Equilibrium statistical mechanics on correlated random graphs

J. Stat. Mech. (2011) P02027

(http://iopscience.iop.org/1742-5468/2011/02/P02027)

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Equilibrium statistical mechanics on correlated random graphs

Adriano Barra[1,2] **and Elena Agliari**[3,4,5]

[1] Dipartimento di Fisica, Sapienza Università di Roma, Italy
[2] GNFM, Gruppo Nazionale per la Fisica Matematica, Italy
[3] Dipartimento di Fisica, Università di Parma, Italy
[4] INFN, Gruppo Collegato di Parma, Italy
[5] Theoretische Polymerphysik, Albert-Ludwigs-Universität, Freiburg, Germany
E-mail: adriano.barra@roma1.infn.it and elena.agliari@fis.unipr.it

Received 7 September 2010
Accepted 24 January 2011
Published 18 February 2011

Online at stacks.iop.org/JSTAT/2011/P02027
doi:10.1088/1742-5468/2011/02/P02027

Abstract. Biological and social networks have recently attracted great attention from physicists. Among several aspects, two main ones may be stressed: a non-trivial topology of the graph describing the mutual interactions between agents and, typically, imitative, weighted, interactions. Despite such aspects being widely accepted and empirically confirmed, the schemes currently exploited in order to generate the expected topology are based on *a priori* assumptions and, in most cases, implement constant intensities for links.

Here we propose a simple shift $[-1,+1] \to [0,+1]$ in the definition of patterns in a Hopfield model: a straightforward effect is the conversion of frustration into dilution. In fact, we show that by varying the bias of pattern distribution, the network topology (generated by the reciprocal affinities among agents, i.e. the Hebbian rule) crosses various well-known regimes, ranging from fully connected, to an extreme dilution scenario, then to completely disconnected. These features, as well as small-world properties, are, in this context, emergent and no longer imposed *a priori*.

The model is throughout investigated also from a thermodynamics perspective: the Ising model defined on the resulting graph is analytically solved (at a replica symmetric level) by extending the double stochastic stability technique, and presented together with its fluctuation theory for a picture of criticality. Overall, our findings show that, at least at equilibrium, dilution (of whatever kind) simply decreases the strength of the coupling felt by the spins, but leaves the paramagnetic/ferromagnetic flavors unchanged. The main difference with respect to previous investigations is that, within our approach, replicas

1742-5468/11/P02027+29$33.00

do not appear: instead of (multi)-overlaps as order parameters, we introduce a class of magnetizations on all the possible subgraphs belonging to the main one investigated: as a consequence, for these objects a closure for a self-consistent relation is achieved.

Keywords: phase diagrams (theory), disordered systems (theory), random graphs, networks

Contents

1. Introduction to social and biological networks

Since the early investigations by Milgram [51], several efforts have been made in the context of social sciences in order to understand the structure of interactions within a social system. Granovetter defined this kind of investigation as an attempt to link 'micro- and macrolevels of sociological theories' and gave fundamental prescriptions [42]; in particular, he noticed that the stronger the link between two agents, the larger (on average) the overlap among the relevant neighborhoods, i.e. the higher the degree of cliquishness. Furthermore, he noticed that weak ties play a fundamental role, acting as bridges among sub-clusters of highly connected interacting agents [42]–[44]. As properly pointed out by Watts and Strogatz [59], such features render the Erdös–Rényi (ER) graphs [26] unable

J. Stat. Mech. (2011) P02027

to describe social systems, due to the lack of correlation between links, which constraints the resulting degree of cliquishness to be relatively small [12]. Through a mathematical technique (rewiring), they obtained a first attempt at defining the so-called 'small-world' graph [59]. This structure was also used as a substrate for statistical mechanics models and, in particular, for the Ising model: by treating the network as a clustered chain (solvable via e.g. the transfer matrix) overlapped on a sparse ER graph (solvable via e.g. the replica trick) some rigorous results could be obtained [53, 21].

Indeed, beyond topological investigations, which provide a description of mutual interactions among the components of a system, in the past decades the collective behavior of the system itself has also started to be analyzed in econometrics, by exploiting tools based on statistical mechanics. For instance, McFadden described the discrete choice as a one-body theory with external fields [49], Brock and Durlauf went over and considered also the effects of interactions among agents, all couplings meant to be positive, that is imitative [28, 33]. Of course, the role of anti-imitative ties might be fundamental for collective decision capabilities (see e.g. [18]), however, the greatest proportion of interactions is imitative and this prescription will be followed throughout this paper.

Somewhat close to social breakthrough, after the revolution of Watson and Crick, biological studies in the past fifty years have given rise to completely new fields of science such as genomics [35] and proteomics [38], which are ultimately strongly based on graph theories[6] [13]. Furthermore, in biology, networks appear at various levels: for instance, in immunology the complementary matching between the epitopes of immunoglobulins generates the so-called 'Jerne network' [47, 54, 15, 1]; at larger scales, we can cite food web [52], metabolic networks [48, 29] and virus spreading worldwide [22].

In general, one notices that the magnitude of links is not constant and, as a consequence, one postulates that it follows from a proper probability distribution. The independence of couplings seems to be a simplifying starting point. For example, Blake pointed out [25] that exons in hemoglobin correspond both to structural and functional units of protein, implicitly suggesting a non-null level of correlation among couplings; not that different is the viewpoint of Coolen and coworkers [31, 58]. Indeed, a correlated degree of disorder can dramatically influence the overall behavior of the system and should therefore be explicitly taken into account when trying a statistical mechanics approach.

We finally present the Hopfield model [46], which provides the paradigmatic model for neural networks. Interestingly, in this model there is a scalar product among the bit strings (the so-called Hebbian rule [45]) which can be seen as a measure of the strength of the ties. In the neural context the coupling can be either positive or negative as, in order to share statically memories over all neurons [9], it must use properties of spin glasses [14, 19, 50] as the key for having several minima in the fitness landscape. By varying tunable parameters (level of noise and amount of storage memories) the Hopfield model displays a region where it is paramagnetic, a region where it is a spin glass, and a region where it is a 'working memory' [10, 11].

Now, our aim in this work is to introduce and develop an approach to model discrete systems made up of many interacting components with inner degrees of freedom and able to capture the intrinsic connection between the kind of interactions among components

[6] It is in fact well established that complex organisms share roughly the same number of genes with simpler ones. As a result, a pure reductionism approach (according to which more complexity requires more genes) seems to fail and the interest in how the interactions among genes are arranged is increasing rapidly.

J. Stat. Mech. (2011) P02027

and the emergent topology describing the system itself. As we will see, this is realized through a redefinition of patterns in a Hopfield-like model: instead of using positive and negative values for the coupling in the Hebbian rule of the Hopfield model, we use positive and null values.

We show that, even in this context, by varying the tunable parameters, we recover several topologies, ranging from fully connected (i.e. every node is connected with any other node), to linearly diverging connectivity (i.e. the average number of neighbors per node $\bar{z}$ scales linearly with the system size V), to extreme dilution (i.e. $\bar{z}$ grows with V but sublinearly), to fully disconnected (i.e. no edges at all), possibly featuring non-trivial coupling distributions. Despite a wide range of topologies being recoverable, from the equilibrium statistical mechanics perspective we find that all these networks behave qualitatively the same, with strong differences restricted to dynamical features (in agreement with intuition), which we plan to investigate soon.

We stress that such properties are not due to *a priori* assumptions, but simply result from the kind of interactions (imitative and based on similarity) assumed. Hence, the relationship between such 'inner' details and the 'macroscopic' layout of the system can be highlighted.

The paper is organized as follows: in section 2, we present the model and provide the reader with all the related definitions. Section 3 addresses the topological analysis, while section 4 deals with the thermodynamical analysis based on techniques from statistical mechanics. In section 5 we present our discussion and outlook for the future.

2. The model: definitions

Let us consider V agents, each associated with a 'spin' variable $\pm 1 \ni \sigma_i, i \in (1, \dots, V)$. In a social framework (e.g. discrete choice in econometrics) $\sigma_i = +1$ ($\sigma_i = -1$) may mean that the ith agent agrees (disagrees) with a particular choice. In biological networks, i can label Kauffman genes (assuming undirected links) or Jerne lymphocytes in such a way that $\sigma_i = +1$ represents expression or firing state respectively, while quiescence is denoted by $\sigma_i = -1$.

The influence of external stimuli, e.g. the media in social networks, environmental variations imposing phenotypic changes via gene expression in proteomics, or viruses in immune networks, can be encoded by means of a one-body Hamiltonian term $H = \sum_i^V h_i \sigma_i$, with h_i suitable for the particular phenomenon (as brilliantly done by McFadden [49, 37], Eigen [56] and Burnet [30, 16], for social, gene and immunological systems respectively). When interactions among agents are also allowed, modeling is far harder; in the following we show how our model accounts for this issue.

First of all, each agent $i \in (1, \dots, V)$ is endowed with a set of L characters denoted by a binary string ξ_i of length L. For example, in the social context this string may represent a set of attributes characterizing the ith agent (e.g. $\xi_i^{\mu=1}$ may take into account attitudes toward the opposite sex, such that if $\xi_i^{\mu=1} = 1$, σ_i likes the opposite sex, otherwise $\xi_i^{\mu=1} = 0$; similarly $\xi_i^{\mu=2}$ may take into account attitudes toward smoking and so on up to $\mu = L$). In gene networks the overlap between bit strings may yield a measure of phylogenetic distance, while in immunological context it may generate the affinity matrix specifying the interaction strength between different lymphocytes.

doi:10.1088/1742-5468/2011/02/P02027

J. Stat. Mech. (2011) P02027

Now we want to associate a weighted link between two agents by comparing how many (positive) similarities they share, namely

$$J_{ij} = \sum_{\mu=1}^{L} \xi_i^\mu \xi_j^\mu. \tag{2.1}$$

This description naturally leads to the emergence of a hierarchical partition of the whole population into a series of layers, each layer being characterized by the sharing of an increasing number of characters. Of course, group membership, beyond defining individual identity, is a primary basis both for social and biological interactions. As a result, the interaction strength between individual i and j increases with increasing similarity.

Hence, including both one-body and two-body terms, the model we are describing reads as

$$H_V(\sigma;\xi) = \frac{1}{V}\sum_{i<j}^{V} J_{ij}(\xi)\sigma_i\sigma_j + \sum_{i}^{V} h_i\sigma_i, \tag{2.2}$$

formally identical to the Hopfield model.

It is worth stressing that the way strings ξ are generated is crucial for the overall performance of the system; here we adopt the following distribution:

$$P(\xi_i^\mu = +1) = \frac{1+a}{2}, \qquad P(\xi_i^\mu = 0) = \frac{1-a}{2}, \tag{2.3}$$

in such a way that, by tuning the parameter $a \in [-1,+1]$, the concentration of non-null entries for the ith string $\rho_i = \sum_\mu \xi_i^\mu$ can be varied. When $a \to -1$ there is no network, namely spins are non-interacting, while when $a \to +1$ we have that $J_{ij} = L$ for any couple and (renormalization through L^{-1} apart) we recover the standard Curie–Weiss (CW) model. On the other hand, when $a \neq 0$ the pattern distribution is biased, somehow similarly to the correlations investigated by Amit and coworkers in neural scenarios [11]. Moreover, from equations (2.3) we get $\langle \xi_i^\mu \xi_i^\nu \rangle = ((1+a)/2)[\delta_{\mu\nu} + ((1+a)/2)(1-\delta_{\mu\nu})]$.

As we will see, small values of a give rise to highly correlated, diluted networks, while, as a gets larger the network gets more and more connected and correlation among links vanishes. We also anticipate that in the thermodynamic limit and under the hypothesis of L linearly dependent on V, if we fix $a \in [-1,1]$, we end up with trivial topologies (either fully connected or completely disconnected), while a significant way to take a is by the scaling $a = -1 + \gamma/V^\theta$, where $\gamma > 0$ and $\theta > 0$ are finite, tunable parameters controlling (finely and coarsely, respectively) the degree of dilution (see section 3.3 for full coverage of this point).

In fact, although the theory is defined for any finite V and L, as standard in statistical mechanics, we are interested in the large V behavior (such that, under central limit theorem permissions, deviations from averaged values become negligible and the theory predictive). For this task we find it meaningful to let even L diverge linearly with the system size (to bridge conceptually to high storage neural networks), such that $\lim_{V\to\infty} L/V = \alpha$ defines α as another control parameter. Finally, since we are interested in the regime of large V and large L we will often use V as $V-1$ and L as $L-1$.

J. Stat. Mech. (2011) P02027

3. The emergent network

The set of strings $\{\xi_i^\mu\}_{i=1,\dots,V;\mu=1,\dots,L}$ together with the rule in equation (2.1) generates a weighted graph $\mathcal{G}(V, L, a)$ describing the mutual interactions among nodes. The following investigation is aimed at the study of its topological features, which, as well known, are intimately connected with the dynamical properties of phenomena occurring on the network itself (e.g. diffusion, transport, critical properties, coherent propagation, relaxation, only to cite a few [23, 27, 5, 7, 4, 6]). We first focus on the topology, neglecting the role of weights, and we say that two nodes i and j are connected whenever J_{ij} is strictly positive; disorder on couplings will be addressed in section 3.2.

It is immediate to see that the number ρ of non-null (i.e. equal to one) entries occurring in a string ξ is Bernoulli-distributed, namely

$$P_1(\rho; a, L) = \binom{L}{\rho} \left(\frac{1+a}{2}\right)^{\rho} \left(\frac{1-a}{2}\right)^{L-\rho}, \tag{3.1}$$

with average and variance, respectively,

$$\bar{\rho}_{a,L} = \sum_{\rho=0}^{L} \rho P_1(\rho; a, L) = \left(\frac{1+a}{2}\right) L, \tag{3.2}$$

$$\sigma^2_{a,L} = \overline{\rho^2}_{a,L} - \bar{\rho}^2_{a,L} = \left(\frac{1-a^2}{4}\right) L. \tag{3.3}$$

Moreover, the probability that a string is made up of null entries only is $\prod_{\mu=1}^{L} P(\xi_i^\mu = 0) = [(1-a)/2]^L$, thus, since we are allowing repetitions among strings, the lower bound for the number of isolated nodes is $V[(1-a)/2]^L$.

Let us consider two strings ξ_i and ξ_j of length L, with ρ_i and ρ_j non-null entries, respectively. Then, the probability $P_{\text{match}}(k; \rho_i, \rho_j, L)$ that such strings display k matching entries is

$$P_{\text{match}}(k; \rho_i, \rho_j, L) = \frac{\binom{L}{k}\binom{L-k}{\rho_i-k}\binom{L-\rho_i}{\rho_j-k}}{\binom{L}{\rho_i}\binom{L}{\rho_j}}, \tag{3.4}$$

which is the number of arrangements displaying k matchings over the number of all possible arrangements. As anticipated, for two agents to be connected it is sufficient that their coupling (see equation (2.1)) is larger than zero, i.e. that they share at least one trait. Therefore, we have the following link probability

$$P_{\text{link}}(\rho_i, \rho_j, L) = \sum_{k=1}^{L} P_{\text{match}}(k; \rho_i, \rho_j, L) = 1 - P_{\text{match}}(0; \rho_i, \rho_j, L) = 1 - \frac{(L-\rho_i)!(L-\rho_j)!}{L!(L-\rho_i-\rho_j)!}. \tag{3.5}$$

The previous expression shows that, in general, the link probability between two nodes does depend on the nodes considered through the related parameters ρ_i and ρ_j: when ρ_i and ρ_j are both large, the nodes are likely to be connected and vice versa. Another

J. Stat. Mech. (2011) P02027

kind of correlation, intrinsic to the model, emerges due to the fact that, given $\xi_i^\mu = 1$, the node 1 will be connected with all strings with non-null μth entry; this gives rise to a large (local) clustering coefficient c_i (see section 3.4). Such a correlation vanishes when a is sufficiently larger than -1, so that any generic couple has a relatively large probability to be connected; in this case the resulting topology is well approximated by a highly connected random graph. Moreover, when $a \to +1$ we recover the fully connected graph.

Finally, it is important to stress that, according to our assumptions, repetitions among strings are allowed and this, especially for finite L and V, can have dramatic consequences on the topology of the structure. In fact, the suppression of repetitions would spread out the distribution $P_1(\rho; a, L)$, allowing the emergence of strings with a large ρ (with respect to the expected mean value $L(1+a)/2$); such nodes, displaying a large number of connections, would work as hubs. On the other hand, recalling that the number of couples displaying perfect overlapping strings is $\sim V^2/2^L$, we have that in the thermodynamic limit and L growing faster than $\log V$, repetitions among strings have null measure.

3.1. Degree distribution

We focus our attention on an arbitrary string ξ with ρ non-null entries and we calculate the average probability $\bar{P}_{\text{link}}(\rho; a)$ that ξ is connected to another generic string, which reads as

$$\begin{aligned}\bar{P}_{\text{link}}(\rho; a) &= \sum_{\rho_i=0}^{L} P_1(\rho_i; a, L) P_{\text{link}}(\rho, \rho_i; L) \\ &= 1 - \left(\frac{1-a}{2}\right)^L \left(1 + \frac{1+a}{1-a}\right)^{L-\rho} = 1 - \left(\frac{1-a}{2}\right)^\rho . \end{aligned} \tag{3.6}$$

This result is actually rather intuitive, as it states that, in order to be linked to ξ, a generic node has to display at least a non-null entry corresponding to the ρ non-null entries of ξ. Notice that the link probability of equation (3.6) corresponds to a mean-field approach where we treat all the remaining nodes on average; accordingly, the degree distribution $P_{\text{degree}}(z; \rho, a, V)$ for ξ becomes

$$P_{\text{degree}}(z; \rho, a, V) = \binom{V}{z} \left[1 - \left(\frac{1-a}{2}\right)^\rho\right]^z \left(\frac{1-a}{2}\right)^{\rho(V-z)} . \tag{3.7}$$

Therefore, the number of null entries controls the degree distribution of the related node: a large ρ gives rise to narrow (i.e. small variance) distributions peaked at large values of z. Notice that $\bar{P}_{\text{link}}(\rho; a)$ and, accordingly, $P_{\text{degree}}(z; \rho, a, V)$ are independent of L.

More precisely, from equation (3.7), the average degree for a string displaying ρ non-null entries is

$$\bar{z}_\rho = V \left[1 - \left(\frac{1-a}{2}\right)^\rho\right], \tag{3.8}$$

while the related variance is

$$\sigma_\rho^2 = V \left[1 - \left(\frac{1-a}{2}\right)^\rho\right] \left(\frac{1-a}{2}\right)^\rho . \tag{3.9}$$

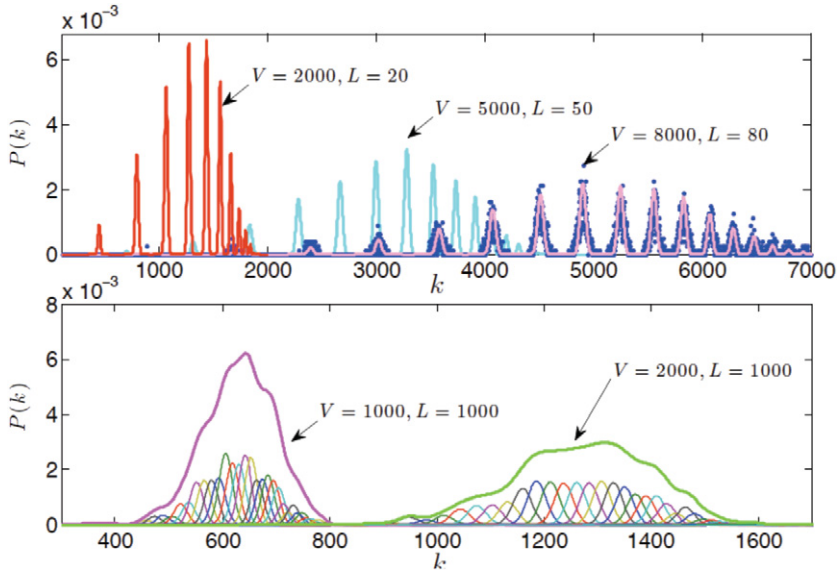

Figure 3.1. Degree distribution $\bar{P}_{\text{degree}}(k)$ for systems displaying small values of L and a multi-modal distribution (upper panel) and large values of L and a distribution collapsing into a unimodal one (lower panel). In the former case we compare systems of different sizes but with fixed $\alpha = 0.01$, where continuous lines represent the analytic estimate of equation (3.10) while symbols (●) represent data from simulation and for clarity are reported only for the case $V = 8000$. In the lower panel we compare systems with same L but different volumes; thicker curves represent $\bar{P}_{\text{degree}}(z; a, L, V)$, while each mode $P_{\text{degree}}(z; \rho, a, V)$ is depicted in different colors.

Now, the overall distribution can be written as a combination of binomial distributions

$$\bar{P}_{\text{degree}}(z; a, L, V) = \sum_{\rho=0}^{L} P_{\text{degree}}(z; \rho, a, V) P_1(\rho; a, L), \tag{3.10}$$

where the overlap among two 'modes', say ρ and $\rho + 1$, can be estimated through $\sigma_\rho/(\bar{z}_{\rho+1} - \bar{z}_\rho)$: exploiting equations (3.8) and (3.9) we get

$$\frac{\sigma_\rho}{\bar{z}_{\rho+1} - \bar{z}_\rho} = \sqrt{1 - \left(\frac{1-a}{2}\right)^\rho} \left[\sqrt{V}\left(\frac{1-a}{2}\right)^{\rho/2}\left(\frac{1+a}{2}\right)\right]^{-1} \sim \sqrt{\frac{L}{V}} = \sqrt{\alpha}, \tag{3.11}$$

where the generic mode ρ is replaced with $\bar{\rho}$ and the approximate result $\sqrt{L/V}$ was derived by using the scaling $a = -1 + \gamma/V^\theta$, with $1/2 < \theta \leq 1$ (both these points are fully discussed in section 3.2); also, the last passage holds rigorously in the thermodynamic limit of the high storage regime (L linearly diverging with V). Interestingly, for systems with different scaling regimes among L and V, for instance $L \propto \log V$ [15, 1], the distribution remains multi-modal because a vanishing overlap occurs among the single distributions $P_{\text{degree}}(z; \rho, a, V)$: $\bar{P}_{\text{degree}}(z; a, L, V)$ turns out to be an $(L+1)$-modal distribution (see figure 3.1, upper panel); vice versa, for $L \propto V$, the overall distribution becomes mono-modal (see figure 3.1, lower panel). Briefly, we mention that for $\theta = 1/2$ the ratio on the

J. Stat. Mech. (2011) P02027

lhs of equation (3.11) still converges to a finite value approaching $\sqrt{\alpha}$ for $\gamma^2 \ll \alpha$, while for $\theta < 1/2$ it diverges.

From equation (3.10), the average degree for a generic node is

$$\bar{z} = \sum_{z=0}^{V} z\, \bar{P}_{\text{degree}}(z; a, L, V) = \sum_{\rho=0}^{L} P_1(\rho; a; L)\bar{z}_\rho = V\left\{1 - \left[1 - \left(\frac{1+a}{2}\right)^2\right]^L\right\}, \tag{3.12}$$

where

$$p = 1 - \left[1 - \left(\frac{1+a}{2}\right)^2\right]^L \tag{3.13}$$

is the average link probability for two arbitrary strings ξ_i and ξ_j, which can be obtained by averaging over all possible string arrangements, namely, recalling equations (3.1) and (3.6),

$$\begin{aligned} p &= \sum_{\rho_i=0}^{L}\sum_{\rho_j=0}^{L} P_1(\rho_i; a, L)P_1(\rho_j; a, L)P_{\text{link}}(\rho_i, \rho_j; a, L) \\ &= 1 - \left(\frac{1-a}{2}\right)^{2L} \sum_{\rho_i=0}^{L}\sum_{\rho_j=0}^{L} \left(\frac{1+a}{1-a}\right)^{\rho_i+\rho_j} \binom{L}{\rho_i}\binom{L-\rho_i}{\rho_j} \\ &= 1 - \left[1 - \left(\frac{1+a}{2}\right)^2\right]^L. \end{aligned} \tag{3.14}$$

Of course, equation (3.14) could be obtained directly by noticing that the probability for the μth entries of two strings not to yield any contribution is $1 - [(1+a)/2]^2$, so that two strings are connected if there is at least one matching.

3.2. Coupling distribution

As explained in section 2, the coupling J_{ij} between nodes i and j is given by the relative number of matching entries between the corresponding strings ξ_i and ξ_j. Equation (3.4) provides the probability for ξ_i and ξ_j to share a link of magnitude $J = k$, namely $P_{\text{coupling}}(J; \rho_i, \rho_j, L) = P_{\text{match}}(k; \rho_i, \rho_j, L)$. Following the same arguments as in section 3.1 we get the probability that a link stemming from ξ_i has magnitude J, that is

$$\begin{aligned} \bar{P}_{\text{coupling}}(J; \rho_i, a) &= \sum_{\rho_j=J}^{L-\rho_i+J} P_{\text{coupling}}(J; \rho_i, \rho_j, L)P_1(\rho_j; a, L) \\ &= \binom{\rho_i}{J}\left(\frac{1-a}{2}\right)^{\rho_i-J}\left(\frac{1+a}{2}\right)^J, \end{aligned} \tag{3.15}$$

which is the probability that J out of ρ_i non-null entries are properly matched with the generic second node.

Similarly to $\bar{P}_{\text{degree}}(z; a, L, V)$, the overall coupling distribution can be written as the superposition $\sum_\rho P_1(\rho; a, L)\bar{P}_{\text{coupling}}(J; \rho, a)$. Each mode has variance $\sigma_\rho^2 = \rho(1-a^2)/4$

J. Stat. Mech. (2011) P02027

and is peaked at

$$\bar{J}_\rho = \rho\,\frac{1+a}{2}, \tag{3.16}$$

which represents the average coupling expected for links stemming from a node with ρ non-null entries. Nevertheless, by comparing $\bar{J}_{\rho+1} - \bar{J}_\rho = (1+a)/2$ and the standard deviation $\sqrt{\rho(1-a^2)}/2$, we find that in the limit, $L = \alpha V$ and $V \to \infty$ the distribution is mono-modal.

Anyhow, we can still define the average weighted degree w_ρ expected for a node displaying ρ non-null entries. Given that for the generic node i, $w \equiv \sum_j J_{ij}$, we get

$$\bar{w}_\rho = V\bar{J}_\rho = V\ \rho\,\frac{1+a}{2}. \tag{3.17}$$

Of course, one expects that the larger the coordination number of a node and the larger its weighted degree; such a correlation is linear only in the regime of low connectivity. In fact, by merging equations (3.8) and (3.17), one gets

$$\bar{w}_\rho = \left(\frac{1+a}{2}\right)\frac{\log\left(1-(\bar{z}_\rho/V)\right)}{\log\left(1-a/2\right)}V \approx \bar{z}_\rho, \tag{3.18}$$

where the last expression holds for $\bar{z}_\rho \ll V$ and $a \ll 1$.

It is important to stress that (pathological cases apart, which will be taken into account in the $L \to \infty$ scaling later on) the variance of ρ scales as $\sigma^2_\rho(a;L) = (1-a^2)L/4$ such that, despite the average of ρ is $(1+a)L/2$, substituting ρ/L with $(1+a)/2$ into equation (3.18) becomes meaningless in the thermodynamic limit as the variance of $\bar{J}_\rho$ diverges as $\sqrt{L} \propto \sqrt{V}$: this will affect drastically the thermodynamics whenever far from the CW limit.

It should be noted that $\bar{J}_\rho$ represents the average coupling for a link stemming from a node characterized by a string with ρ non-null entries, where the average includes also non-existing links corresponding to zero coupling. On the other hand, the ratio $\bar{w}_\rho/\bar{z}_\rho$ directly provides the average magnitude for existing couplings. Moreover, the average magnitude for a generic link is

$$\bar{J} = \sum_{\rho=0}^{L} P_1(\rho;a,L)\bar{J}_\rho = \left(\frac{1+a}{2}\right)^2. \tag{3.19}$$

By comparing equations (3.16) and (3.19) we notice that the local energetic environment seen by a single node, i.e. $\bar{J}_\rho$, and the overall energetic environment, i.e. $\bar{J}$, scale, respectively, linearly and quadratically with $(1+a)/2$: as we will see in the thermodynamic-dedicated section, despite the self-consistence relation (which is more sensible by local condition) is influenced by $\sqrt{\bar{J}}$, the critical behavior is found at $\beta_c = \bar{J}^{-1}$ consistently with a manifestation of a collective, global effect; of course, for the CW case ($a = 1$), this effect vanishes as global and local effects merge.

Anyhow, when V is large and the coupling distribution is narrowly peaked at the mode corresponding to $\bar{\rho}_{a,L}$, the couplings can be rather well approximated by the average value $\bar{J}_{L(1+a)/2} = [(1+a)/2]^2 = \bar{J}$, so that the disorder due to the weight distribution may be lost; as we will show this can occur in the regime of high dilution ($\theta > 1/2$). As for the other source of disorder (i.e. topological inhomogeneity), this can also be lost if a is sufficiently larger than -1, as we are going to show.

J. Stat. Mech. (2011) P02027

3.3. Scalings in the thermodynamic limit

In the thermodynamic limit and high storage regime, L is linearly divergent with V and the average probability p for two nodes to be connected (see equation (3.13)) approaches a discontinuous function, assuming value 1 when $a > -1$, and value 0 when $a = -1$. More precisely, as $V \to \infty$ there exists a vanishingly small range of values for a giving rise to a non-trivial graph; such a range is here recognized by the following scaling

$$a = -1 + \frac{\gamma}{V^\theta}, \tag{3.20}$$

where $\gamma, \theta \geq 0$ are finite parameters.

First of all, we notice that, following equations (3.2) and (3.3),

$$\bar{\rho}_{-1+\gamma/V^\theta,\alpha V} = \frac{\alpha\gamma}{2V^{\theta-1}} \tag{3.21}$$

$$\sigma^2_{-1+\gamma/V^\theta,\alpha V} = \frac{\alpha\gamma}{2V^{\theta-1}}\left(1 - \frac{\gamma}{2V^\theta}\right) \sim \bar{\rho}_{-1+\gamma/V^\theta,\alpha V}, \tag{3.22}$$

where the last approximation holds in the thermodynamic limit and it is consistent with the convergence of the binomial distribution in equation (3.1) to a Poissonian distribution. For $\theta \leq 1$, $\bar{\rho} \gtrsim \sigma^2 > 1$, so that when referring to a generic mode ρ, we can take without loss of generality $\bar{\rho}$; the case $\theta > 1$ will be neglected as it corresponds to a disconnected graph.

Indeed, the probability for two arbitrary nodes to be connected becomes

$$p = 1 - \left[1 - \left(\frac{1+a}{2}\right)^2\right]^L = 1 - \left[1 - \frac{\gamma^2}{4V^{2\theta}}\right]^{\alpha V} \underset{V\to\infty}{\to} 1 - \mathrm{e}^{-\gamma^2\alpha V^{1-2\theta}/4}, \tag{3.23}$$

so that we can distinguish the following regimes:

- $\theta < 1/2$, $p \approx 1$, $\bar{z} \approx V \Rightarrow$ fully connected (FC) graph.
- $\theta = 1/2$, $p \sim 1 - \mathrm{e}^{-\gamma^2\alpha/4} \sim \gamma^2\alpha/4$, $\bar{z} = O(V) \Rightarrow$ linearly diverging connectivity. Within a mean-field description the Erdös–Rényi (ER) random graph with finite probability $\mathcal{G}(V, p)$ is recovered.
- $1/2 < \theta < 1$, $p \sim \gamma^2\alpha V^{1-2\theta}/4$, $\bar{z} = O(V^{2-2\theta}) \Rightarrow$ extreme dilution regime (ED). In agreement with [61, 62], $\lim_{V\to\infty} \bar{z}^{-1} = \lim_{V\to\infty} \bar{z}/V = 0$.
- $\theta = 1$, $p \sim \frac{\gamma^2\alpha}{4V}$, $\bar{z} = O(V^0) \Rightarrow$ finite connectivity regime. Within a mean-field description $\gamma^2\alpha/4 = 1$ corresponds to a percolation threshold.

Therefore, while θ controls the connectivity regime of the network, γ allows a fine tuning.

As for the average coupling (see equation (3.19)) and the average weighted degree:

$$\bar{J} = \frac{\gamma^2}{4V^{2\theta}}, \tag{3.24}$$

$$\bar{w} = V\bar{J} = \frac{\gamma^2}{4V^{2\theta-1}}. \tag{3.25}$$

J. Stat. Mech. (2011) P02027

Now, the average 'effective coupling' $\tilde{J}$, obtained by averaging only on existing links, can be estimated as

$$\tilde{J} = \bar{J}/p = \begin{cases} \gamma^2/(4V^{2\theta}) & \text{if } \theta < 1/2 \\ \gamma^2/[4V^{2\theta}(1-\mathrm{e}^{-\gamma^2\alpha/4})] & \text{if } \theta = 1/2 \\ 1/(\alpha V) = 1/L & \text{if } 1/2 < \theta \leq 1. \end{cases} \tag{3.26}$$

Interestingly, this result suggests that in the thermodynamic limit, for values of a determined by equation (3.20) with $1/2 < \theta \leq 1$, nodes are pairwise either non-connected or connected due to one single matching among the relevant strings. This can be shown more rigorously by recalling the coupling distributions $\bar{P}_{\text{coupling}}(J;\rho_i,L)$ of equation (3.15): in particular, for $\theta > 1/2$, neglecting higher-order corrections, for $J=0$ the probability is $p_0 \sim \exp(\alpha\gamma^2 V^{1-2\theta}/4) \sim 1-\gamma^2\alpha/(4V^{2\theta-1})$, for $J=1/L$ the probability is $p_1 \sim p_0\gamma^2\alpha/(4V^{2\theta-1}) \sim 1-p_0$. For $\theta = 1/2$ this still holds for $\alpha\gamma^2/4 \ll 1$, which corresponds to a relatively high dilution regime, otherwise some degree of disorder is maintained, being that $p_k \sim (\alpha\gamma^2/4)^k/k!$. On the other hand, for $\theta < 1/2$, while topological disorder is lost (FC), the disorder due to the coupling distribution is still present. However, notice that for $\theta = 0$ and $\gamma = 2$, $\bar{P}_{\text{coupling}}(J;\rho_i,L)$ gets peaked at $J=L$ and, again, disorder on couplings is lost so that a pure CW model is recovered.

This means that, for $L = \alpha V$ and $V \to \infty$, we can distinguish three main regions in the parameter space (θ,α,γ) where the graph presents only topological disorder ($\theta > 1/2$), or only coupling disorder ($\theta < 1/2$), or both ($\theta = 1/2 \wedge \gamma^2\alpha = O(1)$).

In general, we expect that the critical temperature scales as the connectivity times the average coupling and the system can be looked at as fully connected with average coupling equal to $\bar{J}$ or as a diluted network with effective coupling $\tilde{J}$ and connectivity given by $\bar{z}$; in any case we get $\beta_c^{-1} \sim \bar{J}$ (see equation (4.36)).

3.4. Small-world properties

Small-world networks are endowed, by definition, with a high clustering coefficient, i.e. they display sub-networks that are characterized by the presence of connections between almost any two nodes within them, and with small diameter, i.e. the mean-shortest path length among two nodes grows logarithmically (or even slower) with V. While the latter requirement is a common property of random graphs [60, 12], the clustering coefficient deserves much more attention because of the basic role it covers in biological [63, 64] and social networks [42, 43].

The clustering coefficient measures the likelihood that two neighbors of a node are themselves linked; a higher clustering coefficient indicates a greater 'cliquishness' or transitivity. Two versions of this measure exist [60, 12]: global and local; as for the latter, the coefficient c_i associated with a node i describes how well connected the neighborhood of i is. If the neighborhood is fully connected c_i is one, while a value close to zero means that there are hardly any connections in the neighborhood.

The clustering coefficient of a node is defined as the ratio between the number of connections in the neighborhood of that node and the number of connections if the neighborhood was fully connected. Here neighborhood of node i means the nodes that are connected to i but does not include i itself. Therefore we have

$$c_i = \frac{2E_i}{z_i(z_i-1)}, \tag{3.27}$$

J. Stat. Mech. (2011) P02027

where E_i is the number of actual links present, while $z_i(z_i-1)/2$ is the number of connections for a fully connected group of z_i nodes. Of course, for the Erdös–Renyi graph where each link is independently drawn with a probability p, one has $c^{\mathrm{ER}}=p$, regardless of the node considered.

We now estimate the clustering coefficient for the graph $\mathcal{G}(a,L,V)$, focusing our attention on a range of a such that the average number of non-null entries per string is small enough for the link probability to be strictly lower than one, so that the topology is non-trivial; to fix our ideas and recalling the last section, $1/2\leq\theta\leq 1$. Let us consider a string displaying ρ non-null entries, corresponding to the positions $\mu_1,\mu_2,\ldots,\mu_\rho$, and z nearest neighbors; the latter can be divided into ρ groups: strings belonging to the jth group have $\xi^{\mu_j}=1$. Neglecting the possibility that a nearest neighbor can belong to more than one group contemporaneously (in the thermodynamic limit this is consistent with equation (3.26)), we denote by n_j the number of nodes belonging to the jth group, being $\sum_j n_j=z$, whose average value is z/ρ (which, due to the above assumptions is larger than one). Now, nodes belonging to the same group are all connected to each other as they share at least one common trait, i.e. they form a clique; the contribution of intra-group links is

$$E_{\mathrm{intra}}=\frac{1}{2}\sum_{i=1}^{\rho}n_i(n_i-1)=\frac{1}{2}\left(\sum_{i=1}^{\rho}n_i^2-z\right)\approx\frac{1}{2}\left[\left(\frac{z}{\rho}\right)^2\rho-z\right], \tag{3.28}$$

while the contribution of inter-group links can be estimated as

$$E_{\mathrm{inter}}\approx\sum_{i,j=1,i\neq j}^{\rho}n_in_j\tilde{p}\approx\left(\frac{z}{\rho}\right)^2\binom{\rho}{2}\tilde{p}, \tag{3.29}$$

where $\tilde{p}$ is the probability for two nodes linked to i and belonging to different groups to be connected, and the sum runs over all possible $\binom{\rho}{2}$ couples of groups. Hence, the total number of links between neighbors is $E=E_{\mathrm{intra}}+E_{\mathrm{inter}}=\{\sum_{i=1}^{\rho}n_i\sum_{j=1}^{\rho}n_j[\tilde{p}+(1-\tilde{p})\delta_{ij}]-z\}/2$, where δ_{ij} is the Kronecker delta, returning 1 if $i=j$ and zero otherwise; of course, for $\tilde{p}=1$ we have $E=(z^2-z)/2$ and $c_i=1$.

Now, on average, the probability $\tilde{p}$ is smaller than p as it represents the probability for two strings of length $L-1$ and displaying an average number of non-null entries equal to $\rho-1$ to be connected. However, for ρ and L not too small, the two probabilities converge so that by summing the two contributions in equations (3.28) and (3.29) we get

$$E\approx\frac{1}{2}\left[\left(\frac{z}{\rho}\right)^2\rho-z\right]+\left(\frac{z}{\rho}\right)^2\binom{\rho}{2}\tilde{p}\Rightarrow c\approx p+\frac{1}{\rho}-\frac{1}{z-1}>p, \tag{3.30}$$

where in the last inequality we used $\rho<z-1$. Therefore, it follows straightforwardly that c_i is larger than the clustering coefficient expected for an ER graph displaying the same connectivity, that is $c^{\mathrm{ER}}=p$.

From previous arguments it is clear that the small-world effect becomes more evident, with respect to the ER case taken as reference, when the network is highly diluted. This is confirmed by numerical data: figure 3.2 shows in the upper panel the clustering coefficient expected for the analogous ER graph, namely $c^{\mathrm{ER}}=\bar{z}/V$, while in the lower panel it shows

J. Stat. Mech. (2011) P02027

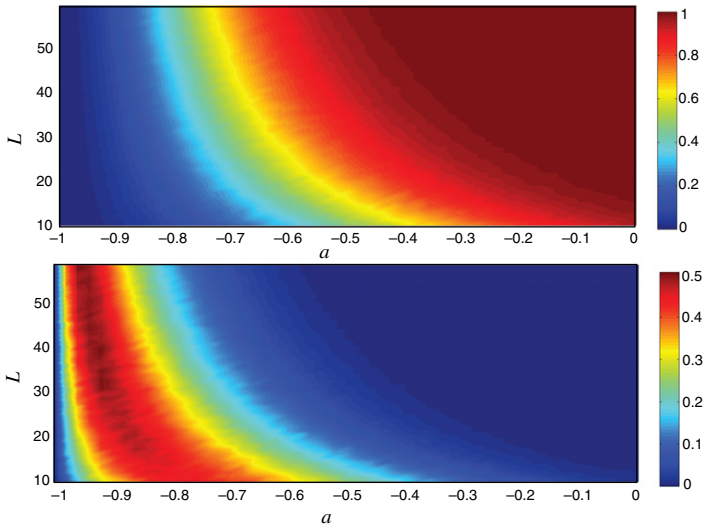

Figure 3.2. Upper panel: average link probability $p = \bar{z}/N$; lower panel: difference between the average clustering coefficient for $\mathcal{G}(V, L, a)$ and for an analogous ER graph just corresponding to p. Both plots are presented as function of a and L and refer to a system of $V = 2000$ nodes.

the difference between the average local clustering coefficient $c = \sum_{i=1}^{V} c_i/V$ and c^{ER} itself. Of course, when a approaches one, the graph becomes fully connected and $c \to c^{\mathrm{ER}} \to 1$.

Actually, a large cliquishness for $\mathcal{G}(V, L, a)$ could be simply derived by the kind of interaction underlying its construction: being the couplings J_{ij} based on string similarity, which is a transitive property[7], the whole graph is expected to exhibit a transitive structure as well, namely, if node i is connected (i.e. displays similar attributes) to both k and j, then k and j are also likely to be connected (i.e. to display similar attributes). Conversely, if we defined the couplings J_{ij} according to intransitive properties, such as complementarity or dissimilarity between strings, the resulting graph would display a small degree of cliquishness [15].

Finally, we mention that when focusing on the low storage regime, a non-trivial distribution for couplings can give rise to interesting effects. Indeed, weak ties can be shown [17, 8] to work as 'bridges', being crucial for maintaining the graph connectivity, as typical of real networks [42, 65]. Also, as often found in technological and biological networks, the graph under study displays a 'disassortative mixing', that is to say, high-degree vertices prefer to attach to low-degree nodes [60, 12] (see [17] for more details).

4. Thermodynamics

So far the emergent network has been exhaustively described by a random, correlated graph whose links are endowed with weights; we now build up a quantitative thermodynamics on such a structure.

[7] Notice that, actually, string similarity is not rigorously transitive, hence in the chain of neighbors $i \sim j \sim \cdots \sim k$, nodes i and k may turn out to display not so close attributes. However, for short series, such as triplets, the argument still works.

J. Stat. Mech. (2011) P02027

Once the Hamiltonian $H_V(\sigma;\xi)$ is given (equation (2.2)), we can introduce the partition function $Z_V(\beta;\xi)$ as

$$Z_V(\beta;\xi) = \sum_{\sigma} \mathrm{e}^{-\beta H_V(\sigma;\xi)}, \tag{4.1}$$

the Boltzmann state ω as

$$\omega(\cdot) = \frac{\sum_{\sigma} \cdot \mathrm{e}^{-\beta H_V(\sigma;\xi)}}{Z_V(\beta;\xi)}, \tag{4.2}$$

and the related free energy as

$$A(\beta,\alpha,a) = \lim_{V\to\infty} \frac{1}{V} \mathbb{E} \log Z_V(\beta;\xi), \tag{4.3}$$

where $\mathbb{E}$ averages over the quenched distributions of the affinities ξ.

Once the free energy (or equivalently the pressure) is obtained, remembering that (calling S the entropy and U the internal energy)

$$A(\beta,\alpha,a) = -\beta f(\beta,\alpha,a) = S(\beta,\alpha,a) - \beta U(\beta,\alpha,a),$$

the full macroscopic properties, thermodynamics, can be derived due the Legendre structure of thermodynamic potentials [55].

4.1. Free energy through extended double stochastic stability

For clarity we now describe in complete generality and detail the whole plan dealing with a generic expectation on ξ (i.e. $\mathbb{E}\xi = (1+a)/2$). Then, we will study more carefully the $L \to \infty$ scaling, in which a must tend to -1.

First of all, let us normalize the Hamiltonian (2.2) more conveniently for this section, that is, by dividing the J_{ij} by L, such that the effective coupling is bounded in $[0,1]$, and let us neglect the external field h (which can later be implemented straightforwardly).

$$H_V(\sigma;\xi) = \frac{1}{VL} \sum_{ij}^{V} \sum_{\mu}^{L} \xi_i^{\mu} \xi_j^{\mu} \sigma_i \sigma_j. \tag{4.4}$$

As a next step, through the Hubbard–Stratonovich transformation [55, 34], we map the partition function of our Hamiltonian into a bipartite Erdös–Rényi ferromagnetic random graph [2, 39], whose parts are the 'old' V dichotomic variables and the 'new' L Gaussian variables z_μ, $\mu \in (1,\dots,L)$:

$$Z(\beta;\xi) = \sum_{\sigma} \exp(-\beta H_V(\sigma;\xi)) = \sum_{\sigma} \int_{-\infty}^{+\infty} \prod_{\mu=1}^{L} \mathrm{d}\mu\,(z_\mu) \exp\left(\sqrt{\frac{\beta}{LV}} \sum_{i}^{V} \sum_{\mu}^{L} \xi_{i,\mu} \sigma_i z_\mu \right), \tag{4.5}$$

where with $\prod_{\mu=1}^{L} \mathrm{d}\mu\,(z_\mu)$ we mean the Gaussian measure on the product space of the Gaussian part. Note that, even when L goes to infinity linearly with V (as in the high storage Hopfield model [10]), due to the normalization encoded into the affinity product of the ξ's, neither does the z-diagonal term contribute to the free energy (as happens in the neural network counterpart [20]), nor (this will be made clear at the end of the section) is there a true dependence on α in the thermodynamics.

J. Stat. Mech. (2011) P02027

Furthermore, notice that the graph we are currently dealing with is bipartite, where links are randomly and uncorrelatedly drawn, and display no weight [12]: hence, the original complex structure for a single party has been transformed into a simpler structure but paying the price of accounting for another party in the interaction. The lack of weight on links will have fundamental importance when defining the order parameters.

Another way to see this is by noticing that if we dilute randomly the original Hopfield model (e.g. to check its robustness as already tested by Amit [9]) we obtain an ER topology, while if we introduce zero in pattern definitions we have to deal with correlated dilution.

Now, assuming the existence of the V limit, we want to solve the following free energy:

$$A(\beta,\alpha,a)=\lim_{V\to\infty}\frac{1}{V}\mathbb{E}\log\sum_{\sigma}\int_{-\infty}^{+\infty}\prod_{\mu}^{L}\mathrm{d}\mu\,(z_\mu)\exp\left(\sqrt{\frac{\beta}{LV}}\sum_{i}^{V}\sum_{\mu}^{L}\xi_{i,\mu}\sigma_i z_\mu\right). \tag{4.6}$$

For this task we extend the method of the double stochastic stability, recently developed in [20] in the context of neural networks. Namely, we introduce independent random fields $\eta_i, i\in(1,\dots,V)$ and $\chi_\mu,\mu\in(1,\dots,L)$, (whose probability distribution is the same as the ξ variables as in every cavity approach), which account for one-body interactions for the agents of the two parties. So our task is to interpolate between the original system and the one left with only these random perturbations: let us use $t\in[0,1]$ for such an interpolation; the trial free energy $A(t)$ is then introduced as follows

$$\begin{aligned}A(t)=\lim_{V\to\infty}\frac{1}{V}\mathbb{E}\log\sum_{\sigma}\int_{-\infty}^{+\infty}\prod_{\mu}^{L}\mathrm{d}\mu\,(z_\mu)\\ \times\ \exp\left[t\sqrt{\frac{\beta}{LV}}\sum_{i\mu}^{VL}\xi_{i\mu}\sigma_i z_\mu+(1-t)\left(\sum_{l_c=1}^{L}b_{l_c}\sum_{i}^{V}\eta_i\sigma_i+\sum_{l_b=1}^{V}c_{l_b}\sum_{\mu}^{L}\chi_\mu z_\mu\right)\right],\end{aligned} \tag{4.7}$$

where now $\mathbb{E}=\mathbb{E}_\xi\mathbb{E}_\eta\mathbb{E}_\chi$ and b_{l_c} (with $l_c\in(1,\dots,L)$) and c_{l_b} (with $l_b\in(1,\dots,V)$) are real numbers (possibly functions of β,α) to be set *a posteriori*.

As the theory is no longer Gaussian, we need infinite sets of random fields, mapping the presence of multi-overlaps in standard dilution [2, 36] and no longer just the first two momenta of the distributions.

Of course, we recover the proper free energy by evaluating the trial $A(t)$ at $t=1$, $(A(\beta,\alpha,a)=A(t=1))$, which we want to obtain by using the fundamental theorem of calculus:

$$A(1)=A(0)+\int_0^1(\partial A(t')/\partial t')_{t'=t}\,\mathrm{d}t. \tag{4.8}$$

To this task we need two objects: the trial free energy $A(t)$ evaluated at $t=0$ and its t-streaming $\partial_t A(t)$.

Before outlining the calculations, some definitions are in order here to lighten the notation: taking g as a generic function of the quenched variables we have

$$\mathbb{E}_\eta g(\eta)=\sum_{l_b=0}^{V}P(l_b)g(\eta_{l_b})=\sum_{l_b=0}^{V}\binom{V}{l_b}\left(\frac{1+a}{2}\right)^{V-l_b}\left(\frac{1-a}{2}\right)^{l_b}g(\eta_{l_b}), \tag{4.9}$$

$$\mathbb{E}_\chi g(\chi) = \sum_{l_c=0}^{L} P(l_c) g(\chi_{l_c}) = \sum_{l_c=0}^{L} \binom{L}{l_c} \left(\frac{1+a}{2}\right)^{L-l_c} \left(\frac{1-a}{2}\right)^{l_c} g(\chi_{l_c}), \tag{4.10}$$

$$\mathbb{E}_\xi g(\xi) = \sum_{l_b=0}^{V} \sum_{l_c=0}^{L} \binom{V}{l_b} \binom{L}{l_c} \left(\frac{1+a}{2}\right)^{l_b+l_c} \left(\frac{1-a}{2}\right)^{V+L-l_b-l_c} \delta_{l_b l_c = l}, \tag{4.11}$$

where $P(l_b)$ is the probability that l_b (out of V random fields) are active, i.e. $\eta = 1$, so that the number of spins effectively contributing to the function g is l_b; analogously, mutatis mutandis, for $P(l_c)$. Moreover, in the last equation we summed over the probability $P(l)$ that in the bipartite graph a number l of links out of the possible $V \times L$ display a non-null coupling, i.e. $\xi \neq 0$; interestingly, equation (4.10) can be rewritten in terms of the above mentioned $P(l_b)$ and $P(l_c)$. In fact, $\xi_{i,\mu}$ can be looked at as a $V \times L$ matrix generated by the product of two given vectors like η and χ, namely $\xi_{i,\mu} = \eta_i \chi_\mu$, in such a way that the number of non-null entries in the overall matrix ξ is just given by the number of non-null entries displayed by η times the number of non-null entries displayed by χ. Hence, $P(l)$ is the product of $P(l_b)$ and $P(l_c)$ conditional to $l_b l_c = l$.

4.2. The 'topologically microcanonical' order parameters

Starting with the streaming of equation (4.7), this operation gives rise to the sum of three terms $\mathcal{A} + \mathcal{B} + \mathcal{C}$. The former when deriving the first contributions into the exponential, the last two terms when deriving the two contributions by all the η and χ.

$$\mathcal{A} = +\frac{1}{V}\sqrt{\frac{\beta}{LV}} \sum_{i,\mu}^{V,L} \mathbb{E}\xi_{i\mu}\omega(\sigma_i z_\mu) = \sqrt{\alpha\beta}\left(\frac{1+a}{2}\right) \sum_{l_b,l_c}^{V,L} P(l_b)P(l_c)M_{l_b}N_{l_c} \tag{4.12}$$

$$\mathcal{B} = -\sum_{l_c=1}^{L} \frac{b_{l_c}}{V} \sum_{i}^{V} \mathbb{E}\eta_i \omega(\sigma_i) = -\sum_{l_c=1}^{L} b_{l_c} \left(\frac{1+a}{2}\right) \sum_{l_b=0}^{V} P(l_b) M_{l_b} \tag{4.13}$$

$$\mathcal{C} = -\sum_{l_b=1}^{V} \frac{c_{l_b}}{N} \sum_{\mu}^{L} \mathbb{E}\chi_\mu \omega(z_\mu) = -\sqrt{\alpha} \sum_{l_b=1}^{V} c_{l_b} \left(\frac{1+a}{2}\right) \sum_{l_c=0}^{L} P(l_c) N_{l_c}, \tag{4.14}$$

where we introduced the following order parameters

$$M_{l_b} = \frac{1}{V} \sum_{i}^{V} \omega_{l_b+1}(\sigma_i), \tag{4.15}$$

$$N_{l_c} = \frac{1}{L} \sum_{\mu}^{L} \omega_{l_c+1}(z_\mu), \tag{4.16}$$

and the Boltzmann states ω_k are defined by taking into account only k terms among the elements of the party involved.

Of course the Boltzmann states are no longer the ones introduced into the definition (4.2) but the extended ones, taking into account the interpolating structure of the cavity fields (which however will recover the originals of statistical mechanics when evaluated at $t = 1$).

J. Stat. Mech. (2011) P02027

Namely, ω_{l_b+1} has only l_b+1 terms of the type $b\sigma$ in the Maxwell–Boltzmann exponential, ultimately accounting for the (all equivalent in distribution) l_b+1 values of $\eta=1$, all the others being zero.

In the same way ω_{l_c+1} has only l_c+1 terms of the type cz in the Maxwell–Boltzmann exponential, ultimately accounting for the (all equivalent in distribution) l_c+1 values of $\chi=1$, all the others being zero.

When dealing with $\xi_{i\mu}$ we can decompose the latter according to what was discussed before. By these 'partial Boltzmann states' we can define the averages of the order parameters as

$$\langle M\rangle=\sum_{l_b}^{V-1}P(l_b)M_{l_b}, \tag{4.17}$$

$$\langle N\rangle=\sum_{l_c}^{L-1}P(l_c)N_{l_c}. \tag{4.18}$$

These objects deserve more explanation because, as a main difference from classical approaches [2, 32, 36], here replicas and their overlaps are not involved (somehow suggesting the implicit correctness of a replica symmetric scenario). Conversely, we do conceptually two (standard) operations when introducing our order parameters: at first we average over the (t-extended) Boltzmann measure, then we average over the quenched distributions. Let us consider only one party for simplicity: during the first operation we do not take the whole party size but only a subsystem, say k spins (whose distribution is symmetric with respect to zero for both the parties, $-1,+1$ for the dichotomic, Gaussians for the continuous one). Then, in the second average, for any k from one to the volume of the party, we consider all the possible links among these k nodes in this subgraph. As the links connecting the nodes are always constant (i.e. equal to one due to the Hubbard–Stratonovich transformation (4.4)) in the intensity, the resulting associated energies are, in distribution and in the thermodynamic limit, all equivalent: we are introducing a family of microcanonical observables which sum up to a canonical one, in some sense close to the decomposition introduced in [19].

4.3. The sum rule

Let us now move on and consider the following source S of the fluctuations of the order parameters, where $\bar{M}_{l_b},\bar{N}_{l_c}$ stand for the replica symmetric values[8] of the previously introduced order parameters:

$$S=\left(\frac{1+a}{2}\right)\sqrt{\alpha\beta}\sum_{l_b}^{V-1}\sum_{l_c}^{L-1}P(l_b)P(l_c)((M_{l_b}-\bar{M}_{l_b})(N_{l_c}-\bar{N}_{l_c})) \tag{4.19}$$

$$=\left(\frac{1+a}{2}\right)\sqrt{\alpha\beta}\langle(M-\bar{M})(N-\bar{N})\rangle. \tag{4.20}$$

[8] Strictly speaking there are no replicas here but configurations over different graphs. However, the expression RS-approximation, meaning that we assume the probability distribution of the order parameters delta-like over their average (denoted with a bar), is a sort of self-averaging and is a hinge in disordered statistical mechanics such that we allow ourselves to retain the same expression with a little abuse of language.

J. Stat. Mech. (2011) P02027

We see that with the choice of the parameters $b_{l_c} = \sqrt{\alpha\beta}\bar{N}_{l_c}$ and $c_{l_b} = \sqrt{\beta/\alpha}\bar{M}_{l_b}$, we can write the t-streaming as

$$\dot{A} = S - \frac{1+a}{2}\sqrt{\alpha\beta}\sum_{l_b}^{V-1}\sum_{l_c}^{L-1} P(l_b)P(l_c)\bar{M}_{l_b}\bar{N}_{l_c}.$$

The replica symmetric solution (which is claimed to be the correct expression in diluted ferromagnets) is simply achieved by setting $S = 0$ and omitting it from future calculations.

We must now evaluate $A(0)$. This term is given by two separate contributions, one for each party. Namely we have

$$\begin{aligned} A(0) &= \frac{1}{V}\mathbb{E}\log\sum_{\sigma} \mathrm{e}^{\sum_{l_c=1}^{L} b_{l_c}\sum_i^V \eta_i\sigma_i} + \frac{1}{V}\mathbb{E}\log\int_{-\infty}^{+\infty}\prod_{\mu=0}^{L}\mathrm{d}\mu\,(z_\mu)\mathrm{e}^{\sum_{l_b=1}^{V} c_{l_b}\sum_\mu^L \chi_\mu z_\mu} \\ &= \log 2 + \left(\frac{1+a}{2}\right)\sum_{l_b=0}^{V-1} P(l_b)\sum_{l_c=0}^{L-1} P(l_c)\log\cosh(\sqrt{\alpha\beta}\bar{N}_{l_c}) \\ &\quad + \left(\frac{1+a}{2}\right)^2\frac{\beta}{2}\sum_{l_b}^{V-1} P(l_b)\bar{M}_{l_b}^2. \end{aligned}$$

Summing $A(0)$ plus the integral of $\partial_t[A(S=0)]$, we finally get

$$\begin{aligned} A(t=1) &= \log 2 + \left(\frac{1+a}{2}\right)\sum_{l_c}^{L-1}\sum_{l_b}^{V-1} P(l_c)P(l_b)\log\cosh(\sqrt{\alpha\beta}\bar{N}_{l_c}) \\ &\quad + \frac{\beta}{2}\left(\frac{1+a}{2}\right)^2\sum_{l_b}^{L-1} P(l_b)\bar{M}_{l_b}^2 - \frac{1+a}{2}\sqrt{\alpha\beta}\sum_{l_b}^{V-1}\sum_{l_c}^{L-1} P(l_b)P(l_c)\bar{M}_{l_b}\bar{N}_{l_c}. \end{aligned} \quad (4.21)$$

As found in [20, 40], for bipartite ferromagnetic models the free energy obeys a min–max principle by which, extremizing the free energy with respect to the order parameters, we can express $\langle\bar{N}\rangle$ through $\langle\bar{M}\rangle$: the trial replica symmetric solution, expressed through $\langle\bar{M}\rangle, \langle\bar{N}\rangle$ is (at fixed $\langle\bar{N}\rangle$) convex in $\langle\bar{M}\rangle$. This defines uniquely a value $\langle\bar{M}(\bar{N})\rangle$ where we get the maximum. Furthermore, $\langle\bar{M}(\bar{N})\rangle$ is increasing and convex in $\langle\bar{N}^2\rangle$ such that the following extremization is a well defined procedure.

$$\sum_k P(k)\frac{\partial A}{\partial\bar{M}_k} = 0 \to \sum_{l_c} P(l_c)\bar{N}_{l_c} = \sum_k P(k)\frac{1+a}{2}\sqrt{\frac{\beta}{\alpha}}\bar{M}_k, \quad (4.22)$$

$$\sum_k P(k)\frac{\partial A}{\partial\bar{N}_{l_k}} = 0 \to \sum_{l_b} P(l_b)\bar{M}_{l_b} = \sum_k P(k)\tanh\left(\sqrt{\alpha\beta}\bar{N}_k\right). \quad (4.23)$$

Due to the mean-field nature of the model, as we can express $\bar{N}_k$ through the average of the $\bar{M}_k$, we can write the free energy of our network through the series of $\bar{M}_k$ alone, as

J. Stat. Mech. (2011) P02027

expected since we started from a one-party system (see equation (2.2))

$$A(\beta,a)=\log 2+\left(\frac{1+a}{2}\right)\log\cosh\left(\tanh^{-1}\left[\sum_{l_b}P(l_b)\bar{M}_{l_b}\right]\right)+\frac{\beta}{2}\left(\frac{1+a}{2}\right)^2\sum_{l_b}P(l_b)\bar{M}_{l_b}^2-\left(\frac{1+a}{2}\right)\sum_{l_b}P(l_b)\bar{M}_{l_b}\tanh^{-1}\left(\sum_{l_b'}P(l_{b'})\bar{M}_{l_b'}\right). \tag{4.24}$$

As anticipated there is no true dependence on α. Note that without normalizing the scalar product among the bit strings we should rescale β accordingly with α, as in the $L\to\infty$ limit we would get an infinite coupling (which is physically meaningless).

Before exploring further properties of these networks, we should recover the well-known limit of Curie–Weiss ($a=+1$) and of the isolated spin system ($a=-1$).

For clarity let us work out the self-consistency in a purely CW style by extremizing equation (4.24), with respect to $\langle\bar{M}\rangle$:

$$\partial_{\langle\bar{M}\rangle}A(\beta)=\left(\frac{1+a}{2}\right)\left[\frac{\langle\bar{M}\rangle}{1-\langle\bar{M}\rangle^2}+\beta\left(\frac{1+a}{2}\right)\langle\bar{M}\rangle\right]-\left(\frac{1+a}{2}\right)\left(\frac{\langle\bar{M}\rangle}{1-\langle\bar{M}\rangle^2}+\tanh^{-1}\langle\bar{M}\rangle\right)=0$$
$$\Rightarrow\langle\bar{M}\rangle=\tanh\left[\beta\left(\frac{1+a}{2}\right)\langle\bar{M}\rangle\right]\Rightarrow\tanh^{-1}\langle\bar{M}\rangle=\beta\left(\frac{1+a}{2}\right)\langle\bar{M}\rangle, \tag{4.25}$$

such that to get the classical magnetization in our model we have to sum over all the contributing graphs, namely $\langle M_{\mathrm{CW}}\rangle=\langle\bar{M}\rangle=\sum_{l_b}P(l_b)\bar{M}_{l_b}$, and we immediately recover

$$a\to-1\Rightarrow A(\beta,a=-1)=\log 2, \tag{4.26}$$

$$a=+1\Rightarrow A(\beta,a=+1)=\log 2+\log\cosh(\beta\langle\bar{M}\rangle)-\frac{\beta}{2}\langle M^2\rangle, \tag{4.27}$$

which are the correct limits (note that in equation (4.26) $J=0$, while in equation (4.27) $J=1$).

Furthermore, we stress that in our 'topological microcanonical' decomposition of our order parameters, when summing over all the possible subgraphs to obtain the CW magnetization, these are all null except the fully connected network, hence the distribution of the order parameters becomes trivially $\propto\delta(\bar{M}-M_{\mathrm{CW}})$, namely, only one order parameter, corresponding to the classical CW magnetization, survives.

4.4. Critical line through fluctuation theory

Developing a fluctuation theory of the order parameters allows one to determine where critical behavior arises and, ultimately, the existence of a phase transition[9].

[9] Strictly speaking this approach holds only for a second-order phase transition, which indeed is the one expected in imitative models, even in presence of dilution [2].

J. Stat. Mech. (2011) P02027

For this task we have at first to work out the general streaming equation with respect to the t-flux. Given a generic observable $\mathcal{O}$ defined on the space of the σ, z variables, it is immediate to check that the following relation holds; for the sake of simplicity we set $\alpha = 1$ as it never appears in the calculations (this can be easily checked by substituting $\langle N\rangle$ with $\langle M\rangle$ through equation (4.22) which changes the prefactor $(1+a/2)\sqrt{\alpha\beta}$ into $(1+a/2)^2\beta$ and express the fluctuations only via the real variables[10] σ):

$$\frac{\partial\langle\mathcal{O}\rangle}{\partial t} = \frac{1+a}{2}\sqrt{\beta}[(\langle\mathcal{O}\mathcal{M}\mathcal{N}\rangle - \langle\mathcal{O}\rangle\langle\mathcal{M}\mathcal{N}\rangle) - \bar{N}(\langle\mathcal{O}\mathcal{M}\rangle - \langle\mathcal{O}\rangle\langle\mathcal{M}\rangle) - \bar{M}(\langle\mathcal{O}\mathcal{N}\rangle - \langle\mathcal{O}\rangle\langle\mathcal{N}\rangle)], \tag{4.28}$$

where we defined the centered and rescaled order parameters:

$$\langle\mathcal{M}\rangle = \sqrt{V}\sum_{l_b} P(l_b)(M_{l_b} - \bar{M}_{l_b}) = \sqrt{V}\langle M - \bar{M}\rangle, \tag{4.29}$$

$$\langle\mathcal{N}\rangle = \sqrt{L}\sum_{l_c} P(l_c)(N - \bar{N}) = \sqrt{L}\langle N - \bar{N}\rangle. \tag{4.30}$$

Now we focus on their squares: we want to obtain the behavior of $\langle\mathcal{M}^2\rangle_{t=1}$, $\langle\mathcal{M}\mathcal{N}\rangle_{t=1}$, $\langle\mathcal{N}^2\rangle_{t=1}$, so as to see where their divergencies (onsetting the phase transition) are located.

By defining the dot operator as

$$\langle\dot{\mathcal{O}}\rangle = \left(\frac{1+a}{2}\right)\sqrt{\beta}\partial_t\langle\mathcal{O}\rangle \tag{4.31}$$

we can write

$$\langle\dot{\mathcal{M}}^2\rangle = [\langle\mathcal{M}^3\mathcal{N}\rangle - \langle\mathcal{M}^2\rangle\langle\mathcal{M}\mathcal{N}\rangle - \bar{N}\langle\mathcal{M}^3\rangle + \bar{N}\langle\mathcal{M}^2\rangle\langle\mathcal{M}\rangle - \bar{M}\langle\mathcal{M}^2\mathcal{N}\rangle + \bar{M}\langle\mathcal{M}^2\mathcal{N}\rangle],$$
$$\langle\dot{\mathcal{M}}\mathcal{N}\rangle = [\langle\mathcal{M}^2\mathcal{N}^2\rangle - \langle\mathcal{M}\mathcal{N}\rangle\langle\mathcal{M}\mathcal{N}\rangle - \bar{N}(\langle\mathcal{M}^2\mathcal{N}\rangle - \langle\mathcal{M}\mathcal{N}\rangle\langle\mathcal{M}\rangle) - \bar{M}(\langle\mathcal{M}\mathcal{N}^2\rangle - \langle\mathcal{M}\mathcal{N}\rangle\langle\mathcal{N}\rangle)],$$
$$\langle\dot{\mathcal{N}}^2\rangle = [\langle\mathcal{N}^3\mathcal{M}\rangle - \langle\mathcal{N}^2\rangle\langle\mathcal{M}\mathcal{N}\rangle - \bar{N}\langle\mathcal{N}^2\mathcal{M}\rangle + \bar{N}\langle\mathcal{N}^2\rangle\langle\mathcal{M}\rangle - \bar{M}\langle\mathcal{N}^3\rangle + \bar{M}\langle\mathcal{N}^2\rangle\langle\mathcal{N}\rangle].$$

Now, for the sake of simplicity, let us introduce alternative labels for the fundamental observables. We define $A(t) = \langle\mathcal{M}^2\rangle_t$, $D(t) = \langle\mathcal{M}\mathcal{N}\rangle_t$ and $G(t) = \langle\mathcal{N}^2\rangle_t$ and let us work out their $t = 0$ value, which is straightforward as at $t = 0$ everything is factorized (alternatively these can be seen as high noise expectations):

$$A(t=0) = 1, \qquad D(t=0) = 0, \qquad G(t=0) = \left(1 + \left(\frac{1+a}{2}\right)^2\frac{\beta}{\alpha}\langle\bar{M}^2\rangle\right) - \langle\bar{N}^2\rangle = 1$$

where we used the self-consistence relation (4.22) and assumed that, at least where everything is completely factorized, the replica solution is the true solution[11]. Following

[10] Another simple argument to understand the uselessness of α is a comparison with neural networks: in that context, and assuming the thermodynamic limit, α controls the velocity by which we add stored memories into the network with respect to the velocity by which we add neurons. If the former are faster than a critical value, by a central limit theorem argument they sum up to a Gaussian before the infinite volume limit has been achieved and the Hopfield model turns into a Sherrington–Kirkpatrick model [20]. Here there is no danger in this as we have only positive, normalized interactions.

[11] Any debate concerning RSB on diluted ferromagnets is however ruled out here as we are approaching the critical line from above.

J. Stat. Mech. (2011) P02027

the technique introduced in [41], starting from the high temperature and, under the Gaussian ansatz for critical fluctuations, we want to take into account correlations between the order parameters. Within this approach, using the Wick theorem to split the four observable averages into a series of couples, the (formal) dynamical system reduces to

$$\dot{A}(t) = 2A(t)D(t), \tag{4.32}$$

$$\dot{D}(t) = A(t)G(t) + D^2(t), \tag{4.33}$$

$$\dot{G}(t) = 2G(t)D(t). \tag{4.34}$$

We must now solve for $A(t), D(t), G(t)$ and evaluate these expressions at $t = (1 + a/2)\sqrt{\beta}$ according to the definition of the dot operator in equation (4.31). Notice at first that

$$\partial_t \log A = \frac{\dot{A}}{A} = 2D = \frac{\dot{G}}{G} = \partial_t \log G.$$

This means that $\partial_t(A/G) = 0$ and, since $A(0)/G(0) = 1$, we already know that $A(t) = G(t)$: the fluctuations of the two order parameters behave in the same way, not surprisingly, as we have already pointed out their mutual interdependence several times. Therefore, let us focus on the remaining

$$\dot{D}(t) = G^2(t) + D^2(t), \tag{4.35}$$

$$\dot{G}(t) = 2G(t)D(t). \tag{4.36}$$

By defining $Y = D+G$ and summing the two equations above we immediately get $\dot{Y} = Y^2$, by which we find $Y(t) = Y(0)/(1 - tY(0))$. As $Y(0) = 1$ we can write

$$D(t = (1 + a)/2\sqrt{\beta}) + G(t = (1 + a)/2\sqrt{\beta}) = \frac{1}{1 - (1 + a/2)\sqrt{\beta}},$$

so there is a regular behavior up to $\beta_c = 1/(1 + a/2)^2$.

We must now solve separately for D and G: this is straightforward by introducing the function $Z = G^{-1}$ and checking that Z obeys

$$-\dot{Z} - 2YZ + 2 = 0,$$

which, once solved with standard techniques (as Y is known) gives $G(t) = [2(1 - t)]^{-1}$ and ultimately, simply by noticing the divergencies of $A(t = (1 + a)\sqrt{\beta}/2), D(t = (1 + a)\sqrt{\beta}/2), G(t = (1 + a)\sqrt{\beta}/2)$, we get the critical line for both the squared order parameters and their relative correlation: all these functions do diverge on the line

$$\beta_c = \frac{1}{(1 + a/2)^2} = \frac{1}{\bar{J}}, \tag{4.37}$$

defining a phase transition according with intuition.

4.5. $L \to \infty$ scaling in the thermodynamic limit

As we understood in section 3.3, in the $V \to \infty$ and $L \to \infty$ limits we need to tune the limit of $a \to -1$ carefully in order to recover the various interesting topologies and to avoid the trivial limits of fully connected/disconnected graphs.

J. Stat. Mech. (2011) P02027

In particular, a must approach -1 as $a = -1 + \gamma/V^{\theta}$. To tackle this scaling it is convenient to use directly γ, θ as tunable parameters and rewrite the Hamiltonian in the following forms[12]

$$H_V(\sigma;\xi) = \frac{1}{2\alpha V^{2(1-\theta)}} \sum_{ij}^{V} \sum_{\mu}^{L} \xi_i^{\mu}\xi_j^{\mu}\sigma_i\sigma_j \Rightarrow \tilde{H}_{V,L}(\sigma, z;\xi) = \frac{\sqrt{\beta/\alpha}}{V^{1-\theta}} \sum_{i,\mu} \xi_{i,\mu}\sigma_i z_{\mu}, \tag{4.38}$$

where the difference between the expressions for H and $\tilde{H}$ arises due to the Hubbard–Stratonovich transformation applied to the related partition functions, as performed earlier through equation (4.4).

Now, our free energy reads off as

$$A(\alpha,\beta,\gamma,\theta) = \lim_{V\to\infty} \frac{1}{V}\mathbb{E}\log \sum_{\sigma} \int \prod_{\mu}^{L} \mathrm{d}\mu\,(z_{\mu}) \exp\left(\frac{\sqrt{\beta/\alpha}}{V^{(1-\theta)}}\right) \sum_{i,\mu} \xi_{i\mu}\sigma_i z_{\mu}, \tag{4.39}$$

where α is the ratio between the size of the two parties, β is the noise felt by the system, θ distinguishes the topological class (see section 3.3) and γ allows a fine tuning over connectivity.

The interpolating scheme remains the same as before: we introduce the random fields and use $t \in (0,1)$ to define

$$A(t) = \frac{1}{N}\mathbb{E}\log \sum_{\sigma} \int \mathrm{d}\mu\,(z_{\mu}) \exp\left[t\tilde{H}_{V,L}(\sigma, z;\xi) + (1-t)\left(\sum_i \tilde{a}\eta_i\sigma_i + \sum_{\mu} \tilde{b}\eta_{\mu} z_{\mu}\right)\right]. \tag{4.40}$$

By performing the t-streaming we get

$$\partial_t A(t) = \frac{\sqrt{\beta}\gamma}{2}\langle (M-\bar{M})(N-\bar{N})\rangle + \frac{\sqrt{\beta}\gamma}{2}\bar{M}\bar{N}, \tag{4.41}$$

such that the replica symmetric sum rule becomes

$$A(1) = A(0) - \frac{\sqrt{\beta}\gamma}{2}\bar{M}\bar{N}, \tag{4.42}$$

and the replica symmetric free energy reads as

$$A(\beta,\gamma,\theta) = \log 2 + \frac{\gamma}{2V^{\theta}}\log\cosh(\sqrt{\beta}\bar{N}V^{\theta}) + \frac{\beta\gamma^2}{8}\bar{M}^2 - \frac{\sqrt{\beta}\gamma}{2}\bar{M}\bar{N}. \tag{4.43}$$

Let us now investigate some limits of this expression and its self-consistency. Note that by extremizing with respect to the order parameters we can express $\bar{N}$ from $\bar{M}$ as

$$\langle \bar{N}\rangle = \frac{\sqrt{\beta}\gamma}{2}\langle \bar{M}\rangle. \tag{4.44}$$

[12] As we are going to see soon there does not exist a unique normalization for the Hamiltonian, able to account simultaneously for both the coupling strength and the volume extensiveness for all the possible graphs. We choose to normalize $H_V(\sigma;\xi)$ in this way in order to tackle immediately the basic limits; apparent divergencies in the couplings can develop and standardly canceled out by properly rescaling the temperature corresponding to the number of nearest neighbors, as in more classical approaches.

J. Stat. Mech. (2011) P02027

4.5.1. $\theta = 0$ case: fully connected, weighted and Curie–Weiss scenario. The case $\theta = 0$ corresponds to a fully connected graph, and in particular when γ reaches its upper bound two, disorder on couplings vanishes and we recover the unweighted CW model (see section 3.3). Now, by setting $\theta = 0$, from equation (4.43) we have

$$A(\beta, \gamma, \theta = 0) = \log 2 + \frac{\gamma}{2} \log \cosh \left(\beta \frac{\gamma}{2} \langle \bar{M} \rangle \right) - \frac{\beta \gamma^2}{8} \langle \bar{M} \rangle^2, \tag{4.45}$$

from which the self-consistency relation follows as

$$\langle \bar{M} \rangle = \tanh \left(\beta \frac{\gamma}{2} \langle \bar{M} \rangle \right).$$

This holds generally for the weighted graph, while when $\gamma = 2$ the (normalized) coupling is constant and equal to one and we get straightforwardly the standard CW limit:

$$A(\beta, \gamma = 2, \theta = 0) = \log 2 + \log \cosh(\beta \langle \bar{M} \rangle) - \beta 2 \langle \bar{M} \rangle^2, \tag{4.46}$$

$$\langle \bar{M} \rangle = \tanh(\beta \langle \bar{M} \rangle). \tag{4.47}$$

4.5.2. $\theta = 1/2$: standard dilution regime. With a scheme perfectly coherent with the previous one we can write down the free energy and relevant self-consistency as

$$A(\beta, \gamma, \theta = 1/2) = \lim_{V \to \infty} \left(\log 2 + \frac{\gamma}{2\sqrt{V}} \log \cosh \left(\frac{\beta \gamma}{2} \sqrt{V} \langle \bar{M} \rangle \right) - \frac{\beta \gamma^2}{8} \langle \bar{M} \rangle^2 \right), \tag{4.48}$$

$$\langle \bar{M} \rangle = \lim_{V \to \infty} \tanh \left(\frac{\beta \gamma}{2} \sqrt{V} \langle \bar{M} \rangle \right). \tag{4.49}$$

Interestingly, we notice that, as $\sqrt{J} = \gamma / 2\sqrt{V}$, the argument of the logarithm of the hyperbolic cosine scales as $\sqrt{J} V \langle \bar{M} \rangle$: this is consistent with the lack of a proper normalization into the Hamiltonian (4.38), because for $\theta = 1/2$ the latter is still divided by V which should not appear. To avoid the lack of a universal normalization, we need to renormalize the local average coupling by a factor V so that we get the correct behavior, namely we write explicitly the free energy, putting in evidence that $p \sim 1 - \exp(-\alpha \gamma^2 / 4)$:

$$A(\tilde{\beta}, \gamma, \theta = 1/2) = \log 2 + \sqrt{J} \log \cosh \left(\frac{\tilde{\beta} \sqrt{J}}{p} \langle \bar{M} \rangle \right) - \frac{\tilde{\beta} J}{2p} \langle \bar{M} \rangle^2, \tag{4.50}$$

such that, having $\tilde{\beta} = \beta p V$ [2], we can easily recover the trivial limits of the CW case when $p \to 1$ (and coherently $J \to 1$, $\tilde{\beta} \to \beta$) and of the fully disconnected network $p \to 0 \Rightarrow A(\tilde{\beta}, \gamma \to 0, \theta = 1/2) = \log 2$ as p is superlinear in γ.

4.5.3. $\theta = 1$: finite connectivity regime. Finally, when $\theta = 1$ the free energy and its coupled self-consistency are

$$A(\beta, \gamma, \theta = 1) = \lim_{V \to \infty} \left(\log 2 + \frac{\gamma}{2V} \log \cosh \left(\frac{\beta \gamma}{2} V \langle \bar{M} \rangle \right) - \frac{\beta \gamma^2}{8} \langle \bar{M} \rangle^2 \right), \tag{4.51}$$

$$\langle \bar{M} \rangle = \lim_{V \to \infty} \tanh \left(\frac{\beta \gamma}{2} V \langle \bar{M} \rangle \right). \tag{4.52}$$

J. Stat. Mech. (2011) P02027

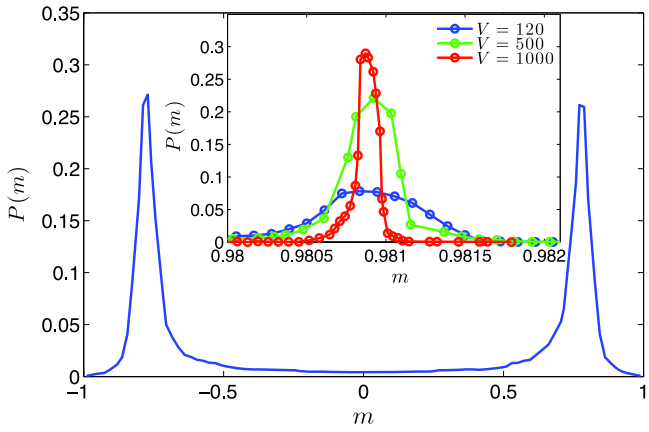

Figure 4.1. Probability distribution for the order parameter $P(m)$. Main figure: system of size $N = 1000$ at a temperature $T = 0.5T_c$; inset: comparison between systems of size $N = 120$, $N = 500$ and $N = 1000$, as shown by the legend, set at a temperature $T = 0.9T_c$.

Of course here, with respect to the previous case, we get even stronger divergencies. Now we need to renormalize the local average coupling by a factor V^2.

4.6. Numerics: probability distribution

As the critical line is obtained, in the fluctuation theory, through the Gaussian ansatz, we double check our finding via numerical simulations.

First of all, we notice that since the interaction matrix J_{ij} is symmetric ($J_{ij} = J_{ji}$), detailed balance holds and it is well known [57, 3] how to introduce a Markov process for the dynamical evolution ruled by Hamiltonian (2.2) and obtain the transition rates for stationarity: Monte Carlo sampling is then meaningful for equilibrium investigation.

The order parameter distribution function has been proved to be a powerful tool for studying the critical line in different kinds of systems; in particular, for magnetic systems, the order parameter can be chosen as the magnetization per spin which, in finite-size systems, is a fluctuating quantity characterized by a probability distribution $P(m)$ [24]. In Ising-like models undergoing a second-order phase transition it is known that at temperatures lower than the critical temperature β_c^{-1}, the distribution $P(m)$ has a double peak, centered at the spontaneous magnetization $+m$ and $-m$. At temperatures greater than β_c^{-1}, $P(m)$ has a single peak at zero magnetization, and exactly at β_c^{-1} a double peak shape is observed.

In figure 4.1 we plot numerical data for the probability distribution, obtained by means of Monte Carlo simulations, where $P(m)$ corresponds to the fraction of the total number of realizations in which the system magnetization is m. In the main figure we show the distribution for a system with $V = 1000$ set at a temperature $\beta^{-1} = 1.1\beta_c^{-1}$, while in the inset we compare systems of different sizes set at a temperature $\beta^{-1} = 2\beta_c^{-1}$. Notice that, for such small temperatures, as the size is increased the distribution is more and more peaked, while the probability to have zero magnetization is vanishing; this corroborates the replica symmetric ansatz.

5. Conclusion

In this paper we introduced an alternative way to generate systems of mutually interacting components; the approach is inspired by the Hopfield model, where we shift the support of the patterns like $[-1, 1] \to [0, 1]$.

From a graph theory perspective we introduced a model which, given a set of V nodes, each corresponding to a set of L attributes encoded by a binary string ξ, defines an interaction coupling J_{ij} for any couple of nodes (i, j). The resulting system can be envisaged by means of a weighted graph displaying non-trivial coupling distributions and correlations between links. In particular, when attributes are extracted according to a discrete uniform distribution, i.e. $P(\xi_i^\mu) = (1 + a)/2$ for any $i \in [1, V]$ and $\mu \in [1, L]$, a being a tunable parameter, we get that when a is sufficiently small the resulting network exhibits a small-world nature, namely a large clustering coefficient; as a is varied, the network behaves as an isolated spin system, an extreme dilute network, a linearly diverging connectivity network, a weighted fully connected network and an unweighted fully connected network, respectively. Moreover, nodes are topologically distinguishable according to the concentration ρ of non-null entries present in their corresponding binary strings: interestingly, if the scaling between L and V is sublinear (e.g. $P \propto \ln V$ or even slower, as in low storage networks [1, 15]) the degree distribution turns out to be multi-modal, each mode corresponding to a different value of ρ. Conversely, whenever the scaling is (at least) linear (i.e. $L \propto V$), the distribution becomes mono-modal.

Then, as diluted models are of primary interest in disordered statistical mechanics, by assuming self-averaging of the order parameters, we solved the thermodynamics of the model: this required a new technique (a generalization to infinitely random fields of the double stochastic stability) which is of complete generality and paves another way for approaching dilution in complex systems.

Furthermore, within this framework, replicas are not necessary, and instead of averaging over these copies of the system (and dealing with the corresponding overlaps) we can obtain observables as magnetization averages over local subgraphs, implicitly accounting for a replica symmetric behavior (which is indeed assumed throughout the study).

An interesting point on which graph and statistical mechanics investigations converge concerns a peculiar non-mean-field effect in the overall fields felt by the spins: the field insisting on a spin scales as $\sqrt{J}$ (see equation (3.17)), while the averaged field on the network scales as J (see equation (3.20)), which corresponds to the canonical mean-field expectation. Furthermore, looking at equation (4.24) we see that in the hyperbolic tangent encoding the response of the spin to the fields, the contribution of the other spins is not weighted by J but by $\sqrt{J}$. As in the thermodynamics the coupling strength has been normalized, $J < 1 \to \sqrt{J} > J$: in complex thermodynamics there is a superlinearity between the interactions: however this does not affect the critical behavior, which is a global feature of the system and is consequently found to scale with J (see equation (4.36)), this may substantially change all the other speculations based on intuition.

Of course, in the Curie–Weiss limit this effect disappears as global and local environments do coincide (i.e. $J = 1$).

It is worth stressing that (microscopic) correlation among bit strings is directly related to the macroscopic behavior (e.g. critical line), suggesting an alternative, intriguing

J. Stat. Mech. (2011) P02027

approach of investigation in e.g. social networks, gene regulatory networks, or immune networks: the inner degrees of freedom can be inferred from the overall properties.

The next steps, which could stem from our approach, could address the clear statistical mechanics of scale-free networks and applications to real systems (first of all a clear investigation on dynamical retrieval properties), both in biology and in sociology.

Acknowledgments

Francesco Guerra, as usual, is acknowledged for priceless scientific and human interchange. This work is supported by the FIRB grant: $RBFR08EKEV$. AB is grateful to the Smart-Life Grant for partial support. INFN and GNFM are acknowledged too for their partial support.

References

[1] Agliari E and Barra A, *Statistical mechanics of idiotypic immune networks*, 2010 submitted
[2] Agliari E, Barra A and Camboni F, *Criticality in diluted ferromagnets*, 2008 *J. Stat. Mech.* P10003
[3] Agliari E, Barra A, Burioni R and Contucci P, *New perspectives in the equilibrium statistical mechanics approach to social and economic sciences*, 2010 *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences* (Basel: Birkhauser)
[4] Agliari E, Burioni R, Cassi D and Neri F M, *Autocatalytic reactions on low-dimensional structures*, 2007 *Thoer. Chem. Acc* **118** 855
[5] Agliari E, Blumen A and Mülken O, *Dynamics of continuous-time quantum walks in restricted geometries*, 2008 *J. Phys. A: Math. Theor.* **41** 445301
[6] Agliari E, Casartelli M and Vezzani A, *Microscopic energy flows in disordered Ising spin systems*, 2010 *J. Stat. Mech.* P10021
[7] Agliari E, Casartelli M and Vivo E, *Metric characterization of cluster dynamics on the Sierpinski gasket*, 2010 *J. Stat. Mech.* 09002
[8] Agliari E, Cioli C and Guadagnini E, *Percolation on correlated random networks*, submitted
[9] Amit D J, 1992 *Modeling Brain Function: The World of Attractor Neural Network* (Cambridge: Cambridge Univerisity Press)
[10] Amit D J, Gutfreund H and Sompolinsky H, *Storing infinite numbers of patterns in a spin glass model of neural networks*, 1985 *Phys. Rev. Lett.* **55** 1530
[11] Amit D J, Gutfreund H and Sompolinsky H, *Information storage in neural networks with low levels of activiy*, 1987 *Phys. Rev.* A **35** 2293
[12] Albert R and Barabasi A L, *Statistical mechanics of complex networks*, 2002 *Rev. Mod. Phys.* **74** 47 and references therein
[13] Barabasi A L and Oltvai Z N, *Network biology: understanding the cell's functional organization*, 2004 *Rev. Nat. Genetics* **5** 101
[14] Barra A, *Irreducible free energy expansion and overlap locking in mean field spin glasses*, 2006 *J. Stat. Phys.* **123** 601
[15] Barra A and Agliari E, *Autopoietic immune networks from a statistical mechanics perspective*, 2010 *J. Stat. Mech.* P07004
[16] Barra A and Agliari E, *Stochastic dynamics for idiotypic immune networks*, 2010 *Physica* A **389** 5903
[17] Barra A and Agliari E, *A statistical mechanics approach to Granovetter theory*, 2010 arXiv:1012.1272 submitted
[18] Barra A and Contucci P, *Toward a quantitative approach to migrants integration*, 2010 *Europhys. Lett.* **89** 68001
[19] Barra A and Guerra F, *About the ergodicity in Hopfield analogical neural network*, 2008 *J. Math. Phys.* **49** 125217
[20] Barra A, Guerra F and Genovese G, *The replica symmetric behavior of the analogical neural network*, 2010 *J. Stat. Phys.* **140** 784
[21] Barrat A and Weigt M, *On the properties of small-world network models*, 2000 *Eur. Phys. J.* B **13** 547
[22] Barrat A, Barthelemy M and Vespignani A, 2008 *Dynamical Processes in Complex Networks* (Cambridge: Cambridge University Press)

J. Stat. Mech. (2011) P02027

[23] ben-Avraham D and Havlin S, 2005 *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge: Cambridge University Press)
[24] Binder K and Landau D P, 2009 *Guide To Monte Carlo Simulations In Statistical Physics* (Cambridge: Cambridge University Press)
[25] Blake C C F, *Do genes in pieces imply proteins in pieces?*, 1978 *Nature* **273** 267
[26] Bollobas B, 1985 *Random graphs* (*Cambridge Studies in Advanced Mathematics*) (Cambridge: Cambridge University Press)
[27] Burioni R and Cassi D, *Random walks on graphs: ideas, techniques, results*, 2005 *J. Phys. A: Math. Gen.* **38** R45
[28] Brock W and Durlauf S, *Discrete choices with social interactions*, 2001 *Rev. Econ. Stud.* **68** 235
[29] Buchanan M, Caldarelli G, De Los Rios P, Rao F and Vendruscolo M (ed), 2010 *Modelling Cell Biology With Networks* (Cambridge: Cambridge University Press)
[30] Burnet F M, 1959 *The Clonal Selection Theory of Acquired Immunity* (Cambridge: Cambridge University Press)
[31] Coolen A C C and Rabello S, *Generating functional analysis of complex formation and dissociation in large protein interaction networks*, 2009 *J. Phys. Conf. Ser.* **197** 012006
[32] Dembo A and Montanari A, *Gibbs measures and phase transitions on sparse random graphs*, 2008 *Course Taught at the 2008 Brazilian School of Probability*
[33] Durlauf S N, *How can statistical mechanics contribute to social science?*, 1999 *Proc. Nat. Acad. Sci.* **96** 10582
[34] Ellis R S, 1985 *Large Deviations and Statistical Mechanics* (New York: Springer)
[35] Francke C, Siezen R J and Teusink B, *Reconstructing the metabolic network of a bacterium from its genome*, 2005 *Trends Microbiol.* **13** 550
[36] De Sanctis L and Guerra F, *Mean field dilute ferromagnet I. High temperature and zero temperature behavior*, 2008 *J. Stat. Phys.* **132** 759
[37] Gallo I, Barra A and Contucci P, *Parameter evaluation of a simple mean-field model of social interaction*, 2008 *Math. Methods Mod. Appl. Sci.* **19** 1427
[38] Gavin A C *et al*, *Functional organization of the yeast proteome by systematic analysis of protein complexes*, 2002 *Nature* **415** 141
[39] Genovese G and Barra A, *A mechanical approach to mean field spin models*, 2009 *J. Math. Phys.* **50** 053303
[40] Genovese G and Barra A, *A certain class of Curie–Weiss models*, 2009 arXiv:0906.4673
[41] Guerra F, *Sum rules for the free energy in the mean field spin glass model*, 2001 *Mathematical Physics in Mathematics and Physics: Quantum and Operator Algebraic Aspects* (*Fields Institute Communications* vol 30) (Providence, RI: American Mathematical Society)
[42] Granovetter M S, *The strength of weak ties*, 1973 *Am. J. Sociol.* **78** 1360
[43] Granovetter M S, *The strength of the weak tie: revisited*, 1983 *Sociol. Theory* **1** 201
[44] Berlow E L, *Strong effects of weak interactions in ecological communities*, 1999 *Lett. Nat.* **398** 330
[45] Hebb D O, 1949 *Organization of Behaviour* (New York: Wiley)
[46] Hopfield J J, *Neural networks and physical systems with emergent collective computational abilities*, 1982 *Proc. Nat. Acad. Sci.* **79** 2554
[47] Jerne N K, *Toward a network theory of the immune system*, 1974 *Ann. Immun.* **125** 373
[48] Martelli C, De Martino A, Marinari E, Marsili M and Perez-Castillo I, *Identifying essential genes in E. coli from a metabolic optimization principle*, 2009 *Proc. Nat. Acad. Sci.* **106** 2607
[49] McFadden D, *Economic choices*, 2001 *Am. Econ. Rev.* **91** 351
[50] Mézard M, Parisi G and Virasoro M A, 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[51] Milgram S, *The small world problem*, 1967 *Psychol. Today* **2** 60
[52] Montoya J M and Sole' R V, *Small world patterns in food webs*, 2002 *J. Theor. Biol.* **214** 3
[53] Nikoletopoulos T, Coolen A C C, Perez-Castillo I, Skantzos N S, Hatchett J P L and Wemmenhove B, *Replicated transfer matrix analysis of ising spin models on 'small world' lattices*, 2004 *J. Phys. A: Math. Gen.* **37** 6455
[54] Parisi G, *A simple model for the immune network*, 1990 *Proc. Nat. Acad. Sci.* **87** 429
[55] Parisi G, 1988 *Statistical Field Theory* (*Frontiers in Physics*) (Reading, MA: Addison-Wesley)
[56] Peliti L, *Introduction to the statistical theory of Darwinian evolution*, 1997 *Lectures at the Summer College on Frustrated System (Trieste, August)*
[57] Perez-Castillo I, Wemmenhove B, Hatchett J P L, Coolen A C C, Skantzos N S and Nikoletopoulos T, *Analytic Solution of Attractor Neural Networks on Scale-free Graphs*, 2004 *J. Phys. A: Math. Gen.* **37** 8789

J. Stat. Mech. (2011) P02027

[58] Rabello S, Coolen A C C, Perez-Vicente C J and Fraternali F, *A solvable model of the genesis of amino-acid sequences via coupled dynamics of folding and slow genetic variation*, 2008 *J. Phys. A: Math. Theor.* **41** 285004
[59] Watts D J and Strogatz S H, *Collective dynamics of small world networks*, 1998 *Nature* **393** 440
[60] Newman M E J, *The structure and function of complex networks*, 2003 *SIAM Rev.* **45** 167 and references therein
[61] Watkin T L H and Sherrington D, 1991 *Europhys. Lett.* **14** 791
[62] Wemmenhove B, Skantzos N S and Coolen A C C, 2004 *J. Phys. A: Math. Gen.* **37** 7653
[63] Said M R, Begley T J, Oppenheim A V, Lauffenberger D A and Samson L D, 2004 *Proc. Nat. Am. Soc.* **101** 18006
[64] Uetz P, Dong Y-A, Zeretzke C, Atzler C, Baiker A, Berger B, Rajagopala S V, Roupelieva M, Rose D, Fossum E and Haas J, 2006 *Science* **311** 239
[65] Breskin I, Soriano J, Moses E and Tlusty T, 2006 *Phys. Rev. Lett.* **97** 188102

J. Stat. Mech. (2011) P02027