Extensive Parallel Processing on Scale-Free Networks

Peter Sollich,[1] Daniele Tantari,[2] Alessia Annibale,[1,3] and Adriano Barra[4]
[1]Department of Mathematics, King's College London, Strand, London WC2R 2LS, United Kingdom
[2]Dipartimento di Matematica, Sapienza Università di Roma, Piazzale Aldo Moro 2, Roma 00185, Italy
[3]Institute for Mathematical and Molecular Biomedicine, King's College London, Hodgkin Building, London SE1 1UL, United Kingdom
[4]Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, Roma 00185, Italy
(Received 13 April 2014; revised manuscript received 10 November 2014; published 5 December 2014)

We adapt belief-propagation techniques to study the equilibrium behavior of a bipartite spin glass, with interactions between two sets of N and $P = \alpha N$ spins each having an arbitrary degree, i.e., number of interaction partners in the opposite set. An equivalent view is then of a system of N neurons storing P diluted patterns via Hebbian learning, in the high storage regime. Our method allows analysis of parallel pattern processing on a broad class of graphs, including those with pattern asymmetry and heterogeneous dilution; previous replica approaches assumed homogeneity. We show that in a large part of the parameter space of noise, dilution, and storage load, delimited by a critical surface, the network behaves as an extensive parallel processor, retrieving all P patterns in parallel without falling into spurious states due to pattern cross talk, as would be typical of the structural glassiness built into the network. Parallel extensive retrieval is more robust for homogeneous degree distributions, and is not disrupted by asymmetric pattern distributions. For scale-free pattern degree distributions, Hebbian learning induces modularity in the neural network; thus, our Letter gives the first theoretical description for extensive information processing on modular and scale-free networks.

DOI: 10.1103/PhysRevLett.113.238106 PACS numbers: 84.35.+i, 87.19.ll, 89.75.Hc

Since their introduction in the pioneering paper by Amit *et al.* [1,2], associative neural networks (NN) have played a central role in the statistical mechanics community, soon becoming one of the most successful offshoots of spin glasses (SG) [3], as proven by the celebrated paper by Hopfield [4] or the excellent book by Amit [5]. Indeed, the Hopfield model for NN [4] can be regarded as a SG where the coupling between each pair of spins σ_i, σ_j, $i, j = 1, \ldots, N$ has the Hebbian form $J_{ij} = N^{-1}\sum_\mu \xi_i^\mu \xi_j^\mu$ and the $\boldsymbol{\xi}^\mu$, $\mu \in \{1, \ldots, P = \alpha N\}$, represent stored patterns with entries $\xi_i^\mu = \pm 1$ distributed as $P(\xi_i^\mu = \pm 1) = 1/2$.

A further connection between NN and SG has been pointed out recently in the context of bipartite SG [6]. Here, we illustrate the connection for a system of two sets of spins, σ_i, $i = 1, \ldots, N$ and τ_μ, $\mu = 1, \ldots, P$, connected by links $\xi_i^\mu = \pm 1$ that are sparse, so that $P(\xi_i^\mu = \pm 1) = c/2N$ and $P(\xi_i^\mu = 0) = 1 - c/N$ with $c = \mathcal{O}(N^0)$, and described by the SG Hamiltonian $H_{\rm SG}(\boldsymbol{\sigma}, \boldsymbol{\tau}|\boldsymbol{\xi}) \propto -\sum_{i,\mu} \xi_i^\mu \sigma_i \tau_\mu$. Marginalizing over $\boldsymbol{\tau}$ in the partition function $Z = \sum_{\boldsymbol{\sigma},\boldsymbol{\tau}} e^{-\beta H_{\rm SG}(\boldsymbol{\sigma},\boldsymbol{\tau}|\boldsymbol{\xi})} = \sum_{\boldsymbol{\sigma}} e^{-\beta H_{\rm NN}(\boldsymbol{\sigma}|\boldsymbol{\xi})}$ shows that the $\boldsymbol{\sigma}$ represent a NN with Hamiltonian $H_{\rm NN}(\boldsymbol{\sigma}|\boldsymbol{\xi}) = -\beta^{-1}\sum_\mu \ln \times [2\cosh(\beta \sum_i \xi_i^\mu \sigma_i)]$ or, up to an additive constant, $H_{\rm NN}(\boldsymbol{\sigma}|\boldsymbol{\xi}) = -(\beta/2)\sum_{\mu,i,j}(\xi_i^\mu \xi_j^\mu)\sigma_i \sigma_j + \ldots$. Higher-order interactions are not written explicitly here; these are fully absent if the τ_i are continuous rather than discrete and have a Gaussian prior. Remarkably, while standard NN retrieve patterns sequentially (one at time), associative networks with diluted patterns are able to accomplish parallel retrieval in appropriate dilution regimes [7–9]. However, existing studies, which use replica analysis, apply only to homogeneously diluted networks where the degrees in each set of spins have a Poisson distribution. Here we adapt cavity (i.e. belief-propagation) methods to analyze the significantly more general scenario where pattern asymmetry is permitted and degrees in the two sets of spins have arbitrary distributions, thus covering information processing even on scale-free networks [10,11]. Remarkably, Hebbian interactions generated by patterns with such scale-free degree distribution give rise to networks of neurons which are both scale free and contain effective modules, namely highly clustered quasiautonomous communities (see Fig. 1). Their existence, in turn, is crucial at the information-processing level, because it is the basis of extensive parallel retrieval of multiple patterns accomplished by the network. This is in agreement with experimental findings on intracellular protein networks, which have scale-free degrees, but where interactions among hubs are strongly suppressed in favor of a modular structure, which minimizes cross talk among different modules [12,13].

To confirm this scenario, we consider an equilibrated system of N binary neurons $\sigma_i = \pm 1$ at temperature (fast noise) $T = 1/\beta$, with Hamiltonian

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

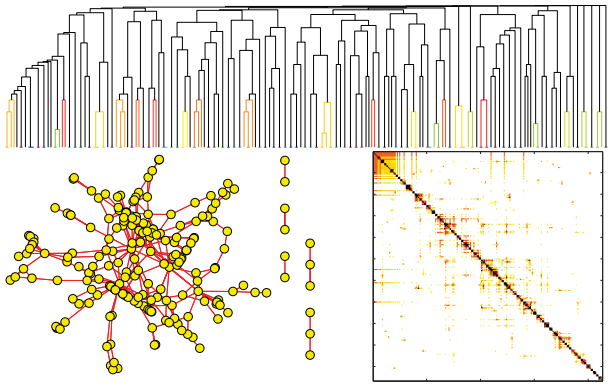

FIG. 1 (color online). (Left) Example of a Hebbian scale-free network of $N = 500$ neurons and $P = \alpha N$ patterns with $\alpha = 3$. Pattern degrees are distributed as $P(e) \propto e^{-\gamma}$ for $e \geq 1$, with $\gamma = 2.2$. (Right) The network structure is modular as shown by the generalized topological overlap matrix [9,10], with (top) modules arranged in a hierarchical fashion. See [14] for further information on network generation.

where pattern entries $\{\xi_i^\mu\}$ are sparse (i.e., the number of nonzero entries of any pattern is finite). We can then use a factor graph representation of the Boltzmann weight as $\prod_\mu F_\mu$, with factors

$$F_\mu = e^{(\beta/2)\sum_{i,j\in O(\mu)}\xi_i^\mu\xi_j^\mu\sigma_i\sigma_j} = \left\langle e^{z\sum_{i\in O(\mu)}\xi_i^\mu\sigma_i}\right\rangle_z, \quad (1)$$

where $O(\mu) = \{i : \xi_i^\mu \neq 0\}$ and z is a zero mean Gaussian variable with variance β [19]. We denote by $e_\mu = |O(\mu)|$ the degree of a pattern μ and by $d_i = |N(i)|$ the degree of a neuron i, with $N(i) = \{\mu : \xi_i^\mu \neq 0\}$. We consider random graph ensembles with given degree distributions $P(d)$ and $P(e)$, and nonzero ξ's independently and identically distributed (IID). Conservation of links demands $N\langle d\rangle = P\langle e\rangle$ where averages are taken over $P(d)$ and $P(e)$. The message from factor μ to node j is the cavity distribution $P_\mu(\sigma_j)$ of σ_j when this is coupled to factor μ only, which we can parameterize by an effective field $\psi_{\mu\to j}$. The message from node j to factor μ is the cavity distribution $P_{\backslash\mu}(\sigma_j)$ of σ_j when coupled to all factors except μ, which we can parameterize by the field $\phi_{j\to\mu}$. The cavity equations are then [20]

$$P_\mu(\sigma_j) = \mathrm{Tr}_{\{\sigma_k\}} F_\mu(\sigma_j, \{\sigma_k\}) \prod_{k\in O(\mu)\backslash j} P_{\backslash\mu}(\sigma_k), \quad (2)$$

$$P_{\backslash\nu}(\sigma_j) = \prod_{\mu\in N(j)\backslash\nu} P_\mu(\sigma_j). \quad (3)$$

Given the site factorization, conditional on z, of the factors (1), translating these equations into ones for the effective fields is straightforward,

$$\psi_{\mu\to j} = \tanh^{-1}\langle\sigma_j\rangle_\mu = \tanh^{-1}\frac{\langle\sinh(z\xi_j^\mu)\prod_{k\in O(\mu)\backslash j}\cosh(\phi_{k\to\mu} + z\xi_k^\mu)\rangle_z}{\langle\cosh(z\xi_j^\mu)\prod_{k\in O(\mu)\backslash j}\cosh(\phi_{k\to\mu} + z\xi_k^\mu)\rangle_z}, \quad (4)$$

$$\phi_{j\to\nu} = \sum_{\mu\in N(j)\backslash\nu} \psi_{\mu\to j}. \quad (5)$$

These equations, once iterated to convergence, are exact on tree graphs. They will also become exact on graphs sampled from our ensemble in the thermodynamic limit, because the sparsity of the ξ_i^μ makes the graphs locally treelike, with typical loop lengths that diverge logarithmically [14] with N [18,21].

For large N, we can describe the solution of the cavity equations in terms of the distribution of messages or fields, $W_\psi(\psi)$ and $W_\phi(\phi)$. Denoting by $\Psi(\{\phi_{k\to\mu}\}, \{\xi_k^\mu\}, \xi_j^\mu)$ the rhs of Eq. (4), convergence of the cavity iterations then implies the self-consistency equation (see, e.g., [22])

$$W_\psi(\psi) = \sum_e [eP(e)/\langle e\rangle] \times \langle\delta(\psi - \Psi(\phi_1, \ldots, \phi_{e-1}, \xi^1, \ldots, \xi^e))\rangle,$$

where the average is over IID values of the (nonzero) $\xi^1, \ldots, \xi^d$ and over IID $\phi_1, \ldots, \phi_{e-1}$ drawn from $W_\phi(\phi)$, and similarly

$$W_\phi(\phi) = \sum_d [dP(d)/\langle d\rangle]\left\langle\delta\left(\phi - \sum_{\mu=1}^{d-1}\psi_\mu\right)\right\rangle,$$

where the average is over IID $\psi_1, \ldots, \psi_{d-1}$ drawn from $W_\psi(\psi)$. Field distributions can then be obtained numerically by population dynamics (PD) [20]. For symmetric ξ distributions, a delta function at the origin for both W_ψ, W_ϕ is always a solution, and we find this to be stable at low β. At high β, on the other hand, the ψ can become large (see Fig. 2), hence also the ϕ, and the spins σ_i will typically be

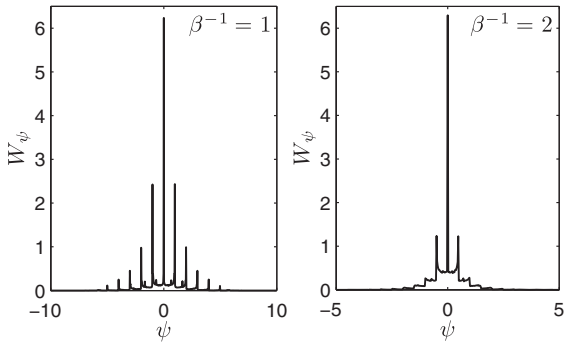

FIG. 2. Histograms $W_\psi(\psi)$ of the field ψ for $\alpha = 8$, $c = 2$, and $\beta^{-1} = 1, 2$, as shown in the figure.

strongly polarized. The fields $\beta\xi_i^\mu\sum_{j\in O(\mu)\setminus i}\xi_j^\mu\sigma_j$ then fluctuate little, and the ψ as suitable averages of these fields cluster near multiples of β (for $\xi=\pm1$).

Our main interest is in the retrieval properties, encoded in the fluctuating pattern overlaps $m_\mu=\sum_{i\in O(\mu)}\xi_i^\mu\sigma_i$. Since the joint distribution of the σ_i in $O(\mu)$ is $F_\mu(\{\sigma_i\})\prod_{i\in O(\mu)}P_{\setminus\mu}(\sigma_i)$, the distribution of the pattern overlap m_μ is

$$\left(\mathrm{Tr}_{\{\sigma_i\}}\left\langle\delta(m_\mu-m)\exp\left[\sum_{i\in O(\mu)}(\xi_i^\mu z+\phi_{i\to\mu})\sigma_i\right]\right\rangle_z\right)\Big/\left(\mathrm{Tr}_{\{\sigma_i\}}\left\langle\exp\left[\sum_{i\in O(\mu)}(\xi_i^\mu z+\phi_{i\to\mu})\sigma_i\right]\right\rangle_z\right). \quad (6)$$

Defining this as $\mathcal{P}(m,\{\phi_{i\to\mu}\},\{\xi_i^\mu\})$, in the graph ensemble we have

$$P(m)=\sum_e P(e)\langle\mathcal{P}(m,\phi_1,\ldots,\phi_e,\xi_1,\ldots,\xi_e)\rangle. \quad (7)$$

The average here can be read as $P(m|e)$, the overlap distribution for patterns with fixed degree e. Whenever $W_\phi(\phi)=\delta(\phi)$, $P(m|e)$ is the overlap distribution for an "effectively isolated" subsystem of size e: the neurons storing each pattern $\boldsymbol{\xi}^\mu$ can retrieve this independently of other patterns, even though the number of patterns is extensive. Retrieval within each group of neurons is strongest at low temperatures [see Fig. 3 (left)] as expected on general grounds. Once nonzero ϕ appear neuron groups are no longer independent and cross talk interference between patterns emerges.

Bifurcation.—When the "parallel processor" solution with zero cavity fields ϕ becomes unstable, a bifurcation to a different stable solution occurs. Depending on the external parameters, this can be seen in the first or second moment of the field distribution. Expanding for small fields we get

$$\Psi(\{\phi_{k\to\mu}\},\{\xi_k^\mu\},\xi_j^\mu)\approx\sum_{k\in O(\mu)\setminus j}\phi_{k\to\mu}\Xi(\xi_k^\mu,\xi_j^\mu,\{\xi_l^\mu\}),$$

with coefficients $\Xi(\xi_k^\mu,\xi_j^\mu,\{\xi_l^\mu\})$ given by

$$\left\langle\sinh(z\xi_j^\mu)\sinh(z\xi_k^\mu)\prod_{l\in O(\mu)\setminus\{j,k\}}\cosh(z\xi_l^\mu)\right\rangle_z\Big/\left\langle\prod_{l\in O(\mu)}\cosh(z\xi_l^\mu)\right\rangle_z.$$

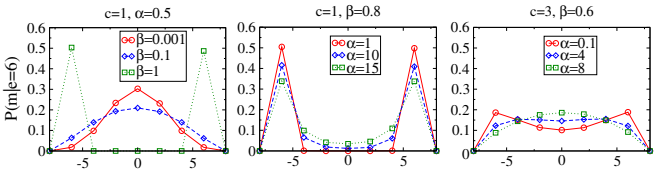

FIG. 3 (color online). $P(m|e=6)$ above (left) and crossing (middle and right) the critical line for different values of β and α, respectively. Full red (dashed blue and dotted green) curves in the middle and right panels refer to temperatures above (below) the critical line.

The self-consistency relations for the field distributions W_ψ and W_ϕ then show that as long as the mean fields are small, they are related to leading order by

$$\langle\psi\rangle=\langle\phi\rangle\sum_e P(e)[e(e-1)/\langle e\rangle]\langle\Xi(\xi_1,\ldots,\xi_e)\rangle, \quad (8)$$

$$\langle\phi\rangle=B_d\langle\psi\rangle, \quad (9)$$

where $B_d=\sum_d P(d)d(d-1)/\langle d\rangle$ is one of the two branching ratios of our locally treelike graphs, the other being $B_e=\sum_e P(e)e(e-1)/\langle e\rangle$. If the means are zero then the onset of nonzero fields is detected by the variances, which are related to leading order by

$$\langle\psi^2\rangle=\langle\phi^2\rangle\sum_e P(e)[e(e-1)/\langle e\rangle]\langle\Xi^2(\xi_1,\ldots,\xi_e)\rangle, \quad (10)$$

$$\langle\phi^2\rangle=B_d\langle\psi^2\rangle. \quad (11)$$

Symmetric pattern distributions.—When the ξ are symmetrically distributed, then the field distributions are also always symmetric and there can be no instability from growing means, cf. Eq. (8). The bifurcation has to result from the growth of the variances, which from Eq. (11) occurs at $A=1$ with

$$A=B_d\sum_e P(e)[e(e-1)/\langle e\rangle]\langle\Xi^2(\xi_1,\ldots,\xi_e)\rangle. \quad (12)$$

This factorizes as $A=B_dA_e(\beta)$ with the dependence on the noise and on the topology (i.e., the distribution of the e's) contained in the second factor $A_e(\beta)$. For $\beta\to0$ the variance of z goes to zero and $A_e(0)=0$. For $\beta\to\infty$, the z averages are dominated by large values of z where $\sinh^2(z)\approx\cosh^2(z)$, so $A_e(\infty)=B_e$. Hence there is no bifurcation when $B_dB_e<1$, in agreement with the general percolation condition for bipartite trees [23].

The key advantage of our method is that we can easily investigate the parallel processing capabilities of a bipartite graph with arbitrary degrees $\{e_\mu\}$. Here we have a pattern-dependent dilution of the links $P(\boldsymbol{\xi})\propto\prod_{i,\mu}P(\xi_i^\mu)\prod_\mu\delta_{e_\mu,\sum_i|\xi_i^\mu|}$ with

$$P(\xi_i^\mu)=\frac{e_\mu}{2N}(\delta_{\xi_i^\mu,1}+\delta_{\xi_i^\mu,-1})+\left(1-\frac{e_\mu}{N}\right)\delta_{\xi_i^\mu,0} \quad (13)$$

leading to $P(d)=\mathrm{Poisson}(\alpha\langle e\rangle)$ while $P(e)=P^{-1}\sum_\mu\delta_{e,e_\mu}$. If we keep the mean degree fixed at $\langle e\rangle=c$, the critical point for $\beta\to\infty$ is found at

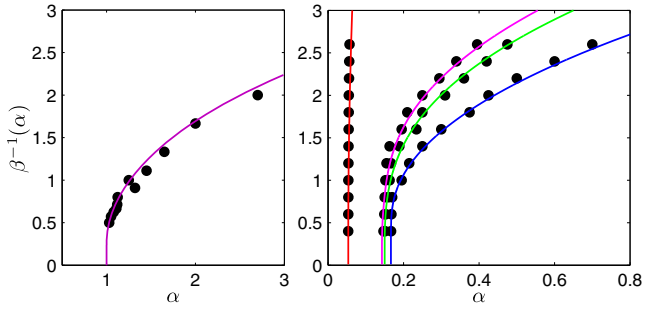

FIG. 4 (color online). Transition lines (theory, with symbols from PD numerics) for different pattern degree distributions. Left: $e \sim \text{Poisson}(c = 1)$. Right: Changing $P(e)$ at constant $\langle e \rangle = 3$: $P(e) = \delta_{e,3}$ (blue); $P(e) = (\delta_{e,2} + \delta_{e,3} + \delta_{e,4})/3$ (green); $P(e) = (\delta_{e,2} + \delta_{e,4})/2$ (pink); $P(e)$ power law as in preferential attachment graphs, with $\langle e^2 \rangle = 21.66$ (orange).

$$B_d B_e = \alpha c(\langle e^2 \rangle / c - 1) = \alpha[c(c-1) + \text{Var}(e)] = 1$$

while for large α one obtains for the critical line $\beta_c^{-1}(\alpha) \approx \sqrt{\alpha}\sqrt{c(c-1) + \text{Var}(e)}$. Similar results are obtained with soft constraints e_μ on the degrees, i.e., by dropping the delta function constraint in $P(\boldsymbol{\xi})$ before Eq. (13): one now finds $B_d B_e = \alpha[c^2 + \text{Var}(e)]$ and $\beta_c^{-1}(\alpha) \approx \sqrt{\alpha}\sqrt{c^2 + \text{Var}(e)}$. In both cases, the region where parallel retrieval is obtained is larger for degree distributions with smaller variance; the optimal situation occurs when all patterns have exactly the same number c of nonzero entries [Fig. 4 (right)]. Notably, then, scale-free networks, which perform best for information spreading [23], are not optimal for information processing [14]. For the special case of homogeneous dilution $P(\xi_i^\mu = \pm 1) = c/(2N)$ we easily recover previous results [8]: the distributions of pattern degrees e and neuron degrees d are Poisson(c) and Poisson(αc), respectively, so $B_d = \alpha c$, $B_e = c$ so that there is no bifurcation for $\alpha c^2 < 1$. The network acts as a parallel processor here for any β because the bipartite network consists of finite clusters of interacting spins in which there is no interference between different patterns [8]. At higher connectivity, the critical line defined above by $A = 1$ indicates the temperature above which this lack of interference persists even though the network now has a giant connected component. Figure 4 (left) compares theory to PD results, where we locate the transition as the onset of nonzero second moments of the field distributions. The impact of the transition on the overlap probability distribution of a pattern with fixed e can be seen from the PD results in Fig. 3 (middle and right panels). Crossing the transition line, parallel retrieval is accomplished at low temperatures, but it degrades when α is increased (see shrinking peaks in the middle panel), or c is increased, eventually fading away for sufficiently large α and c (right panel).

We can also analyze the case of asymmetric patterns, where we take for the nonzero pattern entries $P(\xi_i^\mu = \pm 1) = (1 \pm a)/2$ with a degree of asymmetry $a \in [-1, +1]$. One

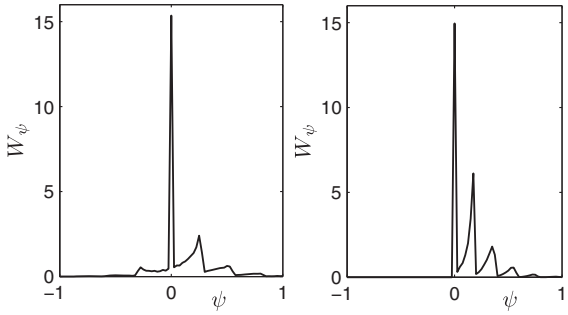

FIG. 5. Histogram of the fields ψ in the ferromagnetic region, for $c = 1$, $\beta = 1$, and different levels of bias: $a = 0.9$ with $\alpha = 9$ (left) and $a = 1$ with $\alpha = 8$ (right). Field distributions are obtained by PD starting from positive fields, to break the gauge symmetry. For $a = 1$ (right) there are only positive fields, as expected: when all patterns have positive entries there are no conflicting signals, even above the percolation threshold.

can show [14] that at zero temperature the bifurcation occurs when $B_d B_e = a^{-2}$; when a tends to zero the transition point goes to infinity and we retrieve the symmetric case. Beyond the bifurcation, noncentered field probability distributions (see Fig. 5) produce a nonzero global magnetization typical of ferromagnetic systems. However, a bifurcation towards growing field variances at zero means can also still occur. The physical bifurcation is the one taking place first on increasing β; Fig. 6 shows that at large α this is the one to growing means, at small α to growing variances.

To our knowledge, ours is the first study to quantify analytically the impact of heterogeneous degree distributions on the resilience of (parallel) processes on graphs. Degree heterogeneities in monopartite graphs [24] are well known to affect their resilience [25,26] and the dynamics of processes that they support (transport, epidemics, etc.) [27,28], due to the fact that hubs enhance the spread of information across the network. Our method paves the way for exploring similar qualitatively important phenomena in bipartite systems.

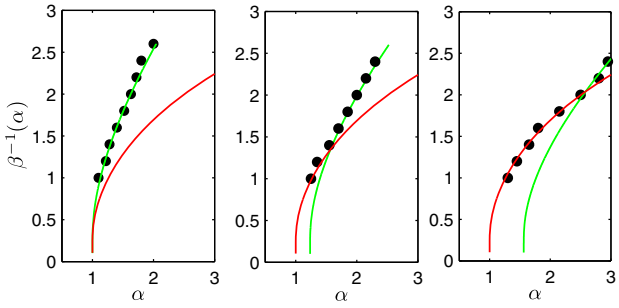

FIG. 6 (color online). Transition lines to growing field means (theory, green) and variances (theory, red), showing a good match to numerical PD data (dots); here $c = 1$ and pattern bias $a = 1, 0.95, 0.9$ from left to right. The first line to be crossed from high $T = \beta^{-1}$ gives the physical transition.

A. B. thanks GNFM-INdAM [Gruppo Nazionale per la Fisica Matematica (Istituto Nazionale d'Alta Matematica)] (Progetto Giovani Agliari) and D. T. acknowledges Sapienza University (Avvio alla Ricerca) for financial support. P. S. acknowledges funding from the EU under REA (Research Executive Agency) Grant No. 290038 (NETADIS). The authors also acknowledge insightful discussions with Elena Agliari and Yamir Moreno.

[1] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. Lett. **55**, 1530 (1985).
[2] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Ann. Phys. (N.Y.) **173**, 30 (1987).
[3] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[4] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).
[5] D. J. Amit, *Modeling Brain Function* (Cambridge University Press, Cambridge, England, 1992).
[6] E. Agliari, A. Barra, F. Guerra, and F. Moauro, J. Theor. Biol. **287**, 48 (2011).
[7] E. Agliari, A. Barra, A. Galluzzi, F. Guerra, and F. Moauro, Phys. Rev. Lett. **109**, 268101 (2012).
[8] E. Agliari, A. Annibale, A. Barra, A. C. C. Coolen, and D. Tantari, J. Phys. A **46**, 415003 (2013).
[9] E. Agliari, A. Annibale, A. Barra, A. C. C. Coolen, and D. Tantari, J. Phys. A **46**, 335101 (2013).
[10] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A. L. Barabasi, Science **297**, 1551 (2002).
[11] S. Boccaletti, V. Latora, Y. Moreno, and M. Chaves, Phys. Rep. **424**, 175 (2006).
[12] S. Maslov and K. Sneppen, Science **296**, 910 (2002).
[13] L. H. Hartwell, J. J. Hopfield, S. Leibler, and A. W. Murray, Nature (London) **402**, C47 (1999).
[14] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.113.238106, which includes Refs. [15–18].
[15] E. Agliari and A. Barra, Europhys. Lett. **94**, 10002 (2011).
[16] B. Bollobas, *Random Graphs* (Springer Press, New York, 1998).
[17] N. C. Wormald, in *Survey in Combinatorics* (Cambridge University Press, Cambridge, England, 1997).
[18] U. Ferrari, C. Lucibello, F. Morone, G. Parisi, F. Ricci-Tersenghi, and T. Rizzo, Phys. Rev. B **88**, 184201 (2013).
[19] Equation (1) corresponds to Gaussian τ_i in the bipartite SG; for discrete τ_i one would average over $z = \pm\beta$ with probability $1/2$ each.
[20] M. Mezard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).
[21] M. Mezard and A. Montanari, *Information, Physics and Computation* (Oxford University Press, New York, 2009).
[22] M. J. Urry and P. Sollich, J. Mach. Learn. Res. **14**, 1801 (2013), http://jmlr.csail.mit.edu/papers/volume14/urry13a/urry13a.pdf.
[23] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).
[24] J. S. Lee, K.-I. Goh, B. Kahng, and D. Kim, Eur. Phys. J. B **49**, 231 (2006).
[25] K.-I. Goh, E. Oh, H. Jeong, B. Kahng, and D. Kim, Proc. Natl. Acad. Sci. U.S.A. **99**, 12583 (2002).
[26] A. Annibale, A. C. C. Coolen, and G. Bianconi, J. Phys. A **43**, 395001 (2010).
[27] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001).
[28] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).