^{-c}\sum_{k\geqslant|M|}\dfrac{c^k}{k!}\left(\dbinom{k}{(k-M)/2}\right)\mathrm{e}^{-\beta(k^2-M^2)/2c} & \text{if } \psi = 0, M\neq 0\end{cases}\\ &= \begin{cases}\mathrm{e}^{-c} & \text{if } M = 0\\ \theta(M\psi)\,\mathrm{e}^{-c}c^{|M|}/|M|! & \text{if } \psi\neq 0, M\neq 0\\ \dfrac{1}{2}\,\mathrm{e}^{-c}c^{|M|}/|M|! & \text{if } \psi = 0, M\neq 0.\end{cases}\end{aligned} \tag{109}$$

Similarly we have

$$\psi\neq 0: P(M|k,\psi) = \delta_{|M|,k}(\delta_{M,0} + \theta(\psi M)(1-\delta_{M,0})), \tag{110}$$

$$\psi = 0: P(M|k,\psi) = \delta_{|M|,k}\big(\delta_{M,0} + \tfrac{1}{2}(1-\delta_{M,0})\big). \tag{111}$$

For $k > 0$ this describes the error-free activation or inhibition of a stored strategy with k nonzero entries.

For intermediate temperatures a plot of (105) shows that without external fields, $P(M|0)$ acquires two symmetric peaks at large overlaps (in absolute value), as β is increased from $\beta = 0$; see figure 9 (top left panel). Unlike typical magnetic systems in the thermodynamic limit, there is no spontaneous ergodicity breaking at $\psi = 0$; the system acts effectively as an extensive number of independent finite subsystems, each devoted to a single B-clone. Each size-k subsystem oscillates randomly between the two peaks in $P(M|0)$, with a characteristic switching timescale $t_k \sim \mathrm{e}^{\beta k^2/2c}$, which grows with the size k of the subsystem and remains finite at finite temperature.

Introducing a field ψ reduces the overlap peak at M values opposite in sign to the field; this peak will eventually disappear for sufficiently strong fields (figure 9, top right panel). The

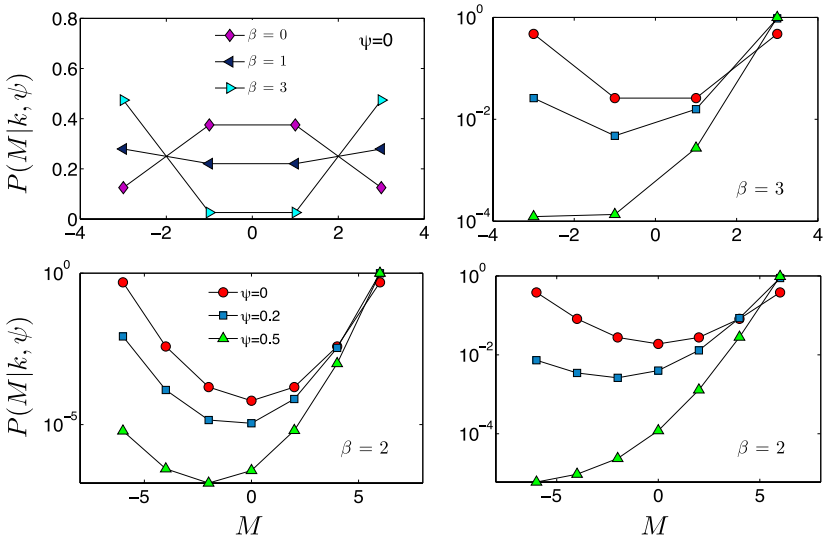

Figure 9. Conditioned overlap distribution $P(M|k, \psi)$ corresponding to the state (76), (99), as given by formula (105). The top panels refer to $k = c = 3$. Left: $\beta = 0, 1, 3$ and $\psi = 0$; Right: $\beta = 3$ and $\psi = 0, 0.2, 0.5$. The bottom panels refer to $\psi = 0, 0.2, 0.5$ and $\beta = 2$. Left: $c = 3, k = 6$. Right: $c = k = 6$. Note that $M \in \{-k, -k + 1, \ldots, k - 1, k\}$, so the lines connecting the markers are only guides to the eye.

field-induced asymmetry in the height of the two peaks increases at smaller temperatures and larger sizes (figure 9, bottom panels). Thus, external fields trigger the system towards either activation or inhibition of a strategy (e.g. clonal expansion versus contraction), whereas in their absence the system oscillates stochastically between the two.

Beyond the multiple clonal expansions, achieved in the present model through activation signalling from the T-cells to B-cells via appropriately diluted cytokine patterns, the apparent emergence of regular inhibitory signals sent to the B-clones that are not expanding (in the absence of external fields triggering those clones) is a biologically fundamental feature for homeostasis. B-cells that are not receiving a significant number of signals undergo a process called 'anergy' [34, 35], and will eventually die. Thus, the ability to support fast switching between positive and negative signals to multiple clones in parallel, which is achieved in a rather natural way in the present multitasking network, has further welcome implications.

4.6. Alternative formulation of the theory before the RS ansatz

The approach developed in the previous section led to transparent formulae for the distribution of overlaps in the RS state (99), and even allows us to derive analytically the condition defining the (continuous) phase transition where (99) ceases to hold (see appendix C). However, the states beyond the transition point are better described within an alternative (but mathematically equivalent) formulation of the theory. This alternative approach is based on formulating our equations first in terms of the following quantities:

$$L(\boldsymbol{\sigma}) = \alpha c \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}\, P(\boldsymbol{\omega}) \cos(\boldsymbol{\omega} \cdot \boldsymbol{\sigma}), \qquad Q(\boldsymbol{\omega}) = \mathrm{e}^{cF(\boldsymbol{\omega})}. \tag{112}$$

Both $P(\boldsymbol{\omega})$ and $Q(\boldsymbol{\omega})$ are only defined for $\boldsymbol{\omega} \in [-\pi, \pi]^n$. In terms of (112) we can write our earlier saddle-point equations (70), (69) as

$$P(\boldsymbol{\omega}) = \left\langle \frac{Q(\boldsymbol{\omega}) \sum_{\mathbf{M}\in\mathbb{Z}^n} \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\sum_\alpha \chi(M_\alpha,\psi)}}{\int_\pi^\pi \mathrm{d}\boldsymbol{\omega}'\, Q(\boldsymbol{\omega}') \sum_{\mathbf{M}\in\mathbb{Z}^n} \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}'\cdot\mathbf{M}+\sum_\alpha \chi(M_\alpha,\psi)}} \right\rangle_\psi, \tag{113}$$

$$\log Q(\boldsymbol{\omega}) = c \frac{\sum_{\boldsymbol{\sigma}\in\{-1,1\}^n} \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})\, \mathrm{e}^{L(\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}\in\{-1,1\}^n} \mathrm{e}^{L(\boldsymbol{\sigma})}}, \tag{114}$$

and the free energy (73) as

$$f[\chi] = -\lim_{n\to 0} \frac{1}{\beta n} \left\{ \log\left(\sum_{\boldsymbol{\sigma}} \mathrm{e}^{L(\boldsymbol{\sigma})-c\alpha} \right) - \frac{\sum_{\boldsymbol{\sigma}} L(\boldsymbol{\sigma})\, \mathrm{e}^{L(\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}} \mathrm{e}^{L(\boldsymbol{\sigma})}} \right.$$
$$\left. + \alpha \left\langle \log\left(\sum_{\mathbf{M}} \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\sum_\alpha \beta\chi(M_\alpha,\psi)} Q(\boldsymbol{\omega}) \right) \right\rangle_\psi \right\}, \tag{115}$$

where we used $\alpha \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}\, P(\boldsymbol{\omega}) \log Q(\boldsymbol{\omega}) = \sum_{\boldsymbol{\sigma}} L(\boldsymbol{\sigma}) \mathrm{e}^{L(\boldsymbol{\sigma})} / \sum_{\boldsymbol{\sigma}} \mathrm{e}^{L(\boldsymbol{\sigma})}$. Clearly $\int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega} P(\boldsymbol{\omega}) = 1$, $Q(\boldsymbol{\omega}) \in \mathbb{R}$, $Q(-\boldsymbol{\omega}) = Q(\boldsymbol{\omega})$, and $Q(\mathbf{0}) = \mathrm{e}^c$. We can now switch from the order parameter $Q(\boldsymbol{\omega})$ to a new order parameter $\tilde{Q}(\mathbf{M})$, defined on $\mathbf{M} \in \mathbb{Z}^n$, via the following one-to-one transformations:

$$\tilde{Q}(\mathbf{M}) = \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} Q(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}}, \qquad Q(\boldsymbol{\omega}) = \sum_{\mathbf{M}\in\mathbb{Z}^n} \tilde{Q}(\mathbf{M})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}}. \tag{116}$$

The validity of these equations follows from the two identities $(2\pi)^{-1} \int_{-\pi}^{\pi} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega m} = \delta_{m0}$ for $m \in \mathbb{Z}$, and $(2\pi)^{-1} \sum_{M\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\omega M} = \delta(\omega)$ for $\omega \in [-\pi, \pi]$. By construction we now have $\sum_{\mathbf{M}} \tilde{Q}(\mathbf{M}) = \mathrm{e}^c$. Moreover, since $Q(-\boldsymbol{\omega}) = Q(\boldsymbol{\omega})$ we also know that $\tilde{Q}(\mathbf{M}) = (2\pi)^{-n} \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega} Q(\boldsymbol{\omega}) \cos(\boldsymbol{\omega}\cdot\mathbf{M}) \in \mathbb{R}$. One can write the saddle-point equations in terms of these order functions (see appendix D for details):

$$\tilde{Q}(\mathbf{M}) = \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega} \cos(\boldsymbol{\omega}\cdot\mathbf{M}) \exp\left[c \frac{\sum_{\boldsymbol{\sigma}} \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})\, \mathrm{e}^{L(\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}} \mathrm{e}^{L(\boldsymbol{\sigma})}} \right], \tag{117}$$

$$L(\boldsymbol{\sigma}) = \alpha c\, \mathrm{e}^{\frac{\beta n}{2c}} \left\langle \frac{\sum_{\mathbf{M}} \tilde{Q}(\mathbf{M})\, \mathrm{e}^{\beta\sum_\alpha \chi(M_\alpha,\psi)} \cosh\left[\beta\left(\frac{1}{c}\mathbf{M}\cdot\boldsymbol{\sigma} + \psi \sum_\alpha \sigma^\alpha\right)\right]}{\sum_{\mathbf{M}} \tilde{Q}(\mathbf{M})\, \mathrm{e}^{\beta\sum_\alpha \chi(M_\alpha,\psi)}} \right\rangle_\psi. \tag{118}$$

and the free energy reads

$$f[\chi] = -\lim_{n\to 0} \frac{1}{\beta n} \left\{ \log \sum_{\boldsymbol{\sigma}} \mathrm{e}^{L(\boldsymbol{\sigma})-c\alpha} - \frac{\sum_{\boldsymbol{\sigma}} L(\boldsymbol{\sigma})\, \mathrm{e}^{L(\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}} \mathrm{e}^{L(\boldsymbol{\sigma})}} \right.$$
$$\left. + \alpha \left\langle \log\left[\sum_{\mathbf{M}} \mathrm{e}^{\sum_\alpha \beta\chi(M_\alpha,\psi)} \tilde{Q}(\mathbf{M}) \right] \right\rangle_\psi \right\}. \tag{119}$$

From (72) we find that the distribution of overlaps can be written as

$$P(M|\psi) = \lim_{n\to 0} \left. \frac{\sum_{\mathbf{M}} \left(\frac{1}{n} \sum_{\gamma=1}^n \delta_{M,M_\gamma}\right) \mathrm{e}^{\beta\sum_\alpha \chi(M_\alpha,\psi)} \tilde{Q}(\mathbf{M})}{\sum_{\mathbf{M}} \mathrm{e}^{\beta\sum_\alpha \chi(M_\alpha,\psi)} \tilde{Q}(\mathbf{M})} \right|_{\chi(M,\psi)=M^2/2c+M\psi}. \tag{120}$$

In appendix E we confirm the correctness of (120) in several special limits.

4.7. The RS ansatz—route II

We now try to construct the RS solution of our new equations (118), (117), by applying the RS ansatz to the functions $L(\boldsymbol{\sigma})$ and $\tilde{Q}(\mathbf{M})$:

$$L(\boldsymbol{\sigma}) = \alpha c \int \mathrm{d}h W(h) \prod_{\alpha=1}^{n} \mathrm{e}^{\beta h \sigma^\alpha}, \qquad \tilde{Q}(\mathbf{M}) = \mathrm{e}^c \int \{\mathrm{d}\pi\} W[\{\pi\}] \prod_{\alpha} \pi(M^\alpha), \tag{121}$$

with $\int \mathrm{d}h\, W(h) = 1, W(h) = W(-h)$, and with a (normalized) functional measure $W[\pi]$ that is only nonzero for functions $\pi(M)$ that are themselves normalized according to $\sum_{M\in\mathbb{Z}} \pi(M) = 1$. This ansatz meets the requirements $L(-\boldsymbol{\sigma}) = L(\boldsymbol{\sigma})$, $L(\mathbf{0}) = \alpha c$ and $\sum_{\mathbf{M}} \tilde{Q}(\mathbf{M}) = \mathrm{e}^c$, and is the most general form of the functions $L(\boldsymbol{\sigma})$ and $\tilde{Q}(\mathbf{M})$, that is invariant under all replica permutations. The advantage of this second formulation of the theory is that it allows us to work with a distribution $W(h)$ of effective fields, instead of functional measures over distributions, which have easier physical interpretations, and are more easy to solve numerically from self-consistent equations.

We relegate to appendix F all the details of the derivation of the RS equations, based on the form (121), the results of which can be summarized as follows. The RS functional measure $W[\pi]$ and the field distribution $W(h)$ obey the following closed equations:

$$W(h) = \int \{\mathrm{d}\pi\} W[\pi] \left\langle \left\langle \delta \left[h - \tau\psi - \frac{1}{2\beta} \log \left(\frac{\sum_M \pi(M)\, \mathrm{e}^{\beta(M^2/2c + M(\psi+\tau/c))}}{\sum_M \pi(M) \mathrm{e}^{\beta(M^2/2c + M(\psi-\tau/c))}} \right) \right] \right\rangle_{\psi} \right\rangle_{\tau=\pm 1}, \tag{122}$$

$$W[\pi] = \mathrm{e}^{-c} \sum_{k\geqslant 0} \frac{c^k}{k!}\, \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int_{-\infty}^{\infty} \mathrm{d}h_1 \ldots \mathrm{d}h_r \left[\prod_{s\leqslant r} W(h_s) \right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \times \prod_M \delta \left[\pi(M) - \frac{\left\langle \mathrm{e}^{\beta \sum_{s\leqslant r} h_s \sigma_{\ell_s}} \delta_{M, \sum_{\ell\leqslant k} \sigma_\ell} \right\rangle_{\sigma_1 \ldots \sigma_k}}{\left\langle \mathrm{e}^{\beta \sum_{s\leqslant r} h_s \sigma_{\ell_s}} \right\rangle_{\sigma_1 \ldots \sigma_k}} \right]. \tag{123}$$

Both $W(h)$ and $W[\pi]$ are correctly normalized, $W(h) = W(-h)$, and $W[\pi]$ allows only for functions π such that $\pi(M) = \pi(-M)$ and $\sum_M \pi(M) = 1$. We can substitute the second equation into the first and eliminate the functional measure $W[\pi]$, leaving us with a compact RS equation for the field distribution $W(h)$ only:

$$W(h) = \mathrm{e}^{-c} \sum_{k\geqslant 0} \frac{c^k}{k!}\, \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int_{-\infty}^{\infty} \mathrm{d}h_1 \ldots \mathrm{d}h_r \left[\prod_{s\leqslant r} W(h_s) \right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \times \left\langle \left\langle \delta \left[h - \tau\psi - \frac{1}{2\beta} \log \left(\frac{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k} \tau_\ell)^2/2c + \beta(\sum_{\ell\leqslant k} \tau_\ell)(\psi+\tau/c) + \beta \sum_{s\leqslant r} h_s \tau_{\ell_s}} \right\rangle_{\tau_1 \ldots \tau_k = \pm 1}}{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k} \tau_\ell)^2/2c + \beta(\sum_{\ell\leqslant k} \tau_\ell)(\psi-\tau/c) + \beta \sum_{s\leqslant r} h_s \tau_{\ell_s}} \right\rangle_{\tau_1 \ldots \tau_k = \pm 1}} \right) \right] \right\rangle_{\psi} \right\rangle_{\tau=\pm 1}. \tag{124}$$

We see that $W(h) = \delta(h)$ is a solution of (124) for any temperature; one easily confirms that this is in fact the earlier state (99), recovered within the alternative formulation of the theory. If we inspect continuous bifurcations of new solutions with moments $m_r = \int \mathrm{d}h h^r W(h)$ different from zero, we find (see appendix G) a second order transition along the critical surface in the (α, β, c)-space defined by

$$1 = \alpha c^2 \sum_{k\geqslant 0} \mathrm{e}^{-c} \frac{c^k}{k!} \left\{ \frac{\int \mathrm{D}z \tanh(z\sqrt{\beta/c} + \beta/c) \cosh^{k+1}(z\sqrt{\beta/c} + \beta/c)}{\int \mathrm{D}z \cosh^{k+1}(z\sqrt{\beta/c} + \beta/c)} \right\}^2. \tag{125}$$

We note that the RHS obeys $0 \leqslant \text{RHS} \leqslant \alpha c^2$, with $\lim_{\beta\to 0} \text{RHS} = 0$ and $\lim_{\beta\to\infty} \text{RHS} = \alpha c^2$. Hence a transition at finite temperature $T_c(\alpha, c) = \beta_c^{-1}(\alpha, c) > 0$ exists to a new state with $W(h) \neq \delta(h)$ as soon as $\alpha c^2 > 1$. The critical temperature becomes zero when $\alpha c^2 = 1$, consistent with the percolation threshold (11) derived from the network analysis. We show in appendix H that the critical surface (125) is indeed identical to the one found in (C.20), within the approach involving functional distributions.

Finally, within the new formulation of the theory, the RS field-conditioned overlap distribution is found to be

$$P(M|\psi) = \lim_{n\to 0} \frac{\int\{\mathrm{d}\pi\}W[\pi]\big(\sum_{M'}\pi(M')\,\mathrm{e}^{\beta(M'^2/2c+\psi M')}\big)^{n-1}\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M)}}{\int\{\mathrm{d}\pi\}W[\pi]\big(\sum_{M'}\pi(M')\,\mathrm{e}^{\beta(M'^2/2c+\psi M')}\big)^{n}}$$
$$= \int\{\mathrm{d}\pi\}W[\pi]\left\{\frac{\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M)}}{\sum_{M'}\pi(M')\,\mathrm{e}^{\beta(M'^2/2c+\psi M')}}\right\}. \tag{126}$$

Insertion of (123) allows us to eliminate the functional measure in favour of effective field distributions:

$$P(M|\psi) = \mathrm{e}^{-c}\sum_{k\geqslant 0}\frac{c^k}{k!}\mathrm{e}^{-\alpha ck}\sum_{r\geqslant 0}\frac{(\alpha c)^r}{r!}\int_{-\infty}^{\infty}\mathrm{d}h_1\ldots\mathrm{d}h_r\left[\prod_{s\leqslant r}W(h_s)\right]\sum_{\ell_1\ldots\ell_r\leqslant k}$$
$$\times\left\{\frac{\big\langle \mathrm{e}^{\beta\sum_{s\leqslant r}h_s\sigma_{\ell_s}}\delta_{M,\sum_{\ell\leqslant k}\sigma_\ell}\big\rangle_{\sigma_1\ldots\sigma_k}\mathrm{e}^{\beta(M^2/2c+\psi M)}}{\sum_{M'}\big\langle \mathrm{e}^{\beta\sum_{s\leqslant r}h_s\sigma_{\ell_s}}\delta_{M',\sum_{\ell\leqslant k}\sigma_\ell}\big\rangle_{\sigma_1\ldots\sigma_k}\mathrm{e}^{\beta(M'^2/2c+\psi M')}}\right\}$$
$$= \mathrm{e}^{-c}\sum_{k\geqslant 0}\frac{c^k}{k!}\mathrm{e}^{-\alpha ck}\sum_{r\geqslant 0}\frac{(\alpha c)^r}{r!}\int_{-\infty}^{\infty}\mathrm{d}h_1\ldots\mathrm{d}h_r\left[\prod_{s\leqslant r}W(h_s)\right]\sum_{\ell_1\ldots\ell_r\leqslant k}$$
$$\times\left\{\frac{\big\langle\delta_{M,\sum_{\ell\leqslant k}\tau_\ell}\,\mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta\psi\sum_{\ell\leqslant k}\tau_\ell+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\big\rangle_{\tau_1\ldots\tau_k=\pm 1}}{\big\langle\mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta\psi\sum_{\ell\leqslant k}\tau_\ell+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\big\rangle_{\tau_1\ldots\tau_k=\pm 1}}\right\}. \tag{127}$$

Again, we can rewrite this result (127) in the form (104), which is more useful for investigating the system's performance since it quantifies the statistics of overlaps relative to their maximum value k, with

$$P(M|k,\psi) = \mathrm{e}^{-\alpha ck}\sum_{r\geqslant 0}\frac{(\alpha c)^r}{r!}\int_{-\infty}^{\infty}\mathrm{d}h_1\ldots\mathrm{d}h_r\left[\prod_{s\leqslant r}W(h_s)\right]\sum_{\ell_1\ldots\ell_r\leqslant k}$$
$$\times\left\{\frac{\big\langle\delta_{M,\sum_{\ell\leqslant k}\tau_\ell}\,\mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta\psi\sum_{\ell\leqslant k}\tau_\ell+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\big\rangle_{\tau_1\ldots\tau_k=\pm 1}}{\big\langle\mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta\psi\sum_{\ell\leqslant k}\tau_\ell+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\big\rangle_{\tau_1\ldots\tau_k=\pm 1}}\right\}. \tag{128}$$

The latter formula shows very clearly that h is to be interpreted as a clonal interference field, which is caused by overlapping signalling strategies in the bi-partite graph $\mathcal{B}$ and leads to clique interactions in the effective H–H graph $\mathcal{G}$. Biologically these interference fields can manifest themselves in unwanted clonal expansions (in the absence of the required antigen), or unwanted clonal reductions (in the presence of the required antigen), due to accidental (frozen) random interactions between clones. Fortunately, we see in figure 10 that even above the percolation threshold $\alpha c^2 = 1$ the system is able to suppress clonal cross-talk (i.e. have $W(h) = \delta(0)$), provided the noise level is nonzero, and that even in the cross-talk phase the signalling performance of the system degrades only smoothly (see the section below).

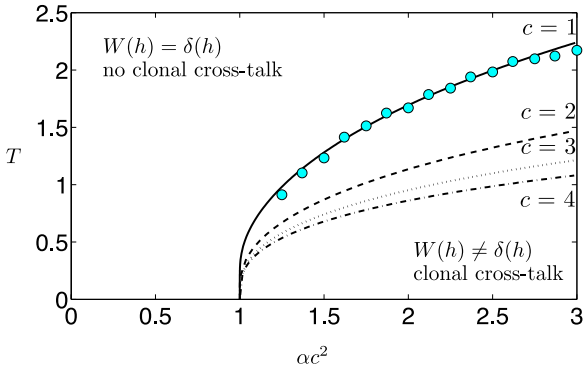

Figure 10. Transition lines (125) for $c = 1, 2, 3, 4$, in the $(\alpha c^2, T)$ plane, with $T = \beta^{-1}$. The distribution $W(h)$ represents the statistics of clonal cross-talk fields, which are caused by increased connectivity in the graph $\mathcal{G}$. If $W(h) = \delta(h)$ the clones are controlled via signalling strategies that can act independently; we see that this is possible even above the percolation threshold if the temperature (i.e. the signalling noise) is nonzero. Circles: transition calculated via numerical solution of (124) for $c = 1$ (see section 5).

5. Numerical results: population dynamics and numerical simulations

5.1. Population dynamics calculation of the cross-talk field distribution

We solve numerically equation (124) for the clonal interference field distribution $W(h)$ with a population dynamics algorithm [50], which is based on interpreting (124) as the fixed-point equation of a stochastic process and simulating this process numerically. One observes that (124) has the structural form

$$W(h) = \langle\langle \delta[h - h(k, r, \mathbf{h}, \boldsymbol{\ell}, \tau, \psi)]\rangle\rangle_{k,r,\mathbf{h},\boldsymbol{\ell},\tau,\psi}, \tag{129}$$

with the following set of random variables:

$k \sim$ Poisson (c)

$r \sim$ Poisson $(\alpha c k)$

$\mathbf{h} = (h_1, \ldots, h_r)$: r i.i.d. random fields with probability density $W(h)$

$\boldsymbol{\ell} = (\ell_1, \ldots, \ell_r)$: r i.i.d. discrete random variables, distributed uniformly over $\{1, \ldots, k\}$

τ : dichotomic random variable, distributed uniformly over $\{-1, 1\}$

ψ : distributed according to $P(\psi)$

and with

$$h(k, r, \mathbf{h}, \boldsymbol{\ell}, \tau, \psi) = \tau\psi + \frac{1}{2\beta}\log\left(\frac{\left\langle e^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta(\sum_{\ell\leqslant k}\tau_\ell)(\psi+\tau/c)+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm1}}{\left\langle e^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta(\sum_{\ell\leqslant k}\tau_\ell)(\psi-\tau/c)+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm1}}\right). \tag{130}$$

We approximate $W(h)$ by the empirical field frequencies computed from a large number (i.e. a population) of fields, which are made to evolve by repeated numerical iteration of a stochastic map. We start by initializing S fields $h_s \in \mathbb{R}$, with $s = 1, \ldots, S$, randomly with uniform probabilities over the interval $[-h_{\max}, h_{\max}]$. Their empirical distribution then represents the zero-step approximation $W_0(h)$ of $W(h)$. We then evolve the fields stochastically via the following Markovian process, giving at each step n an empirical distribution $W_n(h)$ which as n increases gives an increasingly precise approximation of the invariant measure $W(h)$:

- choose randomly the variables $k, r, \boldsymbol{\ell}, \tau, \psi$ according to their (known) probability distributions;
- choose randomly r fields $\mathbf{h} = h_1, \ldots, h_r$ from the S fields available, i.e. draw r fields from the probability distribution $W_{n-1}(h)$ of the previous step;
- compute $h(k, r, \mathbf{h}, \boldsymbol{\ell}, \tau, \psi)$;
- choose randomly one field from the set of the S available, and set its value to $h(k, r, \mathbf{h}, \boldsymbol{\ell}, \tau, \psi)$;

In all population dynamics calculations in this paper we used populations of size $S = 5000$.

We iterate the procedure until convergence, checking that every $\mathcal{O}(S^2)$ steps the distance between different $W_n(h)$, and speed up the computation of $h(k, r, \mathbf{h}, \boldsymbol{\ell}, \tau, \psi)$ by rewriting it as

$$
\begin{aligned}
&h(k, r, \mathbf{h}, \boldsymbol{\ell}, \tau, \psi) \\
&= \tau\psi + \frac{1}{2\beta}\log\left(\frac{\int \mathrm{D}z\left\langle \mathrm{e}^{z\sqrt{\beta/c}\sum_{\ell\leqslant k}\tau_\ell+\beta(\sum_{\ell\leqslant k}\tau_\ell)(\psi+\tau/c)+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm1}}{\int \mathrm{D}z\left\langle \mathrm{e}^{z\sqrt{\beta/c}\sum_{\ell\leqslant k}\tau_\ell+\beta(\sum_{\ell\leqslant k}\tau_\ell)(\psi-\tau/c)+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm1}}\right) \\
&= \tau\psi + \frac{1}{2\beta}\log\left(\frac{\int \mathrm{D}z\prod_{\ell\leqslant k}\cosh[z\sqrt{\beta/c}+\beta(\psi+\tau/c)+\beta\sum_{s\leqslant r}h_s\delta_{\ell\ell_s}]}{\int \mathrm{D}z\prod_{\ell\leqslant k}\cosh[z\sqrt{\beta/c}+\beta(\psi-\tau/c)+\beta\sum_{s\leqslant r}h_s\delta_{\ell\ell_s}]}\right),
\end{aligned}
\tag{131}
$$

which requires Gaussian integration instead of the average over $\{\tau_1, \ldots, \tau_k\}$. Having computed $W(h)$, we can build $P(M|\psi)$ using equation (127). The latter can be rewritten as

$$
\begin{aligned}
P(M|\psi) &= \left\langle\!\!\left\langle \frac{\left\langle \delta_{M,\sum_{\ell\leqslant k}\tau_\ell}\mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta\psi\sum_{\ell\leqslant k}\tau_\ell+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm1}}{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta\psi\sum_{\ell\leqslant k}\tau_\ell+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm1}}\right\rangle\!\!\right\rangle_{k,r,\mathbf{h},\boldsymbol{\ell},\psi} \\
&= \left\langle\!\!\left\langle \frac{\left\langle \delta_{M,\sum_{\ell\leqslant k}\tau_\ell}\mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta\psi\sum_{\ell\leqslant k}\tau_\ell+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm1}}{Z(k, r, \mathbf{h}, \boldsymbol{\ell}, \psi)}\right\rangle\!\!\right\rangle_{k,r,\mathbf{h},\boldsymbol{\ell},\psi},
\end{aligned}
\tag{132}
$$

with $Z(\ldots) = \int \mathrm{D}z\prod_{\ell\leqslant k}\cosh[z\sqrt{\frac{\beta}{c}}+\beta(\psi-\tau/c)+\beta\sum_{s\leqslant r}h_s\delta_{\ell\ell_s}]$ as determined in (131). Hence we can carry out the ensemble average over the parameters $\{\boldsymbol{\tau}, k, r, \mathbf{h}, \boldsymbol{\ell}, \psi\}$ in this last expression as an arithmetic average over a large number L of samples drawn from their joint distribution, for which in this paper we choose $L = \mathcal{O}(10^7)$. The distribution (128) is handled in the same way, and can be rewritten as

$$
P(M|k, \psi) = \left\langle\!\!\left\langle \frac{\left\langle \delta_{M,\sum_{\ell\leqslant k}\tau_\ell}\mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c+\beta\psi\sum_{\ell\leqslant k}\tau_\ell+\beta\sum_{s\leqslant r}h_s\tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm1}}{Z(k, r, \mathbf{h}, \boldsymbol{\ell}, \psi)}\right\rangle\!\!\right\rangle_{r,\mathbf{h},\boldsymbol{\ell},\psi},
\tag{133}
$$

i.e. upon simply omitting the averaging over k.

In the interest of transparency and an intuitive understanding, it helps to identify the physical meaning of the random variables involved in the above stochastic process. Given a subsystem of k spins linked to a particular cytokine pattern (say pattern $\mu = 1$, without loss of generality), we may ask how many other patterns $\mu \neq 1$ interfere with it. This number is the cardinality of the set

$$
R = \left\{\xi_i^\mu,\, i=1,\ldots,N;\, \mu=2,\ldots,\alpha N:\, \xi_i^\mu\xi_i^1 \neq 0\right\}.
\tag{134}
$$

With each of the k spins (labelled i, with $\xi_i^1 \neq 0$) correspond $\alpha N-1$ cytokine variables ξ_i^μ with $\mu > 1$. Hence we have, for a set of k spins, $k(\alpha N-1)$ independent possibilities to generate interfering cytokine signals, each nonzero with probability c/N. Thus, for $N \to \infty$ the number of possible interferences is a Poissonian random variable with mean αck, which is recognized as the variable r. For each value of r we next ask on *which* of the k spins each interference acts, i.e. which are the r indices i such that $\xi_i^\mu\xi_i^1 \neq 0$ for some $\mu > 1$. Each i refers to one of the k

spins selected, so we can describe this situation by r random variables ℓ_s, with $s = 1, \ldots, r$, each distributed uniformly in $\{1, \ldots, k\}$, with are recognized as the vector $\boldsymbol{\ell}$. The parameters k, r and $\boldsymbol{\ell}$ considered so far depend only on the (quenched) structure of the B–H network. By conditioning on these random variables we can write

$$P(M|\psi) = \sum_{k=0}^{\infty} \mathrm{e}^{-c} \frac{c^k}{k!} \sum_{r=0}^{\infty} \mathrm{e}^{-\alpha c k} \frac{(\alpha c k)^r}{r!} \sum_{\ell_1,\ldots,\ell_r=1}^{k} k^{-r} P(M|k, r, \boldsymbol{\ell}, \psi)$$
$$= \left\langle\!\!\left\langle \sum_{\boldsymbol{\sigma}} \delta_{M,\sum_{\ell=1}^{k} \xi_\ell^1 \sigma_\ell} Z^{-1}(k, r, \boldsymbol{\ell}, \psi) \mathrm{e}^{-\beta H(\boldsymbol{\sigma}|k,r,\boldsymbol{\ell},\psi)} \right\rangle\!\!\right\rangle_{k,r,\boldsymbol{\ell}}. \tag{135}$$

Inside the brackets we have the overlap M of a single pattern ($\mu = 1$) with k non-null entries, whose correlation with the other patterns is specified uniquely by the parameters $(k, r, \boldsymbol{\ell})$. We can write the effective Hamiltonian governing this k-spin subsystem by isolating in the Hamiltonian (42) $\mu = 1$ contribution:

$$H_{\mathrm{eff}}(\boldsymbol{\sigma}) = -M_1^2(\boldsymbol{\sigma})/2c - \psi M_1(\boldsymbol{\sigma}) - \sum_{i=1}^{k} \sigma_i \sum_{\mu>2} \xi_i^\mu (M_\mu(\boldsymbol{\sigma})/c + \psi_\mu). \tag{136}$$

Upon transforming $\tau_\ell = \xi_\ell^1 \sigma_\ell$, and defining $h_\ell^\mu(\boldsymbol{\tau}) = \xi_\ell^1 \xi_\ell^\mu (M_\mu/c + \psi_\mu)$, and using the meaning of the parameters r and ℓ_s, we arrive at a description involving r nonzero fields $h_s(\boldsymbol{\tau})$, each acting on a spin ℓ_s:

$$H_{\mathrm{eff}}(\tau_1, \ldots, \tau_k) = -\left(\sum_{\ell \leqslant k} \tau_\ell \right)^2 /2c - \psi \sum_{\ell \leqslant k} \tau_\ell - \sum_{s \leqslant r} h_s(\boldsymbol{\tau}) \tau_{\ell_s}. \tag{137}$$

If we then regard each field $h_s(\boldsymbol{\tau})$ as a independent random field (conditional on $(k, r, \boldsymbol{\ell})$), with probability distribution $W(h_s)$, we arrive at

$$P(M) = \left\langle\!\!\left\langle \int \mathrm{d}\mathbf{h} W(\mathbf{h}) \left\langle \delta_{M,\sum_{\ell=1}^{k} \tau_\ell} \frac{\mathrm{e}^{\beta(\sum_{\ell \leqslant k} \tau_\ell)^2/2c + \beta\psi \sum_{\ell \leqslant k} \tau_\ell + \beta \sum_{s \leqslant r} h_s \tau_{\ell_s}}}{Z(k, r, \boldsymbol{\ell}, \psi)} \right\rangle_{\boldsymbol{\tau}} \right\rangle\!\!\right\rangle_{k,r,\boldsymbol{\ell}}. \tag{138}$$

This is exactly equation (127) obtained within the RS ansatz. Hence the parameters $\mathbf{h}$ in (129) represent the effective fields induced by the cross-talk interference of cytokine patterns. The only difference between the rigorous RS derivation and the above heuristic one is that in the former we effectively find $W(\mathbf{h}) = \prod_{s \leqslant r} W(h_s)$, i.e. the random fields are independent. This may not always be the case: if we recall the definition of the r effective fields, viz. $h_\ell^\mu(\boldsymbol{\tau}) = \xi_\ell^1 \xi_\ell^\mu (M_\mu/c + \psi_\mu)$, we see that as soon as different patterns have more then one spin in common, their interference fields will not be independent. One therefore expects the RS equation to no longer be exact if the bi-partite B–H network is not-tree like but contains loops. Finally we note that in the absence of external fields, the effective fields take values in $\mathbb{Z}/c$, so that $W(h)$ becomes a superposition of delta-functions, consistent with the numerical results in figure 13. This allows in principle a rewriting of the self-consistency equation (124) in terms of the amplitudes of such superposition, which are scalar parameters rather than distribution and may be easier to find numerically.

5.2. Critical line, overlap distributions and the interference field distribution

First we use the population dynamics algorithm to validate the location of the critical line (125). To do so we keep α fixed and compute $W(h)$ for different values of the inverse temperature β. From the solution we compute $m_2 = \int \mathrm{d}h h^2 W(h)$, and determine for which β-value it becomes nonzero (starting from the high temperature phase), i.e. where clonal cross-talk sets

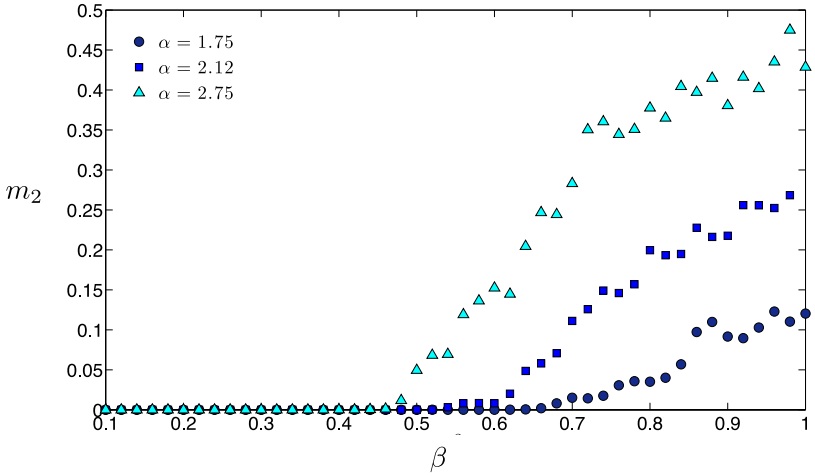

Figure 11. Widths (variances) $m_2 = \int \mathrm{d}h W(h) h^2$ of the distribution of clonal cross-talk fields, shown as markers versus the inverse temperature β for different values of α. In all cases $c = 1$. The values of m_2 are calculated from the population dynamics solution of (129), and are (modulo finite-size fluctuations in population dynamics algorithm) in excellent agreement with (125). The latter predicts that for the α-values considered and for $c = 1$ the widths m_2 should become nonzero at: $\beta_c = 0.6634$ (for $\alpha = 1.75$), $\beta_c = 0.5639$ (for $\alpha = 2.12$), and $\beta_c = 0.4707$ (for $\alpha = 2.75$).

in. The result is shown in figure 11, which reveals excellent agreement between the predicted bifurcation temperatures (125) and those obtained from population dynamics. We also see that there is no evidence for discontinuous transitions. In figure 10 we plotted the bifurcation temperatures obtained via population dynamics versus αc^2 (markers), together with the full transition lines predicted by (125) and again see excellent agreement between the two.

In the under-percolated regime $\alpha c^2 < 1$, there is no possibility of a phase transition and the only solution of (124) is $W(h) = \delta(h)$. Both equations (127), (128) then lose their dependence on α, and after some simple manipulations we recover our earlier results (104), (105). In figure 12 we test our predictions for the overlap statistics against the results of numerical (Monte Carlo) simulations of the spin process defined by Hamiltonian (3), in the absence of external fields. There is excellent agreement between theory and numerical experiment. Comparison of $P(M|k, 0)$ to $P(M|0)$ shows that the former changes shape as the inverse temperature β is increased from zero, from a single peak at $M = 0$ to two symmetric peaks, showing that the system behaviour at high versus low noise levels is very different. In contrast, $P(M|0)$ always has a maximum in $M = 0$, due to the Poissonian distribution of k, and does not capture the two different behaviours. Hence $P(M|k, 0)$ is the most useful quantifier of retrieval behaviour, which from now on we will simply denote in the absence of external fields as $P(M|k)$.

When $\alpha c^2 > 1$, and below the critical line defined by equation (125), the solution of equation (124) in the absence of external fields will exhibit $W(h) \neq \delta(h)$; see figure 13. As a consequence, the effective Boltzmann factor governing the behaviour of a set of k spins, linked to a single pattern, acquires a term $\beta \sum_{s \leqslant r} h_s \tau_{\ell_s}$ (see equation (133)). This term means that each subsystem is no longer isolated as in the under-percolated regime, but feels the interference due to the other patterns in the form of effective random fields. Numerical results for $P(M|k)$ in the over-percolated regime, including comparisons between population dynamics calculations and measurements taken in numerical simulations (involving spin systems with $N = 3.10^4$

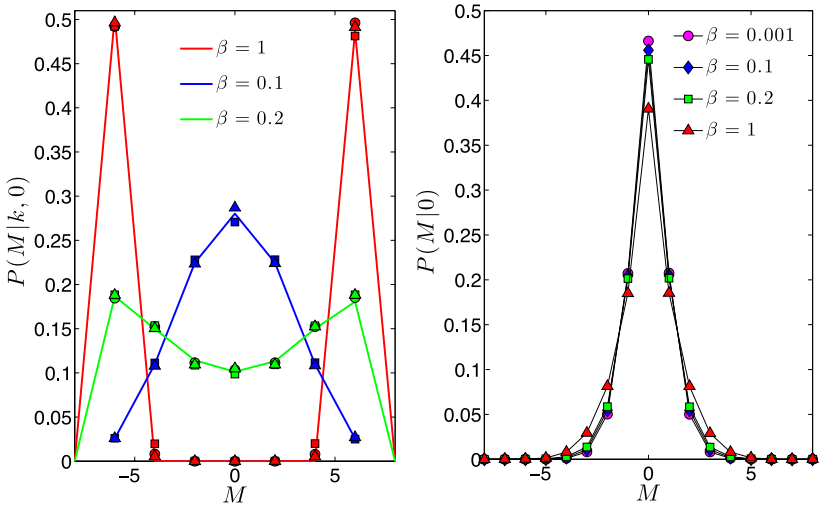

Figure 12. Left: degree-conditioned conditioned overlap distribution $P(M|k,0)$ in the under-percolated regime, for $k = 6$, $c = 1$, and different β values (see legend), without external fields. Solid lines: theoretical predictions. Markers: the results of measuring the overlap statistics in Monte Carlo simulations of the spin system with Hamiltonian (3), with $N = 3.10^4$ H-cells. Different symbols represent different values of α, namely $\alpha = 0.005$ (bullets), $\alpha = 0.008$ (squares) and $\alpha = 0.011$ (triangles). The theory predicts that here $P(M|k,0)$ is independent of α, which we find to be confirmed. Right panel: the overlap distribution $P(M|0)$ at zero field in the under-percolated regime, for $k = 6$, $c = 1$ and $\alpha = 0.5$, and different temperatures (see legend). Note that $M \in \mathbb{Z}$, so the line segments are only guides to the eye.

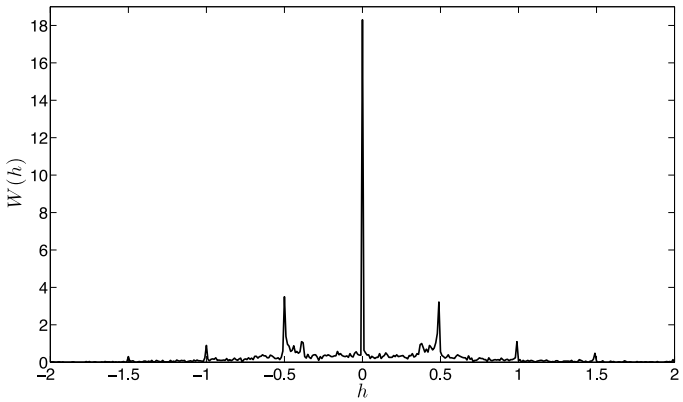

Figure 13. The clonal cross-talk interference field distribution $W(h)$ below the critical temperature and in the absence of external fields, as calculated (approximately) via the population dynamics algorithm, for $c = 2$, $\alpha = 2$ and $\beta = 6.2$. Note that the support of $W(h)$ is $\mathbb{Z}/c$. One indeed observes the weight of $W(h)$ being concentrated on these points; due to the finite population size in the algorithm (here $S = 5000$) one finds small nonzero values for $h \notin \mathbb{Z}/c$ due to finite-size fluctuations.

H-cells) are shown in figure 14. Again we observe excellent agreement. Moreover, we see that in the regime of clonal cross-talk the system's signalling performance degrades only gracefully; provided α is not yet too large, the overlap distribution maintains its bimodal form.

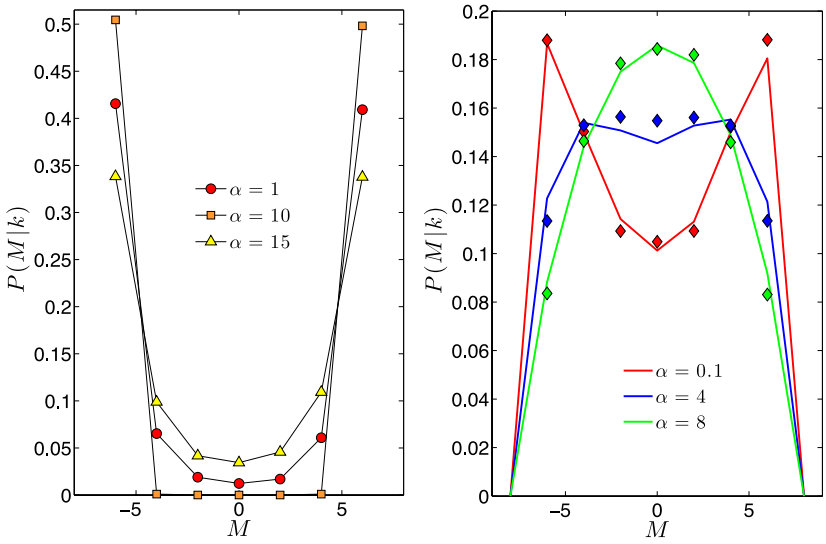

Figure 14. Left panel: overlap distribution $P(M|k)$ at zero field in the over-percolated regime, for $k = 6$, $c = 1$ and $\beta = 0.8$, and different α values (see legend). Right: the same distribution at $\beta = 0.8$, but now for $k = 6$, $c = 3$, and different α values (see legend). Note the different vertical axis scales of the two panels. Solid lines: theoretical predictions, calculated via the population dynamics method. Markers: results of measuring the overlap statistics in Monte Carlo simulations of the spin system with Hamiltonian (3), with $N = 3.10^4$ H-cells. The theory predicts that here $P(M|k, 0)$ is no longer independent of α, which we find confirmed. Note that $M \in \mathbb{Z}$, so the line segments are only guides to the eye.

6. Conclusions

The adaptive immune system consists of a large and diverse ensemble of cells and chemical messengers, such as antibodies and cytokines. Helper and suppressor T-lymphocytes (the coordinator branches) control the activity of B-lymphocytes (the effector branches) through a rich and continuous exchange of cytokines, which elicit or suppress effector actions. From a theoretical point of view, a fascinating feature of the immune system is the ability of T-lymphocytes to manage multiple B-clones at once, which is vital in defending the host from simultaneous attacks by multiple pathogens. We investigated this ability in the present study, as an emergent, collective feature of a spin-glass model, that describes the adaptive response of B-cells under the coordination of T-cells.

In particular, the focus of this paper is on the ability of the T-cells to coordinate very effectively an *extensive* number of B-soldiers, when a suitable degree of dilution in the B-T network is employed. We assumed that the number N_B of B-cells scales with the number N_T of T-cells as $N_B = \alpha N_T$, with $\alpha > 0$, and we modelled the interactions between B-cells and T-cells by means of a finitely connected bi-partite spin-glass, where each B-cell has a likelihood of being connected to a T-cell which scales as c/N_T. This is in agreement with the biological picture of highly-selective touch-interactions among B- and T-cells. The system is thermodynamically equivalent to a diluted monopartite graph $\mathcal{G}$ that describes effective interactions between T-cells, whose topological properties are shown to depend crucially on the parameters α and c. In particular, when $\alpha c^2 < 1$ the typical components in $\mathcal{G}$ are finite sized, and form cliques whose occurrence frequency decays exponentially with their size.

Each clique corresponds to a cytokine signalling pattern, and this kind of arrangement easily allows for the simultaneous recall of multiple patterns. On the other hand, when $\alpha c^2 > 1$, the effective network can exhibit a giant component, which can compromise the system's parallel processing ability.

We have analysed the operation of the system as an effective equilibrated stochastic process of interacting T-cells, by using techniques from the statistical mechanics of finitely connected spin systems. Within the replica-symmetric (RS) ansatz, we found a critical surface $T_c(\alpha, c)$ that separates two distinct phases. For $T > T_c(\alpha, c)$, the system behaves as an extensively large set of independent neural networks, each of finite size and each storing a single undiluted pattern. Here, the only source of noise are the thermal fluctuations within each subsystem. For high temperature, each subsystem behaves as a paramagnet while, at low temperature each subsystem retrieves one particular pattern (or its inverse), representing parallel retrieval (perfectly at zero temperature) of an extensive number of finite-size cytokine patterns. The regulators are able to activate and inhibit independently the whole B repertoire, as they should. In particular, each subsystem will oscillate between positive and negative signalling (with a timescale which increases with its size and only tends to infinity at zero temperature), because there is only weak ergodicity breaking. In the presence of noise (temperature), no clone can be expanded forever and there is no expansion without contraction, unless there is a persistent external stimulus (field) pinning the system to one particular strategy. This may be a key feature for the homeostatic regulation of lymphocytes numbers, as cells that are not signalled in a given time undergo anergy and apoptosis. The critical temperature becomes zero when $\alpha c^2 = 1$, i.e. $T_c(\alpha, 1/\sqrt{\alpha}) = 0 \, \forall \alpha \geqslant 0$, so for $\alpha c^2 < 1$ no transition at finite temperature away from this phase is possible.

When the load increases, i.e. when α becomes larger and we cross the transition line, overlaps among bit entries of the 'cytokine patterns' to be recalled become more and more frequent, and this gives rise to a source of cross-clonal interference which acts as an effective random field on each node. This represents an additional source of noise for the system at any finite temperature, and the only source of noise at zero temperature, and is seen to diminish the parallel processing capabilities. However, the signalling performance is found to degrade only smoothly as one enters further into the clonal cross-talk regime.

Remarkably, the high-temperature phase without clonal cross-talk is the one that gives the desired emerging behaviour of parallel retrieval, in contrast to traditional associative networks. This is due to the fact that the distribution of overlaps, which is the order function of the model, encodes both the thermal fluctuations of the overlap of the system with each pattern, *and* the fluctuations of the overlap across different patterns. Below the percolation threshold, i.e. $\alpha < 1/c^2$, where the system consists of independent subsystems, each dealing with one pattern, fluctuations of the overlap across different patterns (i.e. subsystems) are uncorrelated *even* at zero temperature (when all spins are frozen and ergodicity is broken by each subsystem), so each replica evolves independently. Increasing the temperature restores ergodicity in each subsystem, and the regime of α-values without clonal cross-talk gets wider. From physical arguments and interpretations of our formulae we expect that parallel retrieval without cross-talk is RS, whereas sequential retrieval (or parallel retrieval in the presence of cross-talk) will not be. Our predictions and results are tested against numerical simulations wherever possible, and we consistently find perfect agreement.

Finally, we are tempted to add a last note on the solvability of this model. Despite the graph $\mathcal{G}$ exhibiting many short loops, which are usually an obstacle to statistical mechanical techniques, the present spin model on $\mathcal{G}$ is found to be solvable, due to the separable nature of the effective interaction matrix. This separability allows us to unfold the effective network

into a bi-partite network $\mathcal{B}$, where loops are few or absent and factorization over sites can be achieved.

Acknowledgments

EA, AB and DT acknowledge the FIRB grant RBFR08EKEV and Sapienza Universitá di Roma for financial support. ACCC is grateful for support from the Biotechnology and Biological Sciences Research Council (BBSRC) of the United Kingdom. DT would like to thank King's College London for hospitality.

Appendix A. Simple limits

Here we work out the theory in some simple limits, which can be worked out independently, to test more complicated stages of our general calculation.

- The paramagnetic state at $\beta = 0$:

$$\lim_{\beta\to 0}\beta\overline{f} = -\lim_{N\to\infty}\lim_{n\to 0}\frac{1}{Nn}\log\sum_{\boldsymbol{\sigma}^1\ldots\boldsymbol{\sigma}^n} 1 = -\log 2. \tag{A.1}$$

  The conditioned overlap distribution at $\beta = 0$ would be

$$\begin{aligned}
P(M|\psi) &= \frac{1}{P(\psi)}\lim_{N\to\infty}\frac{1}{\alpha N}\sum_{\mu=1}^{\alpha N}\delta(\psi-\psi_\mu)\int_{-\pi}^{\pi}\frac{\mathrm{d}\phi}{2\pi}\mathrm{e}^{\mathrm{i}M\phi}2^{-N}\sum_{\boldsymbol{\sigma}}\overline{\mathrm{e}^{-\mathrm{i}\phi\sum_i\xi_i^\mu\sigma_i}}\\
&= \lim_{N\to\infty}\int_{-\pi}^{\pi}\frac{\mathrm{d}\phi}{2\pi}\mathrm{e}^{\mathrm{i}M\phi}\left(1+\frac{c}{N}[\cos(\phi)-1]\right)^N\\
&= \int_{-\pi}^{\pi}\frac{\mathrm{d}\phi}{2\pi}\mathrm{e}^{\mathrm{i}M\phi+c[\cos(\phi)-1]} = \mathrm{e}^{-c}\sum_{k\geqslant 0}\frac{c^k}{k!}\int_{-\pi}^{\pi}\frac{\mathrm{d}\phi}{2\pi}\mathrm{e}^{\mathrm{i}M\phi}\langle\mathrm{e}^{\mathrm{i}\phi\sigma}\rangle_{\sigma=\pm1}^k\\
&= \mathrm{e}^{-c}\sum_{k\geqslant 0}\frac{c^k}{k!}\langle\delta_{M,\sum_{\ell\leqslant k}\sigma_k}\rangle_{\sigma_1\ldots\sigma_k=\pm1}.
\end{aligned} \tag{A.2}$$

- The case of external fields only.
  This simply corresponds to removing the M_μ^2 terms, and gives

$$\begin{aligned}
\overline{f} &= -\lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta Nn}\log\overline{\prod_{i\alpha}\left(\sum_\sigma \mathrm{e}^{\beta\sigma\sum_{\mu=1}^{\alpha N}\psi_\mu\xi_i^\mu}\right)}\\
&= -\frac{1}{\beta}\log 2 - \lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta n}\log\overline{\cosh^n\left(\beta\sum_{\mu\leqslant\alpha N}\psi_\mu\xi^\mu\right)}\\
&= -\frac{1}{\beta}\log 2 - \lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta n}\log\int\frac{\mathrm{d}h\,\mathrm{d}\hat{h}}{2\pi}\mathrm{e}^{\mathrm{i}\hat{h}h}\cosh^n(\beta h)\,\overline{\mathrm{e}^{-\mathrm{i}\hat{h}\sum_{\mu\leqslant\alpha N}\psi_\mu\xi^\mu}}\\
&= -\frac{1}{\beta}\log 2 - \lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta n}\log\int\frac{\mathrm{d}h\,\mathrm{d}\hat{h}}{2\pi}\mathrm{e}^{\mathrm{i}\hat{h}h}\cosh^n(\beta h)\prod_{\mu=1}^{\alpha N}\left(1+\frac{c}{N}[\cos(\hat{h}\psi_\mu)-1]\right)\\
&= -\frac{1}{\beta}\log 2 - \lim_{n\to 0}\frac{1}{\beta n}\log\int\frac{\mathrm{d}h\,\mathrm{d}\hat{h}}{2\pi}\mathrm{e}^{\mathrm{i}\hat{h}h}\cosh^n(\beta h)\,\mathrm{e}^{\alpha c\int\mathrm{d}\psi P(\psi)[\cos(\hat{h}\psi)-1]}\\
&= -\frac{1}{\beta}\log 2 - \lim_{n\to 0}\frac{1}{\beta n}\log\int\frac{\mathrm{d}h\,\mathrm{d}\hat{h}}{2\pi}\mathrm{e}^{\mathrm{i}\hat{h}h+\alpha c\int\mathrm{d}\psi P(\psi)[\cos(\hat{h}\psi)-1]}
\end{aligned}$$

$$\times \{1 + n \log\cosh(\beta h) + \mathcal{O}(n^2)\}$$
$$= -\frac{1}{\beta}\log 2 - \frac{1}{\beta}\int dh W(h) \log\cosh(\beta h), \tag{A.3}$$

with the effective field distribution

$$\begin{aligned} W(h) &= \int \frac{\mathrm{d}\hat{h}}{2\pi} \mathrm{e}^{\mathrm{i}\hat{h}h + \alpha c \int \mathrm{d}\psi P(\psi)[\cos(\hat{h}\psi) - 1]} \\ &= \mathrm{e}^{-\alpha c} \sum_{k \geqslant 0} \frac{(\alpha c)^k}{k!} \int \left[\prod_{\ell \leqslant k} P(\psi_\ell)\mathrm{d}\psi_\ell\right] \int \frac{\mathrm{d}\hat{h}}{2\pi} \mathrm{e}^{\mathrm{i}\hat{h}h} \prod_{\ell \leqslant k} \cos(\hat{h}\psi_\ell) \\ &= \mathrm{e}^{-\alpha c} \sum_{k \geqslant 0} \frac{(\alpha c)^k}{k!} \int \left[\prod_{\ell \leqslant k} P(\psi_\ell)\mathrm{d}\psi_\ell\right] \left\langle \int \frac{\mathrm{d}\hat{h}}{2\pi} \mathrm{e}^{\mathrm{i}\hat{h}(h - \sum_{\ell \leqslant k} \psi_\ell \sigma_\ell)} \right\rangle_{\sigma_1 \ldots \sigma_k = \pm 1} \\ &= \sum_{k \geqslant 0} \mathrm{e}^{-\alpha c} \frac{(\alpha c)^k}{k!} \left\langle \left\langle \delta\left[h - \sum_{\ell \leqslant k} \psi_\ell \sigma_\ell\right] \right\rangle_{\psi_1 \ldots \psi_k} \right\rangle_{\sigma_1 \ldots \sigma_k = \pm 1}. \end{aligned} \tag{A.4}$$

Appendix B. Normalization of $F(\boldsymbol{\omega})$

In this appendix we derive equation (74). It follows from

$$\begin{aligned} \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega} \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma}) &= \int \frac{\mathrm{d}m\,\mathrm{d}\hat{m}}{2\pi} \mathrm{e}^{\mathrm{i}m\hat{m}} \cos(m) \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}\, \mathrm{e}^{-\mathrm{i}\hat{m}\boldsymbol{\omega}\cdot\boldsymbol{\sigma}} \\ &= \int \frac{\mathrm{d}m\,\mathrm{d}\hat{m}}{2\pi} \mathrm{e}^{\mathrm{i}m\hat{m}} \cos(m) \left[\frac{2c}{\hat{m}} \sin(\hat{m}\pi)\right]^n \\ &= \int \mathrm{d}\hat{m} \frac{\delta(\hat{m}+1) + \delta(\hat{m}-1)}{2} \left[\frac{2c}{\hat{m}} \sin(\hat{m}\pi)\right]^n = 0, \end{aligned} \tag{B.1}$$

where we isolated $\boldsymbol{\sigma}\cdot\boldsymbol{\omega}$ via $1 = (2\pi)^{-1}\int \mathrm{d}m\,\mathrm{d}\hat{m}\,\mathrm{e}^{\mathrm{i}m\hat{m} - i\hat{m}\boldsymbol{\omega}\cdot\boldsymbol{\sigma}}$ and used

$$\begin{aligned} \int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}\, \mathrm{e}^{-\mathrm{i}\hat{m}\boldsymbol{\omega}\cdot\boldsymbol{\sigma}} &= \prod_{\alpha=1}^{n} \int_{-\pi}^{\pi} \mathrm{d}\omega_\alpha\, \mathrm{e}^{-\mathrm{i}\hat{m}\omega_\alpha\sigma_\alpha} = \prod_{\alpha=1}^{n} \left(2\int_0^{\pi} \mathrm{d}\omega_\alpha \cos(\hat{m}\omega_\alpha\sigma_\alpha)\right) \\ &= \prod_{\alpha=1}^{n} \left(\frac{2c\sigma^\alpha}{\hat{m}} \sin(\hat{m}\pi\sigma^\alpha)\right) = \left[\frac{2c}{\hat{m}} \sin(\hat{m}\pi)\right]^n. \end{aligned} \tag{B.2}$$

Appendix C. Continuous RS phase transitions via route I

Here we derive the equation for the continuous phase transitions in the absence of external fields, i.e. for $P(\psi) = \delta(\psi)$, away from solution (99). At the transition, the function $D_0(\omega|\beta)$, which we will denote simply as $D(\omega|\beta)$, still satisfies (75). Continuous bifurcations away from (99) can be identified via a Guzai (or functional moment) expansion [49]. We transform

$$\pi(\omega) \to \cos(\omega) + \Delta(\omega), \tag{C.1}$$

with $f_k(\{\pi_1, \ldots, \pi_\ell\}) \to \tilde{f}_k(\{\Delta_1, \ldots, \Delta_k\})$, $W[\{\pi\}] \to \tilde{W}[\{\Delta\}]$, and $\tilde{W}[\{\Delta\}] = 0$ as soon as $\int \mathrm{d}\omega\Delta(\omega) \neq 0$ (because $\int \mathrm{d}\omega\,\pi(\omega) = 1$), and $\lambda(\theta|W) \to \tilde{\lambda}(\theta|\tilde{W})$. We expand our equations in powers of the functional moments $\varrho(\omega_1, \ldots, \omega_r) = \int\{\mathrm{d}\Delta\}\tilde{W}[\{\Delta\}]\Delta(\omega_1)\ldots\Delta(\omega_r)$. One assumes that close to the transition there exists some small ϵ such that $\varrho(\omega_1, \ldots, \omega_r) = \mathcal{O}(\epsilon^r)$.

If the lowest bifurcation is of order ϵ^1, we obtain, upon multiplying (89) by Δ and subsequently integrating over Δ:

$$\varrho(\omega) = \int \{\mathrm{d}\Delta\}\Delta(\omega) \int \frac{\mathrm{d}\theta}{2\pi}\tilde{\lambda}(\theta|\tilde{W}) \prod_{\omega} \delta[\Delta(\omega) + \cos(\omega) - \cos(\omega-\theta)]$$
$$= \int \frac{\mathrm{d}\theta}{2\pi}\tilde{\lambda}(\theta|\tilde{W})[\cos(\omega-\theta) - \cos(\omega)] = \cos(\omega)\int \frac{\mathrm{d}\theta}{2\pi}\tilde{\lambda}(\theta|\tilde{W})[\cos\theta - 1], \tag{C.2}$$

where we used the invariance under $\theta \to -\theta$ of

$$\tilde{\lambda}(\theta|\tilde{W}) = \sum_{m\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}m\theta + c\alpha\sum_{k\geqslant 0}\frac{c^k \mathrm{e}^{-c}}{k!}\int\prod_{\ell=1}^{k}[\{\mathrm{d}\Delta_\ell\}\tilde{W}[\{\Delta_\ell\}]]\{\cos[m\arctan\tilde{f}_k(\{\Delta_1,\ldots,\Delta_k\})]-1\}}. \tag{C.3}$$

The solution of (C.2) is clearly $\varrho(\omega) = \phi\cos(\omega)$, with

$$\phi = \int \frac{\mathrm{d}\theta}{2\pi}\tilde{\lambda}(\theta|\tilde{W})[\cos(\theta) - 1], \tag{C.4}$$

which we need to evaluate further by expanding $\tilde{\lambda}(\theta|\tilde{W})$ for small ϵ. Conversely, if the lowest bifurcating order is ϵ^2 one must focus on

$$\varrho(\omega_1,\omega_2) = \int \{\mathrm{d}\Delta\}\Delta(\omega_1)\Delta(\omega_2)\int \frac{\mathrm{d}\theta}{2\pi}\tilde{\lambda}(\theta|\tilde{W})\prod_{\omega}\delta[\Delta(\omega) - \cos(\omega) - \cos(\omega-\theta)]$$
$$= \cos(\omega_1)\cos(\omega_2)\int \frac{\mathrm{d}\theta}{2\pi}\tilde{\lambda}(\theta|\tilde{W})[\cos(\theta)-1]^2$$
$$+ \sin(\omega_1)\sin(\omega_2)\int \frac{\mathrm{d}\theta}{2\pi}\tilde{\lambda}(\theta|\tilde{W})\sin^2(\theta). \tag{C.5}$$

We first inspect (C.4). Transforming each π_ℓ in (98) according to (C.1), we have

$$\prod_{\ell=1}^{k}\pi_\ell(\omega) = \prod_{\ell=1}^{k}[\cos(\omega) + \Delta_\ell(\omega)] = \cos^k(\omega)\left[1 + \sum_{\ell=1}^{k}\frac{\Delta_\ell(\omega)}{\cos(\omega)}\right] + \mathcal{O}(\Delta^2). \tag{C.6}$$

Inserting this result into (98) and using the properties (75) allows us to expand $\tilde{f}_k(\{\Delta_1,\ldots,\Delta_k\})$:

$$\tilde{f}_k(\{\Delta_1,\ldots,\Delta_k\}) = \frac{\sum_{\ell=1}^{k}\int_{-\pi}^{\pi}\mathrm{d}\omega\sin(\omega)\cos^{k-1}(\omega)\Delta_\ell(\omega)D(\omega|\beta)}{\int_{-\pi}^{\pi}\mathrm{d}\omega\cos^{k+1}(\omega)D(\omega|\beta)} + \mathcal{O}(\Delta^2). \tag{C.7}$$

We substitute the above into (C.3) and expand $\cos(m\arctan(x)) = 1 - \frac{1}{2}m^2x^2 + \mathcal{O}(x^4)$. Upon introducing

$$\mathcal{I}_k = \int\prod_{\ell=1}^{k}[\{\mathrm{d}\Delta_\ell\}\tilde{W}[\{\Delta_\ell\}]]\left[\sum_{s=1}^{k}\int_{-\pi}^{\pi}\mathrm{d}\omega\sin(\omega)\cos^{k-1}(\omega)\Delta_s(\omega)D(\omega|\beta)\right]^2, \tag{C.8}$$

$$A_k = \int_{-\pi}^{\pi}\mathrm{d}\omega\cos^{k+1}(\omega)D(\omega|\beta), \tag{C.9}$$

we see that $\mathcal{I}_k = \mathcal{O}(\epsilon^2)$, so we can now expand $\tilde{\lambda}(\theta|\tilde{W})$ as

$$\tilde{\lambda}(\theta|\tilde{W}) = \sum_{m\in\mathbb{Z}}\exp\left[\mathrm{i}m\theta - \frac{c\alpha}{2}m^2\sum_{k\geqslant 0}\frac{\mathrm{e}^{-c}c^k}{k!}\frac{\mathcal{I}_k}{A_k^2} + \mathcal{O}(\epsilon^4)\right]$$
$$= \sum_{m\in\mathbb{Z}}\mathrm{e}^{\mathrm{i}m\theta}\left[1 - \frac{c\alpha}{2}m^2\sum_{k\geqslant 0}\frac{\mathrm{e}^{-c}c^k}{k!}\frac{\mathcal{I}_k}{A_k^2} + \mathcal{O}(\epsilon^4)\right]$$
$$= 2\pi\delta(\theta) + \pi\alpha c\delta''(\theta)\sum_{k\geqslant 0}\frac{\mathrm{e}^{-c}c^k}{k!}\frac{\mathcal{I}_k}{A_k^2} + \mathcal{O}(\epsilon^4). \tag{C.10}$$

Next we need to work out the factors $\mathcal{I}_k$. Using the functional moment definition $\varrho(\omega_1,\ldots,\omega_r)=\int\{\mathrm{d}\Delta\}\tilde{W}[\{\Delta\}]\Delta(\omega_1)\ldots\Delta(\omega_r)$, one may write

$$\begin{aligned}
&\int\prod_{\ell=1}^{k}[\{\mathrm{d}\Delta_\ell\}\tilde{W}[\{\Delta_\ell\}]]\sum_{r,s=1}^{k}\Delta_r(\omega')\Delta_s(\omega'')\\
&=\sum_r\int\{\mathrm{d}\Delta_r\}\tilde{W}[\{\Delta_r\}]\Delta_r(\omega')\Delta_r(\omega'')+\sum_{r\neq s}\int\{\mathrm{d}\Delta_r\}\{\mathrm{d}\Delta_s\}\tilde{W}[\{\Delta_r\}]\tilde{W}[\{\Delta_s\}]\Delta_r(\omega')\Delta_s(\omega'')\\
&=k\varrho(\omega',\omega'')+k(k-1)\varrho(\omega')\varrho(\omega''). \qquad \text{(C.11)}
\end{aligned}$$

This allows us to work out (C.8) further:

$$\begin{aligned}
\mathcal{I}_k&=k\int_{-\pi}^{\pi}\mathrm{d}\omega'\mathrm{d}\omega''\sin(\omega')\cos^{k-1}(\omega')D(\omega'|\beta)\sin(\omega'')\cos^{k-1}(\omega'')D(\omega''|\beta)\psi(\omega',\omega'')\\
&\quad+k(k-1)\left[\int_{-\pi}^{\pi}\mathrm{d}\omega'\sin(\omega')\cos^{k-1}(\omega')D(\omega'|\beta)\psi(\omega')\right]^2\\
&=k\int_{-\pi}^{\pi}\mathrm{d}\omega'\mathrm{d}\omega''D(\omega'|\beta)D(\omega''|\beta)\sin(\omega')\cos^{k-1}(\omega')\psi(\omega',\omega'')\sin(\omega'')\cos^{k-1}(\omega''),
\end{aligned} \qquad \text{(C.12)}$$

where in the last equality we have used the symmetry of $D(\omega|\beta)$ and $\varrho(\omega)=\phi\cos(\omega)$. Inserting this last expression in (C.10) and shifting the summation index $k\to k+1$ then leads to

$$\tilde{\lambda}(\theta|\tilde{W})=2\pi\delta(\theta)+\pi\alpha c^2\delta''(\theta)S(\{\varrho\})+\mathcal{O}(\epsilon^4), \qquad \text{(C.13)}$$

$$\begin{aligned}
S(\{\varrho\})&=\sum_{k\geqslant 0}\frac{\mathrm{e}^{-c}c^k}{k!}\\
&\quad\times\frac{\int_{-\pi}^{\pi}\mathrm{d}\omega'\mathrm{d}\omega''D(\omega'|\beta)D(\omega''|\beta)\sin(\omega')\cos^{k}(\omega')\varrho(\omega',\omega'')\sin(\omega'')\cos^{k}(\omega'')}{\left[\int_{-\pi}^{\pi}\mathrm{d}\omega D(\omega|\beta)\cos^{k+2}(\omega)\right]^2}.
\end{aligned} \qquad \text{(C.14)}$$

To make further progress we need to calculate $\varrho(\omega',\omega'')$. We can first simplify (C.5) using (C.4), giving

$$\varrho(\omega_1,\omega_2)=\phi'\sin(\omega_1)\sin(\omega_2)-(2\phi+\phi')\cos(\omega_1)\cos(\omega_2), \qquad \text{(C.15)}$$

where we defined

$$\phi'=\int_{-\pi}^{\pi}\frac{\mathrm{d}\theta}{2\pi}\tilde{\lambda}(\theta|\tilde{W})\sin^2(\theta). \qquad \text{(C.16)}$$

With this we can simplify (C.14) to

$$S(\{\varrho\})=\phi'\sum_{k\geqslant 0}\frac{\mathrm{e}^{-c}c^k}{k!}\frac{\left[\int_{-\pi}^{\pi}\mathrm{d}\omega D(\omega|\beta)\sin^2(\omega)\cos^{k}(\omega)\right]^2}{\left[\int_{-\pi}^{\pi}\mathrm{d}\omega D(\omega|\beta)\cos^{k+2}(\omega)\right]^2}. \qquad \text{(C.17)}$$

Together with (C.13), this allows us to establish equations from which the two amplitudes ϕ and ϕ' can be solved, by substitution into (C.4) and (C.16). This results in, after integration by parts over θ:

$$\begin{aligned}
\phi&=\frac{1}{2}\alpha c^2S(\{\varrho\})\int_{-\pi}^{\pi}\mathrm{d}\theta[\cos(\theta)-1]\delta''(\theta)+\mathcal{O}(\epsilon^4)=-\frac{1}{2}\alpha c^2S(\{\varrho\})+\mathcal{O}(\epsilon^4)\\
&=-\frac{1}{2}\alpha c^2\phi'\sum_{k\geqslant 0}\frac{\mathrm{e}^{-c}c^k}{k!}\frac{\left[\int_{-\pi}^{\pi}\mathrm{d}\omega D(\omega|\beta)\sin^2(\omega)\cos^{k}(\omega)\right]^2}{\left[\int_{-\pi}^{\pi}\mathrm{d}\omega D(\omega|\beta)\cos^{k+2}(\omega)\right]^2}+\mathcal{O}(\epsilon^4)
\end{aligned} \qquad \text{(C.18)}$$

$$\phi' = \frac{1}{2}\alpha c^2 S(\{\varrho\}) \int_{-\pi}^{\pi} \mathrm{d}\theta \sin^2(\theta)\delta''(\theta) + \mathcal{O}(\epsilon^4)$$
$$= \alpha c^2 \phi' \sum_{k\geqslant 0} \frac{\mathrm{e}^{-c}c^k}{k!} \frac{\left[\int_{-\pi}^{\pi} \mathrm{d}\omega D(\omega|\beta)\sin^2(\omega)\cos^k(\omega)\right]^2}{\left[\int_{-\pi}^{\pi} \mathrm{d}\omega D(\omega|\beta)\cos^{k+2}(\omega)\right]^2} + \mathcal{O}(\epsilon^4). \tag{C.19}$$

Since $\phi' = 0$ immediately implies that $\phi = 0$, the only possible continuous bifurcation must be the first instance where $\phi' \neq 0$. According to the above equation this $\mathcal{O}(\epsilon^2)$ bifurcation happens when

$$1 = \alpha c^2 \sum_{k\geqslant 0} \frac{\mathrm{e}^{-c}c^k}{k!} \left[\frac{\int_{-\pi}^{\pi} \mathrm{d}\omega \sin^2(\omega)\cos^k(\omega)D(\omega|\beta)}{\int_{-\pi}^{\pi} \mathrm{d}\omega \cos^{k+2}(\omega)D(\omega|\beta)}\right]^2, \tag{C.20}$$

with $D(\omega|\beta) = (2\pi)^{-1}\sum_{m\in\mathbb{Z}}\cos(m\omega)\mathrm{e}^{\beta m^2/2c}$. Equation (C.20) defines the transition point, where the system will leave the state (99). The RHS of (C.20) obeys $\lim_{\beta\to 0}\mathrm{RHS} = 0$. In Appendix G we show that $\lim_{\beta\to\infty}\mathrm{RHS} = \alpha c^2$, so a transition at finite temperature $T_c = \beta_c^{-1} > 0$ exists in a new state with $W[\{\pi\}] \neq \prod_\omega \delta[\pi(\omega) - \cos(\omega)]$ as soon as $\alpha c^2 > 1$. The critical temperature becomes zero when $\alpha c^2 = 1$.

Appendix D. Saddle-point equations in terms of $L(\sigma)$

Here we derive equation (118), starting from the definition (112) and relation (113):

$$L(\boldsymbol{\sigma}) = \alpha c \left\langle \frac{\int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\omega}\cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})Q(\boldsymbol{\omega})\sum_{\mathbf{M}}\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\sum_\alpha \chi(M_\alpha,\psi)}}{\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}Q(\boldsymbol{\omega})\sum_{\mathbf{M}}\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}+\sum_\alpha\chi(M_\alpha,\psi)}}\right\rangle_\psi$$
$$= \alpha c \left\langle \frac{\int \mathrm{d}\boldsymbol{\omega}\cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})\sum_{\mathbf{M}'}\tilde{Q}(\mathbf{M}')\sum_{\mathbf{M}}\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot(\mathbf{M}-\mathbf{M}')+\sum_\alpha\chi(M_\alpha,\psi)}}{\int \mathrm{d}\boldsymbol{\omega}\sum_{\mathbf{M}'}\tilde{Q}(\mathbf{M}')\sum_{\mathbf{M}}\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot(\mathbf{M}-\mathbf{M}')+\sum_\alpha\chi(M_\alpha,\psi)}}\right\rangle_\psi. \tag{D.1}$$

We can then work out the integrals

$$\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}\cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})\,\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot(\mathbf{M}-\mathbf{M}')} = \frac{1}{2}\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\omega}(\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\boldsymbol{\sigma}}+\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\boldsymbol{\sigma}})\,\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot(\mathbf{M}-\mathbf{M}')}$$
$$= \pi(\delta_{\mathbf{M}',\mathbf{M}+\boldsymbol{\sigma}} + \delta_{\mathbf{M}',\mathbf{M}-\boldsymbol{\sigma}}), \tag{D.2}$$

and substituting into (D.1) gives

$$L(\boldsymbol{\sigma}) = \frac{1}{2}\alpha c\left\langle \frac{\sum_{\mathbf{M}}[\tilde{Q}(\mathbf{M}+\boldsymbol{\sigma})\,\mathrm{e}^{\beta\sum_\alpha\chi(M_\alpha,\psi)} + \tilde{Q}(\mathbf{M}-\boldsymbol{\sigma})\,\mathrm{e}^{\beta\sum_\alpha\chi(M_\alpha,\psi)}]}{\sum_{\mathbf{M}}\tilde{Q}(\mathbf{M})\,\mathrm{e}^{\sum_\alpha\chi(M_\alpha,\psi)}}\right\rangle_\psi$$
$$= \frac{1}{2}\alpha c\left\langle \frac{\sum_{\mathbf{M}}\tilde{Q}(\mathbf{M})[\mathrm{e}^{\beta\sum_\alpha\chi(M_\alpha-\sigma^\alpha,\psi)} + \mathrm{e}^{\beta\sum_\alpha\chi(M_\alpha+\sigma^\alpha,\psi)}]}{\sum_{\mathbf{M}}\tilde{Q}(\mathbf{M})\,\mathrm{e}^{\sum_\alpha\chi(M_\alpha,\psi)}}\right\rangle_\psi$$
$$= c\alpha\left\langle \mathrm{e}^{\frac{\beta n}{2c}}\frac{\sum_{\mathbf{M}}\tilde{Q}(\mathbf{M})\,\mathrm{e}^{\beta\sum_\alpha\chi(M_\alpha,\psi)}\cosh\left[\beta\left(\frac{1}{c}\mathbf{M}\cdot\boldsymbol{\sigma}+\psi\sum_\alpha\sigma^\alpha\right)\right]}{\sum_{\mathbf{M}}\tilde{Q}(\mathbf{M})\,\mathrm{e}^{\beta\sum_\alpha\chi(M_\alpha,\psi)}}\right\rangle_\psi. \tag{D.3}$$

Appendix E. Simple limits to test the replica theory

Here we inspect several simple limits to test our results for the overlap distribution and the free energy.

- *Infinite temperature.* Using $\lim_{\beta\to 0} L(\boldsymbol{\sigma}) = \alpha c$ and $\sum_{\mathbf{M}} \tilde{Q}(\mathbf{M}) = \mathrm{e}^{c}$ in (119) we immediately find the correct free energy

$$\lim_{\beta\to 0} \beta \overline{f}_{\mathrm{RSB}} = -\lim_{n\to 0} \frac{1}{n} \left\{ \log \sum_{\boldsymbol{\sigma}} 1 \right\} = -\log 2. \tag{E.1}$$

Moreover, from (117) we can extract

$$\begin{aligned}
\lim_{\beta\to 0} \tilde{Q}(\mathbf{M}) &= \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \cos(\boldsymbol{\omega}\cdot\mathbf{M}) \mathrm{e}^{c 2^{-n} \sum_{\boldsymbol{\sigma}} \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})} \\
&= \sum_{k\geqslant 0} \frac{c^k}{k!} 2^{-nk} \sum_{\boldsymbol{\sigma}^1 \ldots \boldsymbol{\sigma}^k} \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \cos(\boldsymbol{\omega}\cdot\mathbf{M}) \prod_{\ell\leqslant k} \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma}^{\ell}) \\
&= \sum_{k\geqslant 0} \frac{c^k}{k!} 2^{-nk} \sum_{\boldsymbol{\sigma}^1 \ldots \boldsymbol{\sigma}^k} \delta_{\mathbf{M}, \sum_{\ell\leqslant k} \boldsymbol{\sigma}^{\ell}} = \sum_{k\geqslant 0} \frac{c^k}{k!} \prod_{\alpha=1}^{n} \langle \delta_{M_\alpha, \sum_{\ell\leqslant k} \sigma_\ell} \rangle_{\sigma_1 \ldots \sigma_k = \pm 1}.
\end{aligned} \tag{E.2}$$

Hence, it now follows from (120) that

$$\begin{aligned}
\lim_{\beta\to 0} P(M|\psi) &= \lim_{n\to 0} \frac{1}{n} \sum_{\gamma=1}^{n} \frac{\sum_{\mathbf{M}\in\mathbb{Z}^n} \sum_{k\geqslant 0} \frac{c^k}{k!} \prod_{\alpha=1}^{n} \langle \delta_{M_\alpha, \sum_{\ell\leqslant k} \sigma_\ell} \rangle_{\sigma_1 \ldots \sigma_k = \pm 1} \delta_{M, M_\gamma}}{\sum_{\mathbf{M}\in\mathbb{Z}^n} \sum_{k\geqslant 0} \frac{c^k}{k!} \prod_{\alpha=1}^{n} \langle \delta_{M_\alpha, \sum_{\ell\leqslant k} \sigma_\ell} \rangle_{\sigma_1 \ldots \sigma_k = \pm 1}} \\
&= \mathrm{e}^{-c} \sum_{k\geqslant 0} \frac{c^k}{k!} \langle \delta_{M, \sum_{\ell\leqslant k} \sigma_\ell} \rangle_{\sigma_1 \ldots \sigma_k = \pm 1}.
\end{aligned} \tag{E.3}$$

This coincides with our RS expression, as it should since at high temperature the RS ansatz is exact.

- *External fields only.* In the case of having only external fields we simply remove all terms that come from the interaction energy in (D.3), obtaining

$$L(\boldsymbol{\sigma}) = \alpha c \left\langle \cosh\left[\beta\psi \sum_{\alpha} \sigma_\alpha \right] \right\rangle_{\psi}. \tag{E.4}$$

Inserting this into (117), and introducing the normalized measure

$$\lambda(\boldsymbol{\sigma}) = \frac{\mathrm{e}^{\alpha c \langle \cosh[\beta\psi \sum_\alpha \sigma_\alpha] \rangle_\psi}}{\sum_{\boldsymbol{\sigma}'} \mathrm{e}^{\alpha c \langle \cosh[\beta\psi \sum_\alpha \sigma'_\alpha] \rangle_\psi}}, \tag{E.5}$$

we get

$$\begin{aligned}
\tilde{Q}(\mathbf{M}) &= \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \cos(\boldsymbol{\omega}\cdot\mathbf{M})\, \mathrm{e}^{c \sum_{\boldsymbol{\sigma}} \lambda(\boldsymbol{\sigma}) \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})} \\
&= \sum_{k\geqslant 0} \frac{c^k}{k!} \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}} \left(\sum_{\boldsymbol{\sigma}} \lambda(\boldsymbol{\sigma})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{\omega}\cdot\boldsymbol{\sigma}} \right)^k \\
&= \sum_{k\geqslant 0} \frac{c^k}{k!} \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{M}} \sum_{\boldsymbol{\sigma}_1 \ldots \boldsymbol{\sigma}_k} \left[\prod_{\ell=1}^{k} \lambda(\boldsymbol{\sigma}_\ell) \right] \mathrm{e}^{-\mathrm{i}\boldsymbol{\omega}\cdot\sum_{\ell\leqslant k} \boldsymbol{\sigma}_\ell} \\
&= \sum_{k\geqslant 0} \frac{c^k}{k!} \sum_{\boldsymbol{\sigma}_1 \ldots \boldsymbol{\sigma}_k} \left[\prod_{\ell=1}^{k} \lambda(\boldsymbol{\sigma}_\ell) \right] \delta_{\mathbf{M}, \sum_{\ell\leqslant k} \boldsymbol{\sigma}_\ell}.
\end{aligned} \tag{E.6}$$

This then gives for the free energy, upon removing the interaction energy:

$$
\begin{aligned}
\overline{f}_{\rm RSB} =& -\lim_{n\to 0}\frac{1}{\beta n}\left\{\alpha\left\langle \log\sum_{\mathbf{M}\in\mathbb{Z}^n}\tilde{Q}(\mathbf{M})\,\mathrm{e}^{\beta\psi\sum_\alpha M_\alpha}\right\rangle_\psi\right.\\
&\left.+\log\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\alpha c[\langle\cosh[\beta\psi\sum_\alpha\sigma_\alpha]\rangle_\psi-1]}-\alpha c\sum_{\boldsymbol{\sigma}}\lambda(\boldsymbol{\sigma})\langle\cosh[\beta\psi\sum_\alpha\sigma_\alpha]\rangle_\psi\right\}\\
=& -\lim_{n\to 0}\frac{1}{\beta n}\left\{\alpha\left\langle \log\left[\sum_{k\geqslant 0}\frac{c^k}{k!}\sum_{\boldsymbol{\sigma}_1\ldots\boldsymbol{\sigma}_k}\left[\prod_{\ell=1}^k\lambda(\boldsymbol{\sigma}_\ell)\right]\sum_{\mathbf{M}\in\mathbb{Z}^n}\delta_{\mathbf{M},\sum_{\ell\leqslant k}\boldsymbol{\sigma}_\ell}\mathrm{e}^{\beta\psi\sum_\alpha M_\alpha}\right]\right\rangle_\psi\right.\\
&\left.+\log\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\alpha c[\langle\cosh[\beta\psi\sum_\alpha\sigma_\alpha]\rangle_\psi-1]}-\alpha c\sum_{\boldsymbol{\sigma}}\lambda(\boldsymbol{\sigma})\left\langle\cosh\left[\beta\psi\sum_\alpha\sigma_\alpha\right]\right\rangle_\psi\right\}\\
=& -\lim_{n\to 0}\frac{1}{\beta n}\left\{\alpha\left\langle \log\left[\sum_{k\geqslant 0}\frac{c^k}{k!}\left(\sum_{\boldsymbol{\sigma}}\lambda(\boldsymbol{\sigma})\,\mathrm{e}^{\beta\psi\sum_\alpha\sigma_\alpha}\right)^k\right]\right\rangle_\psi\right.\\
&\left.+\log\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\alpha c[\langle\cosh[\beta\psi\sum_\alpha\sigma_\alpha]\rangle_\psi-1]}-\alpha c\sum_{\boldsymbol{\sigma}}\lambda(\boldsymbol{\sigma})\left\langle\cosh\left[\beta\psi\sum_\alpha\sigma_\alpha\right]\right\rangle_\psi\right\}\\
=& -\lim_{n\to 0}\frac{1}{\beta n}\left\{\alpha c\left\langle\sum_{\boldsymbol{\sigma}}\lambda(\boldsymbol{\sigma})\,\mathrm{e}^{\beta\psi\sum_\alpha\sigma_\alpha}\right\rangle_\psi-\alpha c\right.\\
&\left.+\log\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\alpha c[\langle\cosh[\beta\psi\sum_\alpha\sigma_\alpha]\rangle_\psi-1]}-\alpha c\sum_{\boldsymbol{\sigma}}\lambda(\boldsymbol{\sigma})\left\langle\cosh\left[\beta\psi\sum_\alpha\sigma_\alpha\right]\right\rangle_\psi\right\}\\
=& -\lim_{n\to 0}\frac{1}{\beta n}\left\{\alpha c\sum_{\boldsymbol{\sigma}}\lambda(\boldsymbol{\sigma})\left\langle\cosh\left[\beta\psi\sum_\alpha\sigma_\alpha\right]\right\rangle_\psi-\alpha c\right.\\
&\left.+\log\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\alpha c[\langle\cosh[\beta\psi\sum_\alpha\sigma_\alpha]\rangle_\psi-1]}-\alpha c\sum_{\boldsymbol{\sigma}}\lambda(\boldsymbol{\sigma})\left\langle\cosh\left[\beta\psi\sum_\alpha\sigma_\alpha\right]\right\rangle_\psi\right\}\\
=& -\lim_{n\to 0}\frac{1}{\beta n}\left\{\log\sum_{\boldsymbol{\sigma}}\mathrm{e}^{\alpha c[\langle\cosh[\beta\psi\sum_\alpha\sigma_\alpha]\rangle_\psi-1]}\right\},
\end{aligned}
\tag{E.7}
$$

where in the penultimate step we used $\lambda(\boldsymbol{\sigma})=\lambda(-\boldsymbol{\sigma})$. We next use the following replica identity, which is proved via Taylor expansion of even non-negative analytical functions $F(x)$ that have $F(0)=1$:

$$
\lim_{n\to 0}n^{-1}\log\left\langle F\left(\sum_{\alpha=1}^n\sigma_\alpha\right)\right\rangle_{\sigma_1\ldots\sigma_n=\pm 1}=\sum_{k>0}\frac{F^{(k)}(0)}{k!}\left(\frac{\mathrm{d}^k}{\mathrm{d}x^k}\log\cosh(x)\right)\bigg|_{x=0}. \tag{E.8}
$$

Application to the function $F(z)=\exp[\alpha c\langle\cosh[\beta\psi z]\rangle_\psi-\alpha c]$ gives

$$
\begin{aligned}
\overline{f}_{\rm RSB} &= -\frac{1}{\beta}\log 2-\frac{\mathrm{e}^{-\alpha c}}{\beta}\lim_{x,z\to 0}\sum_{k>0}\frac{1}{k!}\left(\frac{\mathrm{d}^k}{\mathrm{d}x^k}\log\cosh(x)\right)\frac{\mathrm{d}^k}{\mathrm{d}z^k}\,\mathrm{e}^{\alpha c\langle\cosh(\beta\psi z)\rangle_\psi}\\
&= -\frac{1}{\beta}\log 2-\frac{\mathrm{e}^{-\alpha c}}{\beta}\sum_{\ell\geqslant 0}\frac{(\alpha c)^\ell}{\ell!}\lim_{x,z\to 0}\sum_{k>0}\frac{1}{k!}\left(\frac{\mathrm{d}^k}{\mathrm{d}x^k}\log\cosh(x)\right)\frac{\mathrm{d}^k}{\mathrm{d}z^k}\langle\cosh(\beta\psi z)\rangle_\psi^\ell\\
&= -\frac{1}{\beta}\log 2-\frac{\mathrm{e}^{-\alpha c}}{\beta}\sum_{\ell\geqslant 0}\frac{(\alpha c)^\ell}{\ell!}\lim_{x,z\to 0}\sum_{k>0}\frac{1}{k!}
\end{aligned}
$$

$$
\begin{aligned}
&\times \left(\frac{\mathrm{d}^k}{\mathrm{d}x^k}\log\cosh(x)\right)\frac{\mathrm{d}^k}{\mathrm{d}z^k}\langle\langle e^{\beta\psi\sum_{r\leqslant\ell}\sigma_r z_r}\rangle_{\psi_1\ldots\psi_\ell}\rangle_{\sigma_1\ldots\sigma_\ell=\pm1}\\
&= -\frac{1}{\beta}\log 2 - \frac{\mathrm{e}^{-\alpha c}}{\beta}\sum_{\ell\geqslant 0}\frac{(\alpha c)^\ell}{\ell!}\\
&\times \left\langle\left\langle \sum_{k>0}\frac{1}{k!}\left(\lim_{x\to 0}\frac{\mathrm{d}^k}{\mathrm{d}x^k}\log\cosh(x)\right)\left(\beta\sum_{r\leqslant\ell}\sigma_r\psi_r\right)^k\right\rangle_{\psi_1\ldots\psi_\ell}\right\rangle_{\sigma_1\ldots\sigma_\ell=\pm1}\\
&= -\frac{1}{\beta}\log 2 - \frac{\mathrm{e}^{-\alpha c}}{\beta}\sum_{\ell\geqslant 0}\frac{(\alpha c)^\ell}{\ell!}\left\langle\left\langle \log\cosh\left(\beta\sum_{r\leqslant\ell}\sigma_r\psi_r\right)\right\rangle_{\psi_1\ldots\psi_\ell}\right\rangle_{\sigma_1\ldots\sigma_\ell=\pm1}\\
&= -\frac{1}{\beta}\log 2 - \frac{1}{\beta}\int \mathrm{d}h W(h)\log\cosh(\beta h),
\end{aligned}
\tag{E.9}
$$

with

$$
W(h) = \sum_{k\geqslant 0} e^{-\alpha c}\frac{(\alpha c)^k}{k!}\left\langle\left\langle \delta\left[h - \sum_{\ell\leqslant k}\psi_\ell\sigma_\ell)\right]\right\rangle_{\psi_1\ldots\psi_k}\right\rangle_{\sigma_1\ldots\sigma_k=\pm1}. \tag{E.10}
$$

This correctly recovers the solution of external fields only.

Appendix F. Derivation of RS equations via route II

The RS ansatz converts the saddle-point equation (118) into

$$
\begin{aligned}
&\int \mathrm{d}h W(h)\,\mathrm{e}^{\beta h\sum_\alpha\sigma_\alpha}\\
&\quad= \mathrm{e}^{\beta n/2c}\left\langle\left\langle\int\{\mathrm{d}\pi\}W[\pi]\prod_\alpha\left(\sum_M\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M+\tau(\psi+M/c)\sigma_\alpha)}\right)\right\rangle_\psi\right\rangle_{\tau=\pm1}\\
&\quad= \mathrm{e}^{\beta n/2c}\left\langle\left\langle\int\{\mathrm{d}\pi\}W[\pi]\left(\sum_M\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M+\tau(\psi+M/c))}\right)^{\frac{1}{2}n+\frac{1}{2}\sum_\alpha\sigma_\alpha}\right.\right.\\
&\qquad\left.\left.\times\left(\sum_M\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M-\tau(\psi+M/c))}\right)^{\frac{1}{2}n-\frac{1}{2}\sum_\alpha\sigma_\alpha}\right\rangle_\psi\right\rangle_{\tau=\pm1}\\
&\quad= \mathrm{e}^{\beta n/2c}\left\langle\left\langle\int\{d\pi\}W[\pi]\left(\frac{\sum_M\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M+\tau(\psi+M/c))}}{\sum_M\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M-\tau(\psi+M/c))}}\right)^{\frac{1}{2}\sum_\alpha\sigma_\alpha}\right\rangle_\psi\right\rangle_{\tau=\pm1}\\
&\quad= \mathrm{e}^{\beta n/2c}\int\mathrm{d}h\,\mathrm{e}^{\beta h\sum_\alpha\sigma_\alpha}\left\langle\left\langle\int\{\mathrm{d}\pi\}W[\pi]\delta\right.\right.\\
&\qquad\left.\left.\times\left[h-\frac{1}{2\beta}\log\left(\frac{\sum_M\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M+\tau(\psi+M/c))}}{\sum_M\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M-\tau(\psi+M/c))}}\right)\right]\right\rangle_\psi\right\rangle_{\tau=\pm1}.
\end{aligned}
\tag{F.1}
$$

We conclude after sending $n\to 0$ that

$$
W(h) = \left\langle\left\langle\int\{\mathrm{d}\pi\}W[\pi]\delta\left[h-\frac{1}{2\beta}\log\left(\frac{\sum_M\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M+\tau(\psi+M/c))}}{\sum_M\pi(M)\,\mathrm{e}^{\beta(M^2/2c+\psi M-\tau(\psi+M/c))}}\right)\right]\right\rangle_\psi\right\rangle_{\tau=\pm1}. \tag{F.2}
$$

$W(h)$ is indeed symmetric. Next we turn to equation (117), where we require quantities of the form

$$\varrho(\boldsymbol{\omega}) = \sum_{\boldsymbol{\sigma}} \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})\, \mathrm{e}^{L(\boldsymbol{\sigma})} = \sum_{\boldsymbol{\sigma}} \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma})\, \mathrm{e}^{\alpha c \int \mathrm{d}h W(h)\, \mathrm{e}^{\beta h \sum_\alpha \sigma_\alpha}}. \tag{F.3}$$

In fact we will need only the ratio $\varrho(\boldsymbol{\omega})/\varrho(\mathbf{0})$. We note that

$$\varrho(\mathbf{0}) = 2^n \sum_{k\geqslant 0} \frac{(\alpha c)^k}{k!} \int \mathrm{d}h_1 \ldots \mathrm{d}h_k \left[\prod_{\ell\leqslant k} W(h_k)\right] \cosh^n\left(\beta \sum_{\ell\leqslant k} h_\ell\right) = \mathrm{e}^{\alpha c + \mathcal{O}(n)}. \tag{F.4}$$

We can hence write the RS version of our first saddle-point equation as follows, using $W(h) = W(-h)$:

$$\begin{aligned}
&\int \{\mathrm{d}\pi\} W[\pi] \prod_{\alpha=1}^{n} \pi(M_\alpha) = \mathrm{e}^{-c} \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \cos(\boldsymbol{\omega}\cdot\mathbf{M})\, \mathrm{e}^{c\mathrm{e}^{-\alpha c + \mathcal{O}(n)} \varrho(\boldsymbol{\omega})} \\
&= \mathrm{e}^{-c+\mathcal{O}(n)} \sum_{k\geqslant 0} \frac{c^k}{k!} \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \cos(\boldsymbol{\omega}\cdot\mathbf{M}) \langle \cos(\boldsymbol{\omega}\cdot\boldsymbol{\sigma}) \mathrm{e}^{\alpha c \int \mathrm{d}h W(h)[\mathrm{e}^{\beta h \sum_\alpha \sigma_\alpha} - 1]} \rangle_{\boldsymbol{\sigma}}^{k} \\
&= \mathrm{e}^{-c+\mathcal{O}(n)} \sum_{k\geqslant 0} \frac{c^k}{k!} \left\langle \left\langle \int_{-\pi}^{\pi} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^n} \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot(\tau\mathbf{M} - \sum_{\ell\leqslant k} \tau_\ell \boldsymbol{\sigma}^\ell)} \mathrm{e}^{\alpha c \sum_{\ell\leqslant k} \int \mathrm{d}h W(h)[\mathrm{e}^{\beta h \sum_\alpha \sigma_\alpha^\ell} - 1]} \right\rangle_{\boldsymbol{\sigma}^1 \ldots \boldsymbol{\sigma}^k} \right\rangle_{\tau, \tau_1 \ldots \tau_k = \pm 1} \\
&= \mathrm{e}^{-c+\mathcal{O}(n)} \sum_{k\geqslant 0} \frac{c^k}{k!} \mathrm{e}^{-\alpha c k} \langle \mathrm{e}^{\alpha c \sum_{\ell\leqslant k} \int \mathrm{d}h W(h)\, \mathrm{e}^{\beta h \sum_\alpha \sigma_\alpha^\ell}} \delta_{\mathbf{M}, \sum_{\ell\leqslant k} \boldsymbol{\sigma}^\ell} \rangle_{\boldsymbol{\sigma}^1 \ldots \boldsymbol{\sigma}^k} \\
&= \mathrm{e}^{-c+\mathcal{O}(n)} \sum_{k\geqslant 0} \frac{c^k}{k!} \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \left\langle \left(\int \mathrm{d}h W(h) \sum_{\ell\leqslant k} \mathrm{e}^{\beta h \sum_\alpha \sigma_\alpha^\ell} \right)^r \delta_{\mathbf{M}, \sum_{\ell\leqslant k} \boldsymbol{\sigma}^\ell} \right\rangle_{\boldsymbol{\sigma}^1 \ldots \boldsymbol{\sigma}^k} \\
&= \mathrm{e}^{-c+\mathcal{O}(n)} \sum_{k\geqslant 0} \frac{c^k}{k!} \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int \mathrm{d}h_1 \ldots \mathrm{d}h_r \\
&\qquad \times \left[\prod_{s\leqslant r} W(h_s)\right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \prod_{\alpha} \langle \mathrm{e}^{\beta \sum_{s\leqslant r} h_s \sigma_{\ell_s}} \delta_{M_\alpha, \sum_{\ell\leqslant k} \sigma_\ell} \rangle_{\sigma_1 \ldots \sigma_k} \\
&= \mathrm{e}^{-c+\mathcal{O}(n)} \sum_{k\geqslant 0} \frac{c^k}{k!} \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int \mathrm{d}h_1 \ldots \mathrm{d}h_r \\
&\qquad \times \left[\prod_{s\leqslant r} W(h_s)\right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \prod_{\alpha} \left\{ \frac{\langle \mathrm{e}^{\beta \sum_{s\leqslant r} h_s \sigma_{\ell_s}} \delta_{M_\alpha, \sum_{\ell\leqslant k} \sigma_\ell} \rangle_{\sigma_1 \ldots \sigma_k}}{\langle \mathrm{e}^{\beta \sum_{s\leqslant r} h_s \sigma_{\ell_s}} \rangle_{\sigma_1 \ldots \sigma_k}} \right\} \\
&= \int \{\mathrm{d}\pi\} \left(\prod_{\alpha} \pi(M_\alpha) \right) \mathrm{e}^{-c+\mathcal{O}(n)} \sum_{k\geqslant 0} \frac{c^k}{k!}\, \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int \mathrm{d}h_1 \ldots \mathrm{d}h_r \left[\prod_{s\leqslant r} W(h_s)\right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \\
&\qquad \times \prod_{M} \delta \left[\pi(M) - \frac{\langle \mathrm{e}^{\beta \sum_{s\leqslant r} h_s \sigma_{\ell_s}} \delta_{M, \sum_{\ell\leqslant k} \sigma_\ell} \rangle_{\sigma_1 \ldots \sigma_k}}{\langle \mathrm{e}^{\beta \sum_{s\leqslant r} h_s \sigma_{\ell_s}} \rangle_{\sigma_1 \ldots \sigma_k}} \right].
\end{aligned} \tag{F.5}$$

We thus conclude that for $n \to 0$ the following equation for $W[\pi]$ solves our saddle-point problem:

$$\begin{aligned}
W[\pi] = \mathrm{e}^{-c} \sum_{k\geqslant 0} \frac{c^k}{k!} \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int_{-\infty}^{\infty} \mathrm{d}h_1 \ldots \mathrm{d}h_r \left[\prod_{s\leqslant r} W(h_s)\right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \\
\times \prod_{M} \delta \left[\pi(M) - \frac{\langle \mathrm{e}^{\beta \sum_{s\leqslant r} h_s \sigma_{\ell_s}} \delta_{M, \sum_{\ell\leqslant k} \sigma_\ell} \rangle_{\sigma_1 \ldots \sigma_k}}{\langle \mathrm{e}^{\beta \sum_{s\leqslant r} h_s \sigma_{\ell_s}} \rangle_{\sigma_1 \ldots \sigma_k}} \right].
\end{aligned} \tag{F.6}$$

Everything is properly normalized, and if $W(h) = W(-h)$ the measure $W[\pi]$ is seen to permit only real-valued distributions $\pi(M)$ such that $\pi(M) \in [0, \infty)$ and $\pi(-M) = \pi(M)$ for all $M \in \mathbb{Z}$.

Appendix G. Continuous RS phase transitions via route II

Here we work with the order parameter equation that is written in terms of $W(h)$ only, i.e. (124), and look for phase transitions in the absence of external fields. For $P(\psi) = \delta(\psi)$ we must solve $W(h)$ from

$$W(h) = \mathrm{e}^{-c} \sum_{k\geqslant 0} \frac{c^k}{k!}\, \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int_{-\infty}^{\infty} \mathrm{d}h_1 \ldots \mathrm{d}h_r \left[\prod_{s\leqslant r} W(h_s)\right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \times \left\langle \delta\left[h - \frac{1}{2\beta} \log\left(\frac{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c + \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c + \beta\sum_{s\leqslant r} h_s \tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm 1}}{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c - \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c + \beta\sum_{s\leqslant r} h_s \tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm 1}}\right)\right]\right\rangle_{\tau=\pm 1}. \tag{G.1}$$

Clearly $W(h) = \delta(h)$ solves this equation for any temperature. Due to $W(h) = W(-h)$, we will always have $\int \mathrm{d}h W(h) h = 0$, so the first bifurcation away from $W(h) = \delta(h)$ is expected to be in the second moment. We write $h = \epsilon y$, with $0 < \epsilon \ll 1$, and expand in powers of ϵ. Upon setting $W(h) = \epsilon^{-1}\tilde{W}(h/\epsilon)$ we have

$$\tilde{W}(y) = \mathrm{e}^{-c} \sum_{k\geqslant 0} \frac{c^k}{k!}\, \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int_{-\infty}^{\infty} \mathrm{d}y_1 \ldots \mathrm{d}y_r \left[\prod_{s\leqslant r} \tilde{W}(y_s)\right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \times \left\langle \delta\left[y - \frac{1}{2\beta\epsilon} \log\left(\frac{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c + \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c + \beta\epsilon\sum_{s\leqslant r} y_s \tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm 1}}{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c - \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c + \beta\epsilon\sum_{s\leqslant r} y_s \tau_{\ell_s}}\right\rangle_{\tau_1\ldots\tau_k=\pm 1}}\right)\right]\right\rangle_{\tau=\pm 1}. \tag{G.2}$$

Next we expand the logarithm in the last line. To leading orders in ϵ we obtain

$$\frac{1}{2\beta\epsilon}\log(\ldots) = \frac{1}{2\beta\epsilon}\log\left(\frac{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c + \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\left[1 + \beta\epsilon\sum_{s\leqslant r} y_s\tau_{\ell_s}\right]\right\rangle_{\tau_1\ldots\tau_k=\pm 1}}{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c - \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\left[1 + \beta\epsilon\sum_{s\leqslant r} y_s\tau_{\ell_s}\right]\right\rangle_{\tau_1\ldots\tau_k=\pm 1}}\right)$$

$$= \frac{1}{2\beta\epsilon}\log\left(\frac{1 + \beta\epsilon\sum_{s\leqslant r} y_s \frac{\langle \tau_{\ell_s} \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c + \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\rangle_{\tau_1\ldots\tau_k=\pm 1}}{\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c + \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\rangle_{\tau_1\ldots\tau_k=\pm 1}}}{1 + \beta\epsilon\sum_{s\leqslant r} y_s \frac{\langle \tau_{\ell_s} \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c - \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\rangle_{\tau_1\ldots\tau_k=\pm 1}}{\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c - \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\rangle_{\tau_1\ldots\tau_k=\pm 1}}}\right)$$

$$= \frac{1}{2}\sum_{s\leqslant r} y_s \left\{\frac{\left\langle \tau_{\ell_s} \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c + \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\right\rangle_{\tau_1\ldots\tau_k}}{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c + \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\right\rangle_{\tau_1\ldots\tau_k}} - \frac{\left\langle \tau_{\ell_s} \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c - \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\right\rangle_{\tau_1\ldots\tau_k}}{\left\langle \mathrm{e}^{\beta(\sum_{\ell\leqslant k}\tau_\ell)^2/2c - \beta(\sum_{\ell\leqslant k}\tau_\ell)\tau/c}\right\rangle_{\tau_1\ldots\tau_k}}\right\}$$

$$= \tau\sum_{s\leqslant r} y_s \left\{\frac{\int \mathrm{D}z \tanh(z\sqrt{\beta/c} + \beta/c)\cosh^k(z\sqrt{\beta/c} + \beta/c)}{\int \mathrm{D}z \cosh^k(z\sqrt{\beta/c} + \beta/c)}\right\}. \tag{G.3}$$

Hence our order parameter equation (G.2) becomes

$$\tilde{W}(y) = \mathrm{e}^{-c} \sum_{k\geqslant 0} \frac{c^k}{k!}\mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int_{-\infty}^{\infty} \mathrm{d}y_1 \ldots \mathrm{d}y_r \left[\prod_{s\leqslant r} \tilde{W}(y_s)\right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \times \left\langle \delta\left[y - \tau\sum_{s\leqslant r} y_s \left\{\frac{\int \mathrm{D}z \tanh(z\sqrt{\beta/c} + \beta/c)\cosh^k(z\sqrt{\beta/c} + \beta/c)}{\int \mathrm{D}z \cosh^k(z\sqrt{\beta/c} + \beta/c)}\right\}\right]\right\rangle_{\tau=\pm 1}. \tag{G.4}$$

The first potential type of bifurcation away from $W(h) = \delta(h)$ would have $\int \mathrm{d}hW(h)h = \epsilon \int \mathrm{d}y\tilde{W}(y)y \equiv \epsilon m_1 \neq 0$. However, we see that multiplying both sides of (G.4) by y, followed by integration, immediately gives $m_1 = 0$. Thus, as expected, a bifurcation leading to a function $W(h)$ with $\int \mathrm{d}hW(h)h \neq 0$ is impossible.

Any continuous bifurcation will consequently have $\int \mathrm{d}hW(h)h = 0$ and $\int \mathrm{d}hW(h)h^2 = \epsilon^2 \int \mathrm{d}y\tilde{W}(y)y^2 \equiv \epsilon^2 m_2 \neq 0$. Multiplication of equation (G.4) by y^2, followed by integration over y gives

$$m_2 = \mathrm{e}^{-c} \sum_{k\geqslant 0} \frac{c^k}{k!} \mathrm{e}^{-\alpha c k} \sum_{r\geqslant 0} \frac{(\alpha c)^r}{r!} \int_{-\infty}^{\infty} \mathrm{d}y_1 \ldots \mathrm{d}y_r \left[\prod_{s\leqslant r} \tilde{W}(y_s) \right] \sum_{\ell_1 \ldots \ell_r \leqslant k} \times \sum_{s\leqslant r} y_s^2 \left\langle \left\{ \frac{\int \mathrm{D}z \tanh(z\sqrt{\beta/c} + \beta/c) \cosh^k(z\sqrt{\beta/c} + \beta/c)}{\int \mathrm{D}z \cosh^k(z\sqrt{\beta/c} + \beta/c)} \right\}^2 \right\rangle_{\tau=\pm 1}. \tag{G.5}$$

So now we get a bifurcation when

$$1 = \alpha c^2 \sum_{k\geqslant 0} \mathrm{e}^{-c} \frac{c^k}{k!} \left\{ \frac{\int Dz \tanh(z\sqrt{\beta/c} + \beta/c) \cosh^{k+1}(z\sqrt{\beta/c} + \beta/c)}{\int Dz \cosh^{k+1}(z\sqrt{\beta/c} + \beta/c)} \right\}^2. \tag{G.6}$$

We note that the RHS of (G.6) obeys $\lim_{\beta\to 0} \mathrm{RHS} = 0$ and $\lim_{\beta\to\infty} \mathrm{RHS} = \alpha c^2$. Hence a transition at finite temperature $T_c(\alpha, c) > 0$ exists in a new state with $W(h) \neq \delta(h)$ as soon as $\alpha c^2 > 1$. The critical temperature becomes zero when $\alpha c^2 = 1$, so $T_c(\alpha, 1/\sqrt{\alpha}) = 0$ for all $\alpha \geqslant 0$. For large c, using $\tanh x = x + \mathcal{O}(x^3)$ and $\cosh x = 1 + x^2/2$, valid for small x, we have $T_c = \sqrt{\alpha}$.

Appendix H. Coincidence of the two formulae for the transition line

In order to prove that the two expressions (G.6) and (C.20) for the RS transition line are identical, as they should be, we show that

$$\left\{ \frac{\int_{-\pi}^{\pi} \mathrm{d}\omega \sin^2(\omega) \cos^k(\omega) D(\omega|\beta)}{\int_{-\pi}^{\pi} \mathrm{d}\omega \cos^{k+2}(\omega) D(\omega|\beta)} \right\}^2 = \left\{ \frac{\int \mathrm{D}z \tanh(z\sqrt{\beta/c} + \beta/c) \cosh^{k+1}(z\sqrt{\beta/c} + \beta/c)}{\int \mathrm{D}z \cosh^{k+1}(z\sqrt{\beta/c} + \beta/c)} \right\}^2, \tag{H.1}$$

where $\mathrm{D}z = (2\pi)^{-1/2} \mathrm{e}^{-z^2/2}\, dz$. We can rewrite the argument of the curly brackets on the RHS, which we will denote as A, as

$$\begin{aligned} A &= \frac{\int \mathrm{D}z \sinh(z\sqrt{\beta/c} + \beta/c) \cosh^k(z\sqrt{\beta/c} + \beta/c)}{\int \mathrm{D}z \cosh^{k+1}(z\sqrt{\beta/c} + \beta/c)} \\ &= \frac{\int \mathrm{D}z \langle \tau_{k+1} \mathrm{e}^{(z\sqrt{\beta/c} + \beta/c) \sum_{\ell\leqslant k+1} \tau_\ell} \rangle_{\tau_1 \ldots \tau_{k+1} = \pm 1}}{\int \mathrm{D}z \langle \mathrm{e}^{(z\sqrt{\beta/c} + \beta/c) \sum_{\ell\leqslant k+1} \tau_\ell} \rangle_{\tau_1 \ldots \tau_{k+1} = \pm 1}} \\ &= \frac{\langle \tau_{k+1} \mathrm{e}^{(\beta/2c)(\sum_{\ell\leqslant k+1} \tau_\ell)^2 + (\beta/c) \sum_{\ell\leqslant k+1} \tau_\ell} \rangle_{\tau_1 \ldots \tau_{k+1} = \pm 1}}{\langle \mathrm{e}^{(\beta/2c)(\sum_{\ell\leqslant k+1} \tau_\ell)^2 + (\beta/c) \sum_{\ell\leqslant k+1} \tau_\ell} \rangle_{\tau_1 \ldots \tau_{k+1} = \pm 1}}, \end{aligned} \tag{H.2}$$

where we have carried out the Gaussian integrations. Next we insert $1 = \sum_{M\in\mathbb{Z}} \delta_{M, \sum_{\ell\leqslant k+1} \tau_\ell}$, and write the Kronecker delta in integral form. This gives

$$\begin{aligned} A &= \frac{\sum_{M\in\mathbb{Z}} \mathrm{e}^{(\beta/2c)M^2 + (\beta/c)M} \int_{-\pi}^{\pi} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega M} \langle \tau_{k+1} \mathrm{e}^{-\mathrm{i}\omega \sum_{\ell\leqslant k+1} \tau_\ell} \rangle_{\tau_1 \ldots \tau_{k+1} = \pm 1}}{\sum_{M\in\mathbb{Z}} \mathrm{e}^{(\beta/2c)M^2 + (\beta/c)M} \int_{-\pi}^{\pi} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega M} \langle \mathrm{e}^{-\mathrm{i}\omega \sum_{\ell\leqslant k+1} \tau_\ell} \rangle_{\tau_1 \ldots \tau_{k+1} = \pm 1}} \\ &= -\mathrm{i} \frac{\sum_{M\in\mathbb{Z}} \mathrm{e}^{(\beta/2c)M^2 + (\beta/c)M} \int_{-\pi}^{\pi} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega M} \cos^k(\omega) \sin(\omega)}{\sum_{M\in\mathbb{Z}} \mathrm{e}^{(\beta/2c)M^2 + (\beta/c)M} \int_{-\pi}^{\pi} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega M} \cos^{k+1}(\omega)}. \end{aligned} \tag{H.3}$$

By completing the square, $\sum_M \mathrm{e}^{(\beta/2c)M^2+(\beta/c)M} = \mathrm{e}^{-\beta/(2c)} \sum_M \mathrm{e}^{(\beta/2c)(M+1)^2}$, shifting the summation index $M \to M-1$, and using the symmetry properties (75) of $D(\omega|\beta)$ at zero fields, we finally get

$$\begin{aligned} A &= -\mathrm{i}\frac{\sum_{M\in\mathbb{Z}} \mathrm{e}^{(\beta/2c)M^2} \int_{-\pi}^{\pi} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega(M-1)} \cos^k(\omega)\sin(\omega)}{\sum_{M\in\mathbb{Z}} \mathrm{e}^{(\beta/2c)M^2} \int_{-\pi}^{\pi} \mathrm{d}\omega\, \mathrm{e}^{\mathrm{i}\omega(M-1)} \cos^{k+1}(\omega)} \\ &= -\mathrm{i}\frac{\int_{-\pi}^{\pi} \mathrm{d}\omega D(\omega|\beta)\cos^k(\omega)\sin(\omega)[\cos(\omega)-\mathrm{i}\sin(\omega)]}{\int_{-\pi}^{\pi} \mathrm{d}\omega D(\omega|\beta)\cos^{k+1}(\omega)[\cos(\omega)-\mathrm{i}\sin(\omega)]} \\ &= -\frac{\int_{-\pi}^{\pi} \mathrm{d}\omega D(\omega|\beta)\cos^k(\omega)\sin^2(\omega)}{\int_{-\pi}^{\pi} \mathrm{d}\omega D(\omega|\beta)\cos^{k+2}(\omega)}, \end{aligned} \tag{H.4}$$

which proves (H.1).

References

[1] Parisi G 1990 *Proc. Natl Acad. Sci. USA* **87** 429
[2] Barra A and Agliari E 2010 *J. Stat. Mech.* 07004
[3] Agliari E, Asti L, Barra A and Ferrucci L 2012 *Phys. Rev.* E **85** 051909
[4] Agliari E, Barra A, Guerra F and Moauro F 2011 *J. Theor. Biol.* **287** 48
[5] Mora T, Walczak A M, Bialek W and Callan C G 2010 *Proc. Natl Acad. Sci. USA* **107** 5405
[6] Košmrlj A, Chakraborty A K, Kardar M and Shakhnovich E I 2009 *Phys. Rev. Lett.* **103** 068103
[7] Košmrlj A, Jha A K, Huseby E S, Kardar M and Chakraborty A K 2008 *Proc. Natl Acad. Sci. USA* **105** 16671
[8] Chakraborty A K and Košmrlj A 2011 *Annu. Rev. Phys. Chem.* **61** 283
[9] Schmidtchen H, Thüne M and Behn U 2012 *Phys. Rev.* E **86** 011930
[10] Uezu T, Kadano C, Hatchett J P L and Coolen A C C 2006 *Prog. Theor. Phys.* **161** 385
[11] Perelson A S and Weisbuch G 1997 *Rev. Mod. Phys.* **69** 1219
[12] De Boer R J, Kevrekidis I G and Perelson A S 1993 *Bull. Math. Biol.* **55** 745
[13] Nesterenko V G 1988 *Theoretical Immunology* ed A S Perelson (New York: Addison-Wesley)
[14] Farmer J D, Packard N H and Perelson A S 1986 *Physica* D **22** 187
[15] Albert R and Barabasi A L 2002 *Rev. Mod. Phys.* **74** 47
[16] Hatchett J P L, Perez Castillo I, Coolen A C C and Skantzos N S 2005 *Phys. Rev. Lett.* **95** 117204
[17] Skantzos N S and Coolen A C C 2000 *J. Phys. A: Math. Gen.* **33** 5785
[18] Wemmenhove B and Coolen A C C 2003 *J. Phys. A: Math. Gen.* **36** 9617
[19] Agliari E and Barra A 2011 *Europhys. Lett.* **94** 10002
[20] Barra A, Genovese G and Guerra F 2010 *J. Stat. Phys.* **140** 784
[21] Barra A, Bernacchia A, Contucci P and Santucci E 2012 *Neural Netw.* **34** 1
[22] Coolen A C C, Kühn R and Sollich P 2005 *Theory of Neural Information Processing Systems* (Oxford: Oxford University Press)
[23] Agliari E, Barra A, Galluzzi A, Guerra F and Moauro F 2012 *Phys. Rev. Lett.* **109** 268101
[24] Agliari E, Barra A, Bartolucci S, Galluzzi A, Guerra F and Moauro F 2012 *Phys. Rev.* E **87** 042701
[25] Barkai E, Kanter I and Sompolinsky H 1990 *Phys. Rev.* A **41** 1843
[26] Coolen A C C 2001 *Handbook of Biological Physics* vol 4 ed F Moss and S Gielen (Amsterdam: Elsevier) pp 531–59
[27] Perez-Castillo I, Wemmenhove B, Hatchett J P L, Coolen A C C, Skantzos N S and Nikoletopoulos T 2004 *J. Phys. A: Math. Gen.* **37** 8789
[28] Sompolinsky H 1986 *Phys. Rev.* A **34** 2571
[29] Janeway C, Travers P, Walport M and Shlomchik M 2005 *Immunobiology* (New York: Garland Science)
[30] Abbas A K, Lichtman A H and Pillai S 2012 *Cellular and Molecular Immunology* (Philadelphia, PA: Elsevier)
[31] Barra A and Agliari E 2010 *Physica* A **389** 5903
[32] Theze J (ed) 1999 *The Cytokine Network and Immune Functions* (Oxford: Oxford University Press)
[33] Kuchroo V, Sarvetnick N, Hafler D and Nicholson L (ed) 2002 *Cytokine and Autoimmune Diseases* (Totowa, NJ: Humana Press)
[34] Goodnow C C 1992 *Annu. Rev. Immunol.* **10** 489
[35] Goodnow C C 2005 *Nature* **435** 590
[36] Goodnow C C, Vinuesa C G, Randall K L, Mackay F and Brink R 2010 *Nature Immunol.* **8** 681

[37] Schwartz R H 2005 *Nature Immunol.* **4** 327
[38] Agliari E, Annibale A, Barra A, Coolen A C C and Tantari D 2013 *J. Phys. A: Math. Theor.* **46** 335101
[39] Achlioptas D, D'Souza R M and Spencer J 2009 *Science* **323** 1453
[40] Donetti L, Neri F M and Munoz M 2007 *J. Stat. Mech.* P08007
[41] Derrida B, Gardner E and Zippelius A 1987 *Europhys. Lett.* **4** 167
[42] Wemmenhove B, Skantzos N S and Coolen A C C 2004 *J. Phys. A: Math. Gen.* **37** 7653
[43] Amit D J 1992 *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge: Cambridge University Press)
[44] Amit D J, Gutfreund H and Sompolinsky H 1987 *Phys. Rev.* A **35** 2293
[45] Newman M E J 2002 *Phys. Rev.* E **66** 016128
[46] Perez-Vicente C J and Coolen A C C 2008 *J. Phys. A: Math. Theor.* **41** 255003
[47] Agliari E and Barra A 2011 *Europhys. Lett.* **94** 10002
[48] Barra A and Agliari E 2011 *J. Stat. Mech.* P02027
[49] Coolen A C C, Skantzos N S, Pérez Castillo I, Pérez-Vicente C J, Hatchett J P L, Wemmenhove B and Nikoletopoulos T 2005 *J. Phys. A: Math. Gen.* **38** 8289
[50] Mézard M and Parisi G 2001 *Eur. Phys. J.* B **20** 217