Replica cluster variational method: the replica symmetric solution for the 2D random bond Ising model

Alejandro Lage-Castellanos[1]**, Roberto Mulet**[1]**,
Federico Ricci-Tersenghi**[2] **and Tommaso Rizzo**[3]

[1] 'Henri-Poincaré-Group' of Complex Systems and Department of Theoretical Physics,
Physics Faculty, University of Havana, La Habana, CP 10400, Cuba
[2] Dipartimento di Fisica, INFN—Sezione di Roma 1 and CNR—IPCF, UOS di Roma,
Università La Sapienza, P.le A. Moro 5, I-00185 Roma, Italy
[3] CNR—IPCF, UOS di Roma, Università La Sapienza, P.le A. Moro 5, I-00185 Roma, Italy

E-mail: federico.ricci@roma1.infn.it

Abstract

We present and solve the replica symmetric equations in the context of
the replica cluster variational method for the 2D random bond Ising model
(including the 2D Edwards–Anderson spin-glass model). First, we solve a
linearized version of these equations to obtain the phase diagrams of the model
on the square and triangular lattices. In both cases, the spin-glass transition
temperatures and the multicritical point estimations improve largely over the
Bethe predictions. Moreover, we show that this phase diagram is consistent
with the behavior of inference algorithms on single instances of the problem.
Finally, we present a method to consistently find approximate solutions to the
equations in the glassy phase. The method is applied to the triangular lattice
down to $T = 0$, also in the presence of an external field.

PACS numbers: 05.10.−a, 75.50.Lk

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

1. Introduction

Since the celebrated work of Edwards and Anderson in 1975 [1], many efforts have been
devoted to the analytic description of spin glasses. Particularly remarkable is the solution
found by Parisi in 1979 [2, 3] to the Sherrington–Kirkpatrick mean-field model [4]. The
physical interpretation of the Parisi solution [5] gave a solid basis to concepts such as replica
symmetry (RS) and spontaneous replica symmetry breaking (RSB) that became of standard
use in the scientific community. The solutions of many models, not necessarily of mean-field

type, were interpreted along these lines (see e.g. the review in [6] about spin glasses on finite-dimensional lattices).

In this context, the last decade has been very exciting both from the conceptual and the practical point of view. First, Mézard and Parisi [7, 8] were able to solve analytically the spin-glass (SG) model on a Bethe lattice (usually called the Viana–Bray model [9]) with a RSB ansatz. Within this RSB ansatz, the solution is given in terms of populations of fields that contain all the necessary information to describe the low-temperature phase of the model. The extension to other models was immediate [10–12] and the approach was fundamental to the introduction of the survey propagation algorithm [11] that has been successfully applied in the solution of many *single instances* optimization problems [13, 14]. Moreover, it was soon recognized that the well-known belief propagation (BP) algorithm [15] corresponds to the Bethe approximation [16], that is, the replica symmetric solution on the Bethe lattice.

Unfortunately, all the above analytical results concern mean-field models. To go beyond the Bethe approximation, one should also consider loops in the interaction network and this turns out to be a highly non-trivial task (see for example [17–22]). Yedidia and co-workers [23] described how to generalize the cluster variational method (CVM) of Kikuchi [24] that allows us to derive a free energy that improves the Bethe one by considering exactly the contribution of small loops. The minimization of the CVM free energy can be achieved by the use of a generalized belief propagation (GBP) algorithm [23], but the solution found is always replica symmetric.

The idea of merging the CVM with the RSB ansatz was around for some years, but it remained elusive. Probably because the simplest comprehension of the RSB ansatz within the Bethe approximation is based on a probabilistic cavity construction [7], which is hard (or even impossible) to derive for a general CVM. In a recent paper [25], we proposed a formal solution to this problem. The idea was to apply the CVM to an already replicated free energy, and then within the RSB ansatz to send the number of replicas to zero. This formulation allowed us to derive a set of closed equations for some local fields that play the same role of the cavity fields in the Bethe approximation. Unfortunately, these fields enter into the equations in an implicit form and so standard population dynamic algorithms cannot be used for finding the solution. In previous works [25, 26], using linear stability analysis, we showed that these equations improve the Bethe approximation on the location of the phase boundaries. However, the solution of these equations in the low-temperature phase and the interpretation of this solution in terms of the performance of inference algorithms are still important open problems.

The main goal of this work is to extend our previous results in these two directions. On the one hand, using a stability analysis we study the phase diagram in the ρ (density of ferromagnetic couplings) versus T (temperature) plane for the Edwards–Anderson (EA) model on the square and triangular lattices. Moreover, we show that the GBP algorithm stops converging close to the SG temperature predicted by our approximation. On the other hand, we propose an approximated method to deal, at the RS level, with the complex equations that arise in the formalism in the low T phase.

The rest of the work is organized as follows. In the following section, we rederive the equations already obtained in [25] but now limiting its scope to the RS scenario in the average case. In section 3, we present the phase diagram obtained by a linearized version of these equations and in section 4 we study the consequences of this phase diagram for the performance of GBP. Section 5 shows the solution of a nonlinear approximation for the RS equations in the glassy phase. Finally, the conclusions and possible extensions of our approach are outlined in section 6.

2

2. The CVM replica symmetric solution

The EA model is defined by the Hamiltonian $H = -\sum_{(ij)} J_{ij} s_i s_j - h \sum_i s_i$, where the first sum is over neighboring spins on a finite-dimensional lattice, the couplings J_{ij} are quenched random variables and h is the external field. Although the equations we write are valid for generic couplings, our results will be obtained for couplings drawn from the distribution $P(J) = \rho \, \delta(J - 1) + (1 - \rho)\delta(J + 1)$.

In a model with quenched disorder, the free energy of typical samples can be obtained from the $n \to 0$ limit of the replicated free energy

$$
\Phi(n) = -\frac{1}{n\beta N} \ln \mathrm{Tr} \left\langle \exp\left(\sum_{(ij)} \beta J_{ij} \sum_{a=1}^n s_i^a s_j^a + \sum_i \beta h \sum_{a=1}^n s_i^a \right) \right\rangle_J
$$

$$
= -\frac{1}{n\beta N} \ln \mathrm{Tr} \exp\left(\sum_{(ij)} \ln \left\langle \exp \beta J \sum_a s_i^a s_j^a \right\rangle_J + \sum_i \beta h \sum_{a=1}^n s_i^a \right), \quad (1)
$$

where n copies of a system of N spins are considered at inverse temperature β, and the average over the quenched disorder is represented by the angular brackets.

The starting point of the Kikuchi CVM approximation is to choose a set of regions of the graph over which the model is defined. Restricting only to link and node regions, the CVM recovers the Bethe approximation. We will concentrate here on three kinds of regions: plaquettes (square or triangles, depending on the lattice), links and nodes. Using the definition $\psi_r(\sigma_r) \equiv \prod_{i,j \in r} \left\langle \exp \beta J \sum_a s_i^a s_j^a \right\rangle_J$, the energy of region r is

$$
E_r(\sigma_r) = -\ln \prod_{ij \in r} \psi_{ij}(\sigma_i, \sigma_j) - \ln \prod_{i \in r} \psi_i(\sigma_i), \quad (2)
$$

where the products run over all links and nodes (in the presence of a field) contained in region r. Let us also define the belief $b_r(\sigma_r)$ as an estimate of the marginal probability of the configuration σ_r according to the Gibbs measure. Then, within this approximation, the Kikuchi free energy takes the form

$$
F_K = \sum_{r \in R} c_r \left(\sum_{\sigma_r} b_r(\sigma_r) E_r(\sigma_r) + \sum_{\sigma_r} b_r(\sigma_r) \ln b_r(\sigma_r) \right), \quad (3)
$$

where the so-called Moebius coefficient c_r is the over-counting number of region r [23]. In the case of a model defined on the square lattice, the biggest regions are the square plaquettes, and by definition $c_P = 1$. Since each link region is contained in two plaquettes, $c_L = 1 - 2c_P = -1$. Moreover, the spins regions are contained in four plaquettes and four links, and $c_S = 1 - 4c_P - 4c_L = 1$. Similarly for the triangular lattices, $c_P = 1$, $c_L = 1 - 2c_P = -1$ and $c_S = 1 - 6c_P - 6c_L = 1$.

Now, the Kikuchi free energy has to be extremized with respect to the beliefs $b_r(\sigma_r)$, subject to the constraint that they are compatible upon marginalization. For example, $b_{(ij)}(\sigma_i, \sigma_j) = \sum_{\sigma_k, \sigma_l} b_{(ijkl)}(\sigma_i, \sigma_j, \sigma_k, \sigma_l)$ and $b_i(\sigma_i) = \sum_{\sigma_j} b_{(ij)}(\sigma_i, \sigma_j)$ for the square lattice. It can be shown [23, 27] that the free-energy extremization under these constraints can be better obtained by first introducing a set of 'messages' $m_{rs}(\sigma_s)$ from the region r to region s such that

$$
b_r(\sigma_r) \propto \psi_r(\sigma_r) \prod_{(r', s') \in M(r)} m_{r's'}(\sigma_r), \quad (4)
$$

where $M(r)$ is the set of connected pairs of regions (r', s') such that $r' \setminus s'$ is outside r, while s' coincides either with r or with one of its subsets (descendants). For example, if r is link (ij)

3

in a square lattice, the product in (4) contains the messages from the two squares adjacent to the link (ij), and the messages from the six other links connected to spins i and j.

The messages m_{rs} obey the following self-consistency equations:

$$m_{rs}(\sigma_s) \prod_{(r',s')\in M(r,s)} m_{r's'}(\sigma_s) \propto \sum_{\sigma_{r\backslash s}} \psi_{r\backslash s}(\sigma_r) \prod_{(r'',s'')\in M(r)\backslash M(s)} m_{r''s''}(\sigma_r), \qquad (5)$$

where $M(r, s)$ is the set of connected pairs of regions (r', s') such that r' is a descendant of r and s' is either region s or a descendant of s.

For the particular cases we are considering here (2D square and triangular lattices), the general expression (5) translates into the following two couple equations. The first equation is identical for both lattices and reads

$$m_{(ij)\to j}(\sigma_j) \propto \sum_{\sigma_i} \psi_{(ij)}(\sigma_i,\sigma_j) M_{\alpha\to(ij)}(\sigma_i,\sigma_j) M_{\beta\to(ij)}(\sigma_i,\sigma_j) \prod_{k\in\partial i\backslash j} m_{(ki)\to i}(\sigma_i), \qquad (6)$$

where α and β are the two plaquettes sharing the link (ij) and ∂i is the set of neighbors of site i. The notation used in this equation should make clear that messages are sent between a region and one of its descendant. The second equation takes slightly different forms for the square and triangular lattices, and we write it explicitly for the triangular lattice:

$$M_{(ijk)\to(ij)}(\sigma_i,\sigma_j) m_{(ik)\to i}(\sigma_i) m_{(jk)\to j}(\sigma_j) \propto \sum_{\sigma_k} \psi_{(ik)}(\sigma_i,\sigma_k) \psi_{(jk)}(\sigma_j,\sigma_k)$$

$$\prod_{\alpha\in\partial(ik)\backslash(ijk)} M_{\alpha\to(ik)}(\sigma_i,\sigma_k) \prod_{\beta\in\partial(jk)\backslash(ijk)} M_{\beta\to(jk)}(\sigma_j,\sigma_k) \prod_{l\in\partial k\backslash\{i,j\}} m_{l\to k}(\sigma_k), \quad (7)$$

where, in practice, the first two products only contain one message each. For the square lattice, the equation modifies slightly and contains some more products; disregarding all indices and arguments, its schematic form is $M\, m\, m \propto \sum \psi\, \psi\, \psi \prod M \prod M \prod M \prod m \prod m$.

Up to this point, the only difference with the standard CVM method is the introduction of replicated spins σ_i and the non-obvious connection with the average over the disorder, implicitly introduced in $\psi_r(\sigma_r)$. The main contribution of our previous work [25] was to introduce a consistent scheme to write these equations in the limit $n \to 0$ at any level of RSB.

Here, we reproduce the approach for the average case at the RS level. Following [28], we start by parametrizing the link to node messages in the following way:

$$m(\sigma_i) = \int \mathrm{d}u\, q(u) \exp\left[\beta u \sum_{a=1}^{n} \sigma_i^a\right] (2\cosh\beta u)^{-n}, \qquad (8)$$

and extend the same idea to the parametrization of the plaquette to link messages:

$$M(\sigma_i,\sigma_j) \propto \int \mathrm{d}U\, \mathrm{d}u_i\, \mathrm{d}u_j\, Q(U,u_i,u_j) \exp\left[\beta U \sum_{a=1}^{n} \sigma_i^a\sigma_j^a + \beta u_i \sum_{a=1}^{n} \sigma_i^a + \beta u_j \sum_{a=1}^{n} \sigma_j^a\right]. \quad (9)$$

The above parametrization allows us to rewrite the message passing equations (5) in terms of $q(u)$ and $Q(U, u_1, u_2)$. Substituting equations (8) and (9) into (6) and (7) and sending $n \to 0$, we obtain, after some standard algebra,

$$q(u) = \int \prod_{i}^{k} \mathrm{d}q_i \prod_{\alpha}^{p} \mathrm{d}Q_\alpha \langle\delta(u - \hat{u}(\#))\rangle_J,$$

$$R(U, u_a, u_b) \equiv \int \mathrm{d}u_i\, \mathrm{d}u_j\, Q(U, u_i, u_j) q(u_a - u_i) q(u_b - u_j)$$

$$= \int \prod_{i}^{K} \mathrm{d}q_i \prod_{\alpha}^{P} \mathrm{d}Q_\alpha \langle\delta(U - \hat{U}(\#))\delta(u_a - \hat{u}_a(\#))\delta(u_b - \hat{u}_b(\#))\rangle_J, \qquad (10)$$

4

where k (p) and K (P) correspond to the number of small m (large M) messages that enter into each equation. The specific expressions for $\hat{u}(\#)$, $\hat{U}(\#)$, $\hat{u}_a(\#)$, $\hat{u}_b(\#)$ depend on the lattice. The expressions for the triangular lattices are given in the following section and we refer the reader to [29] for similar formulas for the square lattice.

The next step is to solve the self-consistency equations in (10). Then, once q and Q are known, the thermodynamical observables are well defined in terms of these objects [25]. Unfortunately, since in (10) the functions Q and q are convoluted, this problem cannot be straightforwardly approached using a standard population dynamics algorithm. One possible approach is to deconvolve R using Fourier techniques to extract Q. Unfortunately, this approach suffers from strong instability problems. To use any numerical Fourier transform, one must have R and Q in the form of histograms. But since Q is not necessarily positive defined [25], the sampling of the messages becomes hard and the numerical errors due to the discretization of Q combined with the errors due to the Fourier inversion process make the convergence at low temperatures difficult. To bypass these numerical problems, we choose to solve these equations approximately. We perturb them in terms of small parameters around the paramagnetic solution and keep track of the information about the first few moments of the distributions.

3. Phase diagram from the linearized equations

Since the exact computation of $q(u)$ and $Q(U, u_1, u_2)$ is a daunting task, here we concentrate our attention on the calculation of their first two moments:

$$m = \int q(u)\, u\, \mathrm{d}u, \qquad a = \int q(u)\, u^2\, \mathrm{d}u, \qquad a_0(U) = \iint Q(U, u_1, u_2)\, \mathrm{d}u_1\, \mathrm{d}u_2,$$

$$M_i(U) = \iint Q(U, u_1, u_2)\, u_i\, \mathrm{d}u_1 \mathrm{d}u_2, \qquad a_{ij}(U) = \iint Q(U, u_1, u_2)\, u_i\, u_j\, \mathrm{d}u_1 \mathrm{d}u_2, \quad (11)$$

where $i, j \in \{1, 2\}$. With these definitions, the moments are determined by

$$m = \int \prod_i^k \mathrm{d}q_i \prod_\alpha^p \mathrm{d}Q_\alpha \langle \hat{u} \rangle_J$$

$$a = \int \prod_i^k \mathrm{d}q_i \prod_\alpha^p \mathrm{d}Q_\alpha \langle \hat{u}^2 \rangle_J$$

$$M_1(U) = \int \mathrm{d}u_1 \mathrm{d}u_2 R(U, u_1, u_2) \langle u_1 \rangle_J - m\, a_0(U) \qquad (12)$$

$$a_{11}(U) = \int \mathrm{d}u_1 \mathrm{d}u_2 R(U, u_1, u_2) \langle u_1^2 \rangle_J - 2m M_1(U) - a\, a_0(U)$$

$$a_{12}(U) = \int \mathrm{d}u_1 \mathrm{d}u_2 R(U, u_1, u_2) \langle u_1 u_2 \rangle_J - 2m M_1(U) - m^2 a_0(U).$$

However, keep in mind that $R(U, u_1, u_2)$ is still defined in terms of q and Q, see (10), and not directly in terms of the moments. Therefore, in order to compute the integrals in (12), one must introduce some ansatz over these distributions. It is then reasonable to start considering as correct the high temperature solution and to linearize the equations around this solution. At high temperatures and zero external field, one may assume that the system is paramagnetic:

$$q(u) = \delta(u), \qquad Q(U, u_1, u_2) = a_0(U)\delta(u_1)\delta(u_2). \qquad (13)$$

In what follows, we show, first, the linearization of $\hat{u}(\#)$ and $\hat{u}_i(\#)$ for the triangular lattice. Then, as an example, we present the derivation of the expressions for m and a entering (12) and leave for the appendix the expressions for the moments of Q. The algebra associated with
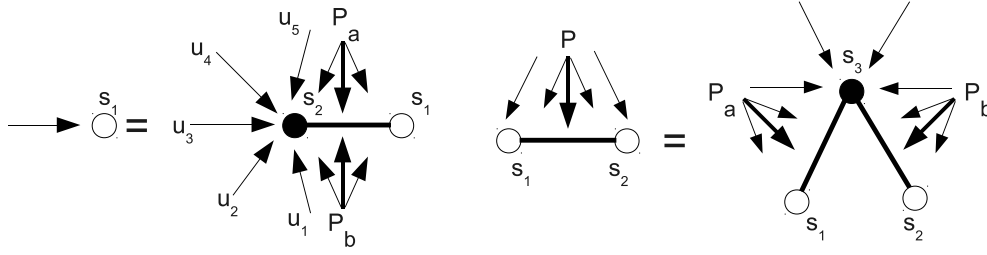
Figure 1. Schematic representation of the message passing equations (10) for the triangular lattice.

the equations for the square lattice is more cumbersome, but is technically equivalent. The interested reader may look for the case with $m = 0$ in [25] and [29].

To compute $\hat{u}(\#)$, it is enough to understand that the first equation in (10) may be interpreted as a standard equation for the Bethe approximation with a renormalized interaction between the spins $\hat{J} = J + U^a + U^b$, where a and b are indices referring to neighboring plaquettes (see the first panel in figure 1). Then, one can follow standard calculations [30] and the expression for the single message $\hat{u}$ reads

$$\hat{u}(\#) = u_1^a + u_1^b + \frac{1}{\beta}\text{arctanh}[\tanh(\beta\hat{J})\tanh(\beta h)], \tag{14}$$

where u_1^a and u_1^b are the small messages sent from the corresponding neighboring plaquettes to the site of interest and $h = u_2^a + u_2^b + \sum_i^5 u_i$, where u_2^a and u_2^b are the messages sent from the same plaquettes to the other end of the link. Considering that all the u and h fields are small, as must be the case close to the paramagnetic transition, the linearized version of the previous expression becomes

$$\hat{u}(\#) = u_1^a + u_1^b + \tanh(\beta\hat{J})h. \tag{15}$$

The messages in the second equation of (10) can be rewritten through the following identities:

$$\hat{U}(\#) = \frac{1}{4\beta}\ln\frac{K(1,1)K(-1,-1)}{K(1,-1)K(-1,1)}$$
$$\hat{u}_1(\#) = \frac{1}{4\beta}\ln\frac{K(1,1)K(1,-1)}{K(-1,1)K(-1,-1)} \tag{16}$$
$$\hat{u}_2(\#) = \frac{1}{4\beta}\ln\frac{K(1,1)K(-1,1)}{K(1,-1)K(-1,-1)},$$

where $K(S_1, S_2) = \sum_{S_3}\exp\left[\beta(\hat{J}_{13}S_1S_3 + \hat{J}_{23}S_2S_3 + u_1^aS_1 + u_2^bS_2 + h_3S_3)\right]$, with $\hat{J}_{13} = J_{13} + U^a$, $\hat{J}_{23} = J_{23} + U^b$ and $h_3 = u_3^a + u_3^b + \sum_i^4 u_i$ (see the second panel in figure 1). Then, after some algebra, it is easy to show that

$$\hat{u}_1(\#) = u_1^a + \frac{1}{4\beta}\ln\frac{\cosh(\beta(h_3 + J_+)))\cosh(\beta(h_3 + J_-)))}{\cosh(\beta(h_3 - J_-)))\cosh(\beta(h_3 - J_+)))} \tag{17}$$
$$\simeq u_1^a + \frac{1}{2}[\tanh(\beta J_+) + \tanh(\beta J_-)]h_3$$

and in a similar way

$$\hat{u}_2(\#) \simeq u_2^b + \frac{1}{2}[\tanh(\beta J_+) - \tanh(\beta J_-)]h_3, \tag{18}$$

6

where $J_{\pm} = \hat{J}_{13} \pm \hat{J}_{23}$. With these expressions, we have all the necessary ingredients to write the linearized form of (12). Next, we show how to derive the linear equations for m and a, and in the appendix we present the results for the other quantities.

The single site magnetization $m = \langle u \rangle$ satisfies

$$m = \langle u \rangle = \int \mathrm{d}u q(u) u = \left\langle \int \mathrm{d}Q_a \, \mathrm{d}Q_b \prod_{i=1}^5 \mathrm{d}q_i u \delta(u - \hat{u}(\#)) \right\rangle_J = \left\langle \int \mathrm{d}Q_a \, \mathrm{d}Q_b \prod_{i=1}^5 \mathrm{d}q_i \hat{u} \right\rangle_J$$

$$= \left\langle \int \mathrm{d}Q_a \, \mathrm{d}Q_b \prod_{i=1}^5 \mathrm{d}q_i (u_1^a + u_1^b + \tanh(\beta\hat{J})h) \right\rangle_J \tag{19}$$

and using the definitions in (11) the last integral can be easily expressed in linear terms of the moments of the distributions. The result is

$$m = 5m \left\langle \int \mathrm{d}U^a \, \mathrm{d}U^b \tanh(\beta\hat{J}) a_0(U^a) a_0(U^b) \right\rangle_J + \int \mathrm{d}U^a M_1(U^a) + \int \mathrm{d}U^b M_1(U^b)$$

$$+ \left\langle \int \mathrm{d}U^a \, \mathrm{d}U^b \tanh(\beta\hat{J}) [a_0(U^b) M_1(U^a) + a_0(U^a) M_1(U^b)] \right\rangle_J . \tag{20}$$

The derivation of a proceeds in a similar way

$$a = \langle u^2 \rangle = \int \mathrm{d}u q(u) u^2 = \left\langle \int \mathrm{d}Q_a \, \mathrm{d}Q_b \prod_{i=1}^5 \mathrm{d}q_i (u_1^a + u_1^b + \tanh(\beta\hat{J})h)^2 \right\rangle_J$$

$$= \left\langle \int \mathrm{d}Q_a \, \mathrm{d}Q_b \prod_{i=1}^5 \mathrm{d}q_i ((u_1^a)^2 + (u_1^b)^2 + 2u_1^a u_1^b \right.$$

$$\left. + 2(u_1^a + u_1^b) \tanh(\beta\hat{J}) + \tanh^2(\beta\hat{J})h^2 \right\rangle_J , \tag{21}$$

which may be rewritten in terms of the moments:

$$a = \int \mathrm{d}U^a a_{11}(U^a) + \int \mathrm{d}U^b a_{11}(U^b) + 2m \int \mathrm{d}U^a \, \mathrm{d}U^b M_1(U^a) M_1(U^b)$$

$$+ 10m \left\langle \int \mathrm{d}U^a \, \mathrm{d}U^b \tanh(\beta\hat{J}) [a_0(U^a) M_1(U^b) + a_0(U^a) M_1(U^a)] \right\rangle_J$$

$$+ 2 \left\langle \int \mathrm{d}U^a \, \mathrm{d}U^b \tanh(\beta\hat{J}) [a_{12}(U^a) + a_{12}(U^b) + M_2(U^b) M_1(U^a) \right.$$

$$\left. + M_1(U^b) M_2(U^a)] \right\rangle_J$$

$$+ (5a + 20m^2) \left\langle \int \mathrm{d}U^a \, \mathrm{d}U^b \tanh^2(\beta\hat{J}) a_0(U^a) a_0(U^b) \right\rangle_J$$

$$+ 10m \left\langle \int \mathrm{d}U^a \, \mathrm{d}U^b \tanh^2(\beta\hat{J}) [a_0(U^a) M_2(U^b) + a_0(U^a) M_2(U^a)] \right\rangle_J$$

$$+ \left\langle \int \mathrm{d}U^a \, \mathrm{d}U^b \tanh^2(\beta\hat{J}) [a_{22}(U^a) + a_{22}(U^b) + 2M_2(U^b) M_2(U^a)] \right\rangle_J . \tag{22}$$

Similar expressions may be derived for M_i and a_{ij}, see the appendix, but note that they are not closed analytical expressions. The form of $a_0(U)$ is unknown, and must be determined for any temperature using population dynamics. Once $a_0(U)$ has been computed, one can study the set of linear equations for the moments and check the local stability of the paramagnetic solution. In order to do this, we start from the paramagnetic solution, i.e. all the moments
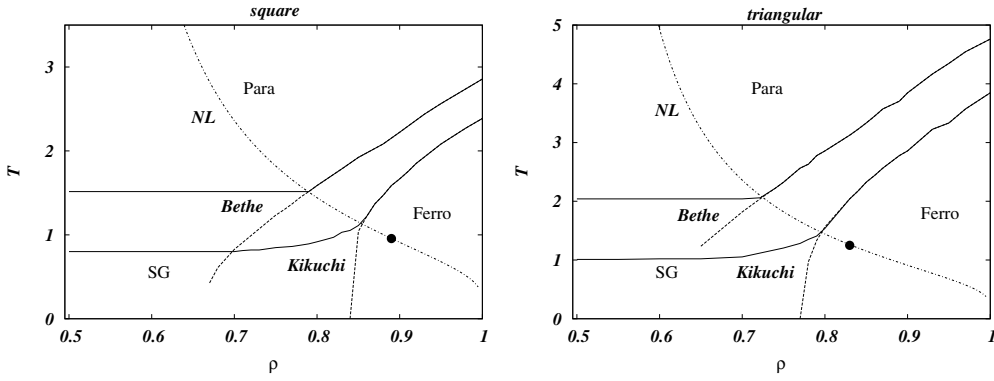
Figure 2. ρ versus T phase diagram for the square and the triangular lattice. We show results for the Bethe approximation (upper curves) and the Kikuchi approximation (lower curves). The Nishimori line (NL) is also shown. The bold circles on the NL are the best analytical predictions for the multicritical points. The dashed lines represent the boundaries for the existence of purely ferromagnetic solutions.

zero, but with $a_0(U)$ being non-trivial. Then, we slightly perturb a and m and check, solving iteratively equations (20), (22) and (A.1)–(A.3), whether these perturbations die out or diverge. Depending on ρ and T we find that under iteration, either both magnitudes diverge, or just a or none diverges. If a and m converge to zero, then the system is in the paramagnetic phase (P). If only a diverges it is in the SG phase, and if both a and m diverge we say that the system is in a ferromagnetic phase (F).

The results of this analysis are reported in figure 2. The phase diagrams must be read in the following way. Below the horizontal lines, we have the SG phase and above the paramagnetic phase. Critical lines meet at the multicritical point (ρ^{cr}, T^{cr}), located on the Nishimori line (NL). On the right of this multicritical point, i.e. if $\rho > \rho_{cr}$, the system is in the ferromagnetic phase at low temperatures and in the paramagnetic phase at high temperatures.

In both cases, the conclusions are similar: the P–SG critical temperature predicted by the Kikuchi approximation is lower than the one predicted by the Bethe approximation. This result was already shown for $\rho = 0.5$ in [25], but here we correct an error in that work where an incomplete range of β was considered during the study of the square lattice. In addition, these results are now extended to larger values of ρ. Moreover, we show that while both approximations correctly predict an SG to F transition at low temperatures and a multicritical point on the NL, the estimation of the latter is much better in the Kikuchi approximation (the big dots on the NL are the exact locations for the multicritical points predicted in [31] and [32]). The following table summarizes the locations of the multicritical points:

Lattice	ρ^{cr}_{Bethe}	$\rho^{cr}_{Kikuchi}$	ρ^{cr}_{exact}
Square	0.79	0.85	0.8894
Triangular	0.74	0.78	0.8358

Finally, we checked the existence of a ferromagnetic transition keeping a zero and perturbing m. Again, the Kikuchi approximation improves the Bethe one. Indeed, the latter predicts an SG–F critical line extending to very low ρ values (well below ρ_{cr}), while the
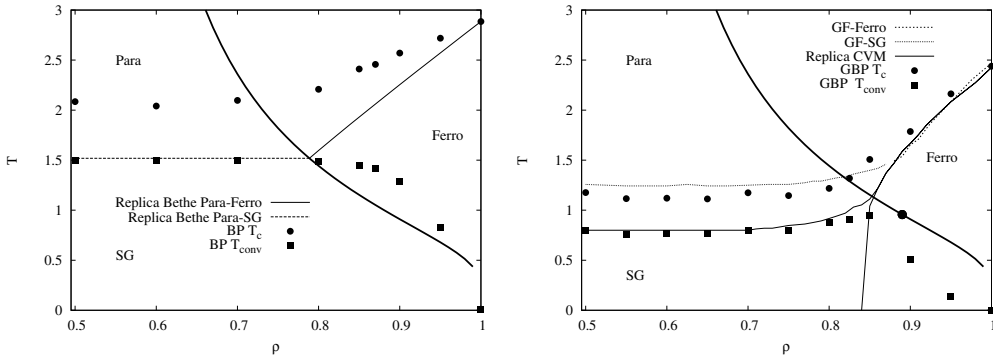
Figure 3. ρ versus T phase diagram for the square lattice in the Bethe (left panel) and plaquette-CVM (right panel) approximations. The circles indicate the temperature at which BP (GBP) finds a non-paramagnetic solution and the squares the temperature below which BP (GBP) does not converge. In the left panel 50 realizations of $N = 64 \times 64$ systems were used for the estimation of the critical temperatures, while in the right panel we have averaged over 10 samples of size $N = 256 \times 256$.

Kikuchi approximation has an SG–F critical line which is almost vertical in the ρ versus T phase diagram (and this behavior is consistent with the theoretical predictions [33]).

4. Connection to the behavior of inference algorithms

The results presented so far are obtained by taking the average over the ensemble and should then correspond to the properties of typical samples in the large N limit. It is known, however, that the predicted SG phase is not present in the 2D EA model at any finite temperature. This mistaken phase transition is a feature of any mean field like approximation (including Bethe and CVM), and therefore is not surprising. Nonetheless, the analytical method developed might keep its validity in relation to the behavior of message passing algorithms in single instances. In this section, we explore this connection for models on the square lattice.

When running BP and GBP for the Bethe and plaquette-CVM approximations on the square lattice we find a paramagnetic solution at high temperatures, characterized by zero local magnetizations $m_i = 0$. Below specific critical temperatures (that we call BP–T_c and GBP–T_c), both algorithms find non-paramagnetic solutions (i.e. with $m_i \neq 0$), as shown by the black circles in figure 3. These critical temperatures in single instances are far from the values predicted by the replica calculation for the para–SG and para–ferro transitions for any $\rho < 1$ value.

The transition that takes place at T_c in both BP and GBP algorithms is between a homogeneous phase, where all local magnetizations are null ($m_i = 0$), and a non-homogeneous phase, where some local magnetizations are different from zero ($m_i \neq 0$), while the global magnetization remains close to zero. This kind of symmetry breaking (space-homogeneous to space-heterogeneous) cannot be detected by the approximation underlying the replica CVM computation, because this method is intrinsically space-homogeneous: indeed, after averaging out the random couplings, one is left out with a 'one-body problem' as in any standard mean-field approximation.

The space heterogeneities arising at T_c are generated by the presence of large unfrustrated regions, where local magnetizations become non-zero [34]. The same unfrustrated regions are

9
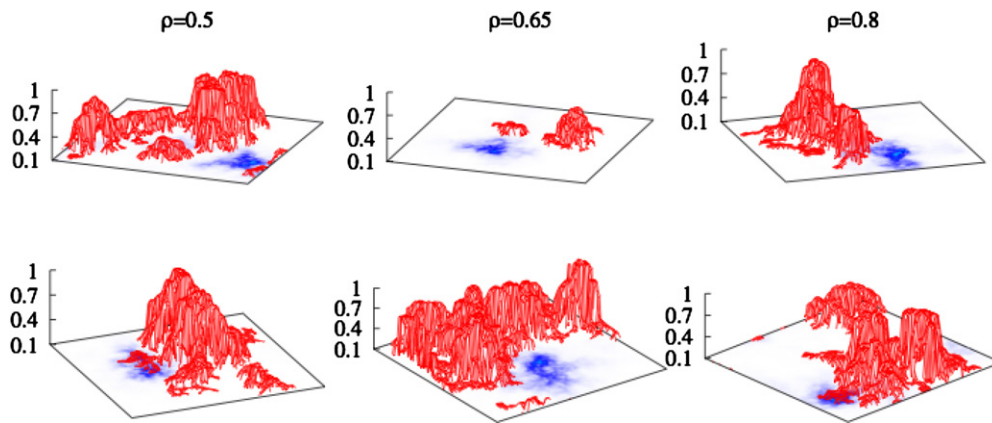
Figure 4. Red surfaces correspond to the absolute value of the magnetization of the spins in a 100×100 system, while the contours (shadowed blue areas) mark the regions where the non-convergence appears for the first time. For each of the three values of ρ, two different realizations of the disorder are shown. The *xy*-plane is set at magnetization $|m| = 0.1$. In most cases, the convergence problems appear in the low magnetization regions.

responsible for the Griffiths singularities [35, 36] that can be found in any finite-dimensional disordered model. It is worth noting that below T_c the solution found by BP has very small magnetizations (especially if compared with those found by GBP below GBP–T_c). This is the main reason why we missed BP–T_c in [26].

On the other hand, both BP and GBP stop converging at a temperature that is quite close to the one predicted by the replica calculations for the para–SG transition in the region $\rho < \rho^{cr}$ (see the black squares in figure 3). Connecting the lack of convergence of an iterative algorithm (as GBP) to the appearance of a flat direction in the CVM free energy is something very desirable: this is what one would call a 'static' explanation to a 'dynamical' behavior. However, here the situation is more subtle, because on any given large sample the message passing algorithm (either BP or GBP) ceases to converge to the paramagnetic fixed point at T_c: below T_c the fixed point reached by BP and GBP has many magnetized variables. So, how can the instability of the paramagnetic fixed point (where all local magnetizations are null) explain the lack of convergence of BP and GBP around the SG fixed point (with non-null magnetization)? We have studied in detail the behavior of GBP close to T_{conv} and we have discovered that in the regions with magnetized spins GBP messages are very stable and show no sign of instability; in contrast, in the regions where spin magnetizations are very close to zero, the GBP messages start showing strong fluctuations and finally produce an instability that leads to the lack of convergence of GBP (see figure 4).

Since in these regions of low local magnetizations the distribution of GBP messages is very similar to the one of the paramagnetic fixed point, then the average case computation for T_{CVM} shown in the previous section may perfectly explain the divergence of GBP messages in these regions. In other words, the space-homogeneous replica CVM calculation is able to predict the presence of a local instability in the message passing algorithms *assuming* that all local magnetizations are null. This assumption turns out to be correct only in those regions where the local symmetry is still unbroken and the local magnetizations are still zero. When this algorithmic local instability arises at a single point, then it propagates to the entire region of low magnetizations, leading to a lack of convergence for the message passing algorithm. Thus, we have a 'static' explanation for a 'dynamical' effect, and this is very desirable.

The above argument explains well the similarity between T_{CVM} and T_{conv} in the region $\rho < \rho_{\text{Kikuchi}}^{\text{cr}}$, where no ferromagnetic long range order is expected to take place. However, for $\rho > \rho_{\text{Kikuchi}}^{\text{cr}}$, the situation is more delicate: indeed there is ferromagnetic long range order below the critical line, and so the above argument cannot hold as it is (there are no large regions with null local magnetizations, where the instability can easily arise). Moreover, if we assume that a GBP instability can mainly grow in a region of low magnetizations, we would conclude that GBP must be much more stable for $\rho > \rho_{\text{Kikuchi}}^{\text{cr}}$. Indeed, what we see in figure 3 is that the behavior of the filled squares drastically changes around $\rho_{\text{Kikuchi}}^{\text{cr}}$, and T_{conv} becomes much smaller in the ferromagnetic phase. This observation supports the idea that an instability of GBP can mainly arise and grow in a region of low local magnetizations: in a ferromagnetic phase these regions are rare and small, and thus GBP is able to converge down to very low temperatures.

Average case with population dynamics

Recently, in [26] we have studied in detail the behavior of GBP on the 2D EA model (i.e. the present model with $\rho = 0.5$). For that case, we reported two important temperatures: a critical temperature T_c where the EA order parameter q_{EA} predicted by GBP becomes different from zero and a lower temperature T_{conv} where GBP stops converging to a fixed point. We noted that the critical temperature found by the replica CVM method was close to T_{conv}, while a critical temperature close to T_c could also be obtained from an average case calculation based on a population dynamics method, similar to the one used in [7] for the Bethe approximation.

The idea underlying this average case population dynamics is to assume that messages arriving on a given plaquette from different directions are independent. Differently from the replica CVM presented above, where the average over the disorder produces an effective one-body problem involving the two functions $q(u)$ and $Q(U, u_1, u_2)$, here we keep explicitly all the correlations between the messages coming from the same side of a given plaquette (actually from the same neighboring region) and disregard the remaining correlations. A population of these joint messages is then evolved by using the GBP equations to compute new messages which in turn replace elements of the population, until the population reaches stationarity. It is worth noting that, at the Bethe approximation level, the population dynamics method provides a result very similar to replica CVM, since the only function appearing in the latter, $q(u)$, is positively defined and can be interpreted as the probability distribution evolved in the population dynamics method.

Thanks to a local gauge symmetry, that is worth breaking in order to improve the convergence properties of the algorithm [26], in the GBP algorithm run on a regular lattice we can always set to zero one of the small-u messages in the triplet (hence the name 4-field for the joint messages). Figure 5 schematically represents a single step of this gauge fixed (GF) population dynamics, where the gray 4-field is computed from the four black 4-fields arriving on a plaquette with randomly chosen couplings. After many iterations of this fundamental step the population stabilizes.

In this average case computation, the correlation between the plaquette-to-link and the link-to-spin fields is accounted in the 4-fields structure, but different 4-fields are considered uncorrelated around the plaquette. One may wonder how strong is such an approximation. We show now that at the fixed point reached by GBP on a typical instance of the 2D EA model the correlation we are disregarding is rather weak, and so the approximation underlying the average case population dynamics is very reasonable. In figure 6, we show the correlation between GBP messages arriving on the same spin, comparing it to the mean absolute value of the messages (the topmost curve); clearly the correlations are all very small and could be
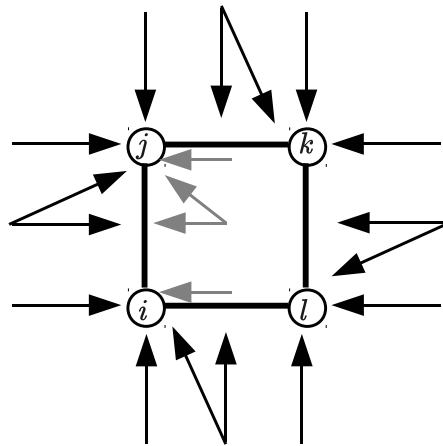
Figure 5. Population dynamics basic step. Four 4-fields $(u_{L\to i}, U_{P\to ij}, u_{P\to j}, u_{L\to j})$ are taken at random from the population, and new 4-fields are computed inside the square with random couplings. The new 4-fields (one of which is shown in gray) are added back to the population.

ignored on a first approximation. One more test that predictions obtained from the stationary population of 4-fields are very similar to those from the GBP algorithm run on single instances of the 2D EA model is provided in figure 7.

Once the population dynamics reaches a stationary state, we take measurements. The critical temperature is defined as the point where non-zero small u messages appear in the population of 4-fields and turns out to be very close to the value of T_c computed in single instances. In [26], these facts were reported as an interesting coincidence that we now extend to other values of ρ. In the right panel of figure 3, the upper dotted line marks the critical temperatures obtained by the GF population dynamics method. The key observation is that this temperature is quite close to the single instance critical temperature T_c of the GBP message passing algorithm. Moreover, in [26] we showed that the small discrepancy between the GF-SG critical temperature and the T_c measured on single samples decreases by increasing the sample size. The closeness of these two temperatures suggests that the messages (4-fields) arriving on a plaquette in a 2D lattice are almost uncorrelated and thus lead to results similar to those obtained by a population dynamics, where messages are uncorrelated by construction. So the critical temperature T_c for a given large sample can very well be estimated from the average case GF population dynamics. At the same time, it suggests that fixing the gauge and keeping the correlation among the four fields in a 4-field message is important to get the right critical temperature, but in the average case replica calculation we cannot fix the gauge and the correlation among the four fields is disregarded, since the distributions $Q(U, u_1, u_2)$ and $q(u)$ are in a product form. This is a weakness of the replica calculation in describing the actual behavior of the message passing algorithm on given samples.

In the Bethe approximation, a population dynamics of link-to-spin fields reproduces exactly the same critical temperature found by the replica method [7]. We tried to implement a new population dynamics, where all messages in a plaquette are updated at the same time, given the messages entering the plaquette, but the critical temperatures found do not compare well with BP–T_c. We also did not get better results by simulating in the Bethe approximation a population of the 2-fields (u_1, u_2) that enter the plaquette from one side.
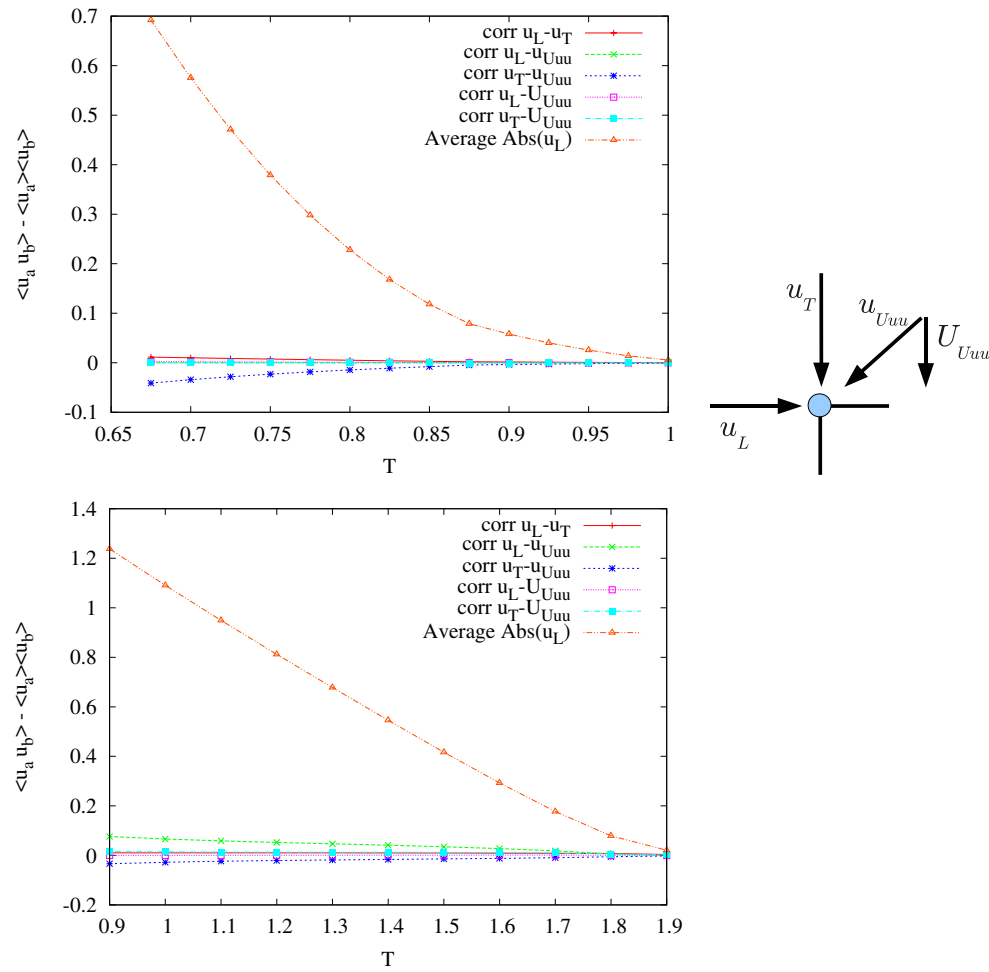
Figure 6. Correlations between u fields acting on a spin from different directions (see the right diagram) in two typical sample with $\rho = 0.5$ (top) and $\rho = 0.85$ (bottom).

These facts point in either of two directions. The first possibility is that the closeness of the GF population dynamics critical temperature (GF-SG in figure 5) to the critical temperature T_c in single instances is completely accidental. The second is that the population dynamics is actually related to the single instance behavior. In this latter case, the fact that in the Bethe approximation the population dynamics is useless in identifying T_c implies that not only the correlation kept in the 4-fields is crucial, but also the presence of the U-fields, somehow overruling the original couplings in the plaquettes, is very important.

5. Nonlinear regime

Supported by the positive results of the previous sections, we look for the solution of equations (10) in the nonlinear regime, below T_{CVM}. Still, the complete deconvolution of the second equation is beyond our technical capabilities and we reduce again the problem to
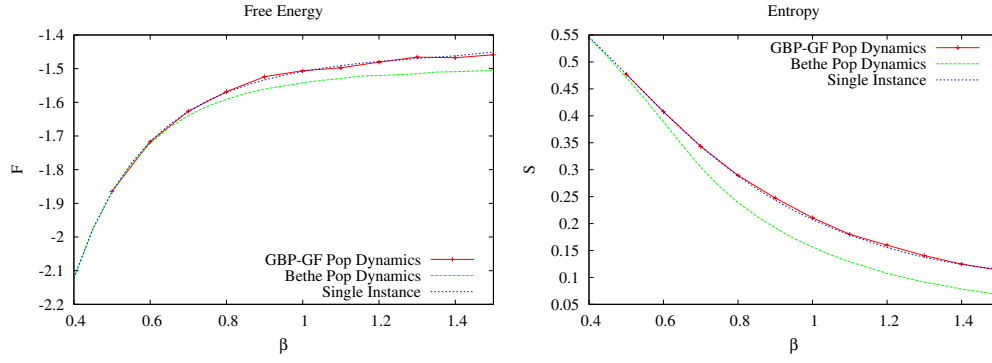
Figure 7. Free energy (left) and entropy (right) computed by GBP on 100 single instances of a 64×64 square 2D EA model ($\rho = 0.5$), by the GBP-GF population dynamics (with a population made of 5000 4-fields) and by the population dynamics at the Bethe level.
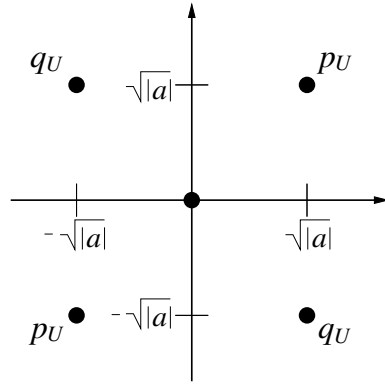
Figure 8. Schematic representation of the parametrization for $Q(U, u_1, u_2)$.

that of computing the different moments of the functions q and Q. However, now we keep the effect of the small messages beyond the linear regime. We show results for $\rho = 0.5$ such that m and $M_i(U)$ are zero. But the extension to more general cases is straightforward.

We start parametrizing $Q(U, u_1, u_2)$ in the following way:

$$Q(U, u_1, u_2) = a_0(U)\phi(u_1, u_2), \tag{23}$$

where

$$\begin{aligned}
\phi(u_1, u_2) = (1 - p_U - q_U)\delta(u_1)\delta(u_2) &+ p_U[\delta(u_1 - \sqrt{a})\delta(u_2 - \sqrt{a}) \\
&+ \delta(u_1 + \sqrt{a})\delta(u_2 + \sqrt{a})] + q_U[\delta(u_1 - \sqrt{a})\delta(u_2 + \sqrt{a}) \\
&+ \delta(u_1 + \sqrt{a})\delta(u_2 - \sqrt{a})].
\end{aligned} \tag{24}$$

This parametrization is sketched in figure 8. It is important to point out that the function ϕ is not necessarily positive, and that the parameters p_U and q_U depend on U, so are functions themselves. We proceed writing these parameters in terms of the moments of the distribution Q. This is easily done by substituting (23) in (11)

14
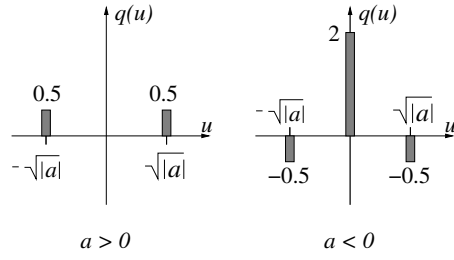
Figure 9. Schematic representation of the parametrization for $q(u)$.

$$a_{11}(U) = \int du_1\, du_2 Q(U, u_1, u_2) u_1^2 = a_0(U)[2a(p_U + q_U)], \tag{25}$$

$$a_{12}(U) = \int du_1\, du_2 Q(U, u_1, u_2) u_1 u_2 = a_0(U)[2a(p_U - q_U)], \tag{26}$$

such that

$$p_U = \frac{a_{11}(U) + a_{12}(U)}{4\, a\, a_0(U)}, \tag{27}$$

$$q_U = \frac{a_{11}(U) - a_{12}(U)}{4\, a\, a_0(U)}. \tag{28}$$

Within this parametrization, a fixes the deviation from the paramagnetic solution of the distribution Q. But a is defined by the distribution of the small messages $q(u)$. Therefore, we keep the consistency in the equations, without loosing physical insight, parametrizing also $q(u)$ in terms of a. The simplest parametrization is sketched in figure 9. It reads

$$q(u) = \tfrac{1}{2}[\delta(u - \sqrt{a}) + \delta(u - \sqrt{a})] \qquad\qquad \text{if } a > 0$$
$$q(u) = 2\delta(u) - \tfrac{1}{2}[\delta(u - \sqrt{|a|}) + \delta(u - \sqrt{|a|})] \quad \text{if } a < 0.$$

Note that the case $a < 0$ must be taken into consideration because, since Q is not necessarily positive defined [25], during the message passing procedure a may become negative. Now, specializing the computations to the case of the triangular lattice, the integrals over $R(U, u_1, u_2)$ in (12) take the form

$$a_{11}(U) = \int du_1\, du_2 R(U, u_1, u_2) u_1^2 - aa_0(U) = \int dU^a\, dU^b a_0(U^a) a_0(U^b)$$
$$\times \int d\bar{u}^a\, d\bar{u}^b \phi(u_1^a, u_2^a) \phi(u_1^b, u_2^b) \prod_{i=1}^4 q(u_i) \hat{u}_1^2 \delta(U - \hat{U}(\#)) - aa_0(U), \tag{29}$$

$$a_{12}(U) = \int du_1\, du_2 R(U, u_1, u_2) u_1 u_2 = \int dU^a\, dU^b a_0(U^a) a_0(U^b)$$
$$\times \int d\bar{u}^a\, d\bar{u}^b \phi(u_1^a, u_2^a) \phi(u_1^b, u_2^b) \prod_{i=1}^4 q(u_i) \hat{u}_1 \hat{u}_2 \delta(U - \hat{U}(\#)), \tag{30}$$

and a satisfies

$$a = \int dU^a\, dU^b a_0(U^a) a_0(U^b) \int d\bar{u}^a\, d\bar{u}^b \phi(u_1^a, u_2^a) \phi(u_1^b, u_2^b) \prod_{i=1}^5 q(u_i) \hat{u}^2(\#), \tag{31}$$
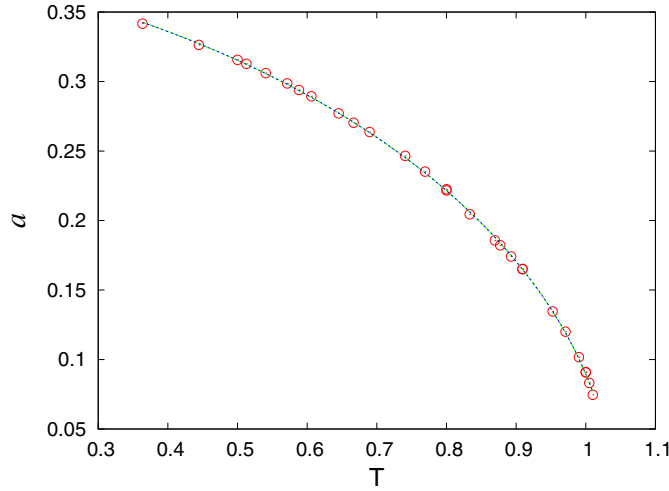
Figure 10. The *a* parameter for the triangular lattice in the low-temperature phase. The curve is a fit behaving as $(T_{\mathrm{CVM}} - T)^{1/2}$ close to the critical point.

where the integrals over U are done using a standard population dynamics and the integrals over $\vec{u}$ can be computed exactly thanks to the previous ansatz (keep in mind that $\phi(u_1^a, u_2^a)$ is given by (24)). The analysis for any other lattice is completely equivalent. Independently of the structure of the plaquettes or the lattice dimensions, the previous ansatz is always valid and the fixed point equations can always be reduced to expressions similar to (29)–(31). Only the computational effort may change. For example, while in equations (29) and (30) we integrate over two U messages, U^a and U^b, in the square lattice we will need a third message to integrate over. However, from the results obtained in the previous section we do not expect any gain in physical insight from studying the square lattice and we concentrate our efforts on the triangular lattice.

Our first result is presented in figure 10 where we present the dependence of a with T below T_{CVM}. Note that the data are compatible with a behavior of the form $a \propto (T_{\mathrm{CVM}} - T)^{1/2}$, although analytical arguments would suggest a linear behavior in $(T_{\mathrm{CVM}} - T)$, much as in the Bethe approximation case. It may be that the linear coefficient is actually very large, but we did not further investigate this point because it would require a consistent increase of numerical precision in the critical region.

In the presence of an external field, the symmetry which allows for the existence of a polynomial algorithm to solve the 2D EA model [33, 37] breaks down. On the other hand, the method based on the replica CVM equations can also be perfectly used in the presence of an external field: the equations remain practically the same, with the only difference that the external field must be added to the local field h in all the expressions above. We leave it for the interested reader to prove this.

We study the model in the presence of an external magnetic field, near, but below, the transition temperature and show (see figure 11) that both a and the energy E go as H^2 close to the transition. Our results, although approximate, can be considered as a good starting point to study the role of the external field in finite-dimensional lattices.

Finally, we prove that this technique can be extended to zero temperature and also provides non-trivial information in that limit. In figure 12, we show the structure of $a_0(U)$ considering a
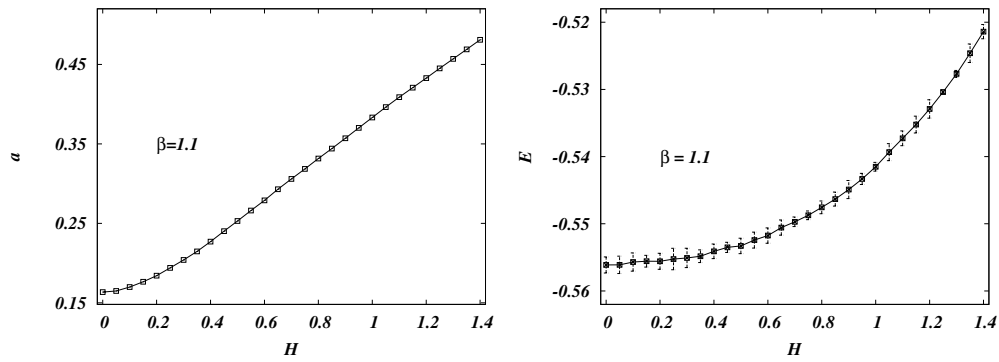
Figure 11. The *a* parameter (left) and the energy E (right) versus the external field H in the triangular lattice. The temperature is slightly below the paramagnetic–spin-glass phase transition ($\beta = 1.1$). Errors are due to the population finite size and to histograms discretization.
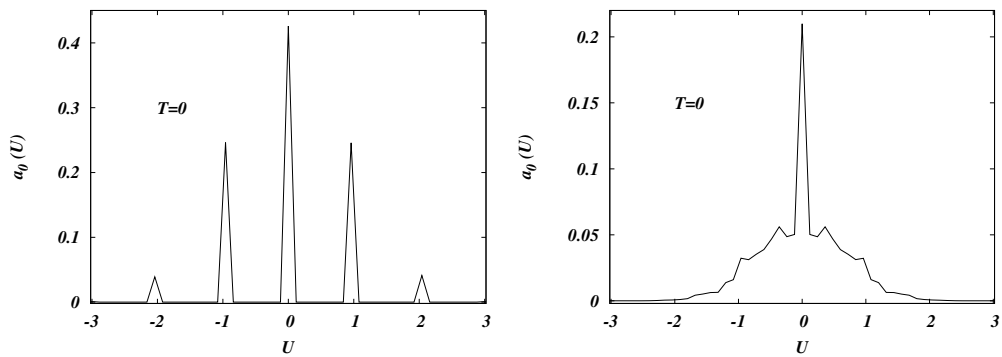
Figure 12. Distribution $a_0(U)$ of the messages U at zero temperature, using the paramagnetic ansatz (left) and the nonlinear ansatz of equation (23) (right) in the triangular lattice.

paramagnetic ansatz (left) where $\phi(u_1, u_2) = \delta(u_1)\delta(u_2)$ and $q(u) = \delta(u)$, and after reaching the fixed point of equations (29)–(31) (right). The paramagnetic solution has a structure very similar to the one found in the study of the EA model on a Bethe lattice [8]. This is not surprising since within the paramagnetic ansatz the problem is equivalent to a Bethe approximation on the dual lattice (see our previous work [29] for a larger discussion on this subject). On the other hand, the structure of $a_0(U)$ when nonlinear effects are considered is richer. While the $U = 0$ peak still dominates the distribution, and there is some reminiscence of other peaks, now the distribution spreads over non-integer values. This is probably one of the more remarkable mathematical consequences of the Kikuchi approximation. It is enough to consider the equation for $\hat{U}$ in the presence of small us, to understand that it is not possible to keep the self-consistency of the equations with distributions supported in the integers (even at $T = 0$). This unavoidable fact makes the computations at $T = 0$ as heavy as the computations at finite temperature, and further contributes to making the Kikuchi approximation harder to deal with than the Bethe approximation.

To further explore the role of $\phi(u_1, u_2)$, we plot $a_{11}(U)$ (left) and $a_{12}(U)$ (right) in figure 13. It is interesting to note that while $a_{11}(U)$ presents a structure with multiple peaks resembling the structure of $a_0(U)$, there is no clear evidence of such a structure in $a_{12}(U)$.
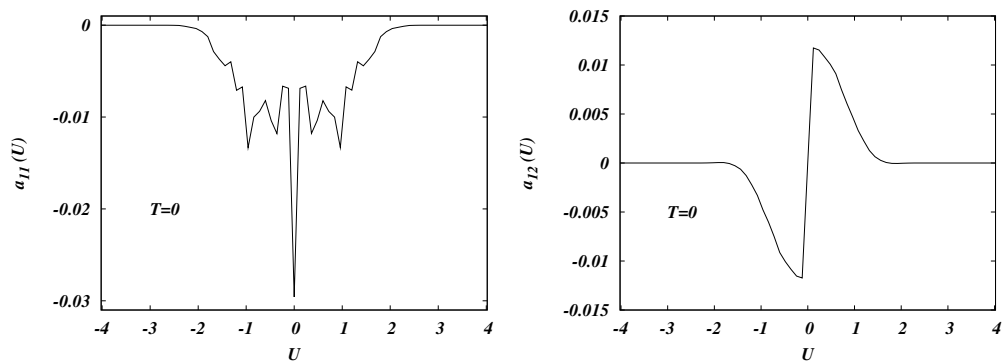
Figure 13. Self-correlations $a_{11}(U)$ (left) and cross-correlations $a_{12}(U)$ (right) of small u messages, using the nonlinear ansatz in equation (23) in the triangular lattice.

This suggests that self-correlations of the small u messages strongly depend on $a_0(U)$, much more than the cross-correlations do (these show a smooth curve at every U).

6. Conclusions

We study the typical properties of the 2D Edwards–Anderson model with the replica cluster variational method (CVM) at the RS level. Using a linearized version of the self-consistency equations, we have obtained the ρ versus T phase diagram on the square and triangular lattices. We show that this phase diagram resembles much better the theoretical predictions than the one obtained using the Bethe approximation: the SG critical temperature is lower, the multicritical point is closer to the exact value and the SG–ferro phase boundary looks similar to theoretical expectations. Moreover, we present numerical evidence supporting the idea that the temperature below which the average case computation predicts the existence of a SG phase (T_{CVM}) is also the temperature at which GBP algorithms stop converging. We apply to the triangular lattice a method to solve the RS equations in the nonlinear regime, i.e. at very low temperatures. The method does work and we show results at $T = 0$ and in the presence of an external magnetic field. All these results suggest that the replica CVM can be used to study finite-dimensional spin glasses, and hopefully in higher dimensions ($D > 2$) the approximation should provide an even better description of the low-temperature phase.

Acknowledgments

Appendix. Triangular lattice

We report here the expressions for the first and second moments of $Q(U, u_1, u_2)$ in the case of the triangular lattice:

$$M_1(U) = \left\langle \int \mathrm{d}U^a\, \mathrm{d}U^b\, \delta\big(U - \mathrm{arctanh}[\tanh(\beta(J^a + U^a))\tanh(\beta(J^b + U^b))]/\beta\big) \right.$$

$$\times \left[\frac{M_1(U^a)a_0(U^b) + a_0(U^a)M_1(U^b)}{2} + 2m\tanh(\beta J_+) \right.$$

$$\left.\left. + \tanh(\beta J_+)\frac{M_1(U^a) + M_1(U^b)}{2} \right] \right\rangle_J, \tag{A.1}$$

$$a_{11}(U) = \left\langle \int \mathrm{d}U^a \mathrm{d}U^b \delta\big(U - \mathrm{arctanh}[\tanh(\beta(J^a + U^a))\tanh(\beta(J^b + U^b))]/\beta\big) \right.$$

$$\times \left[\frac{a_{11}(U^a)a_0(U^b) + a_0(U^a)a_{11}(U^b)}{2} + 2m\tanh(\beta J_+)[M_1(U^a) + M_1(U^b)] \right.$$

$$+ 2m\tanh(\beta J_-)[M_1(U^a) - M_1(U^b)] + \tanh(\beta J_+)M_1(U^a)M_1(U^b)$$

$$+ \frac{1}{2}\tanh(\beta J_+)[a_{12}(U^a) + a_{12}(U^b)] + \frac{1}{2}\tanh(\beta J_-)[a_{12}(U^a) - a_{12}(U^b)]$$

$$+ (a + 3m^2)[\tanh^2(\beta J_+) + \tanh^2(\beta J_-)] + 2m[\tanh^2(\beta J_+)$$

$$+ \tanh^2(\beta J_-)][M_1(U^a) + M_1(U^b)] + \frac{1}{4}[\tanh^2(\beta J_+) + \tanh^2(\beta J_-)]$$

$$\left.\left. \times [a_{11}(U^a) + a_{11}(U^b) + 2M_1(U^a)M_1(U^b)] \right] \right\rangle_J, \tag{A.2}$$

$$a_{12}(U) = \left\langle \int \mathrm{d}U^a \mathrm{d}U^b \delta\big(U - \mathrm{arctanh}[\tanh(\beta(J^a + U^a))\tanh(\beta(J^b + U^b))]/\beta\big) \right.$$

$$\times \left[M_1(U^a)M_1(U^b) + 2m[\tanh(\beta J_+)(M_1(U^b) + M_1(U^a)) \right.$$

$$+ \tanh(\beta J_-)(M_1(U^b) - M_1(U^a))]$$

$$+ \frac{1}{2}\tanh(\beta J_+)[2M_1(U^a)M_1(U^b) + a_{12}(U^a) + a_{12}(U^b)]$$

$$+ \tanh(\beta J_-)\frac{a_{12}(U^b) - a_{12}(U^a)}{2}$$

$$+ (a + 3m^2)(\tanh^2(\beta J_+) - \tanh^2(\beta J_-))$$

$$+ 2m[\tanh^2(\beta J_+) - \tanh^2(\beta J_-)][M_1(U^a) + M_1(U^b)]$$

$$\left.\left. + \frac{1}{4}[\tanh^2(\beta J_+) - \tanh^2(\beta J_-)][a_{11}(U^a) + a_{11}(U^b) + 2M_1(U^a)M_1(U^b)] \right] \right\rangle_J, \tag{A.3}$$

where $J_+ = (J^a + U^a) + (J^b + U^b)$, $J_- = (J^a + U^a) - (J^b + U^b)$.

References

[1] Edwards S F and Anderson P W 1975 *J. Phys. F: Met. Phys.* **5** 965
[2] Parisi G 1979 *Phys. Lett*. A **73** 203
[3] Parisi G 1980 *J. Phys. A: Math. Gen.* **13** L115
[4] Kirkpatrick S and Sherrington D 1978 *Phys. Rev.* B **17** 4384
[5] Parisi G, Mézard M and Virasoro M 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[6] Marinari E, Parisi G, Ricci-Tersenghi F, Ruiz-Lorenzo J J and Zuliani F 2000 *J. Stat. Phys.* **98** 973
[7] Mézard M and Parisi G 2001 *Eur. Phys. J.* B **20** 217

 [8] Mézard M and Parisi G 2003 *J. Stat. Phys.* **111** 1
 [9] Viana L and Bray A J 1985 *J. Phys. C: Solid State Phys.* **18** 3037
[10] Mezard M, Parisi G and Zecchina R 2002 *Science* **297** 812
[11] Mézard M and Zecchina R 2002 *Phys. Rev.* E **66** 056126
[12] Mulet R, Pagnani A, Weigt M and Zecchina R 2002 *Phys. Rev. Lett.* **89** 268701
[13] Braunstein A, Mulet R, Pagnani A, Weigt M and Zecchina R 2003 *Phys. Rev.* E **68** 036702
[14] Zhou H 2003 *Eur. Phys. J.* B **32** 265
[15] Pearl J 1988 *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (San Francisco, CA: Morgan Kaufmann Publishers)
[16] Kabashima Y and Saad D 1998 *Europhys. Lett.* **44** 668
[17] Montanari A and Rizzo T 2005 *J. Stat. Mech.* P10011
[18] Chertkov M and Chernyak V Y 2006a *Phys. Rev.* E **73** 065102
[19] Chertkov M and Chernyak V Y 2006b *J. Stat. Mech.* P06009
[20] Mooij J M and Kappen H J 2007 *J. Mach. Learn. Res.* **8** 1113
[21] Gomez V, Kappen H J and Chertkov M 2010 *J. Mach. Learn. Res.* **11** 1273
[22] Xiao J-Q and Zhou H 2011 *J. Phys. A: Math. Theor.* **44** 425001
[23] Yedidia J, Freeman W T and Weiss Y 2005 *IEEE Trans. Inform. Theory* **51** 2282
[24] Kikuchi R 1951 *Phys. Rev.* **81** 988
[25] Rizzo T, Lage-Castellanos A, Mulet R and Ricci-Tersenghi F 2010 *J. Stat. Phys.* **139** 375
[26] Domínguez E, Lage A, Mulet R, Ricci-Tersenghi F and Rizzo T 2011 *J. Stat. Mech.* P12007
[27] Pelizzola A 2005 *J. Phys. A: Math. Gen.* **38** R309
[28] Monasson R 1998 *J. Phys. A: Math. Gen.* **31** 513
[29] Lage-Castellanos A, Mulet R, Ricci-Tersenghi F and Rizzo T 2011 *Phys. Rev.* E **84** 046706
[30] Mézard M and Montanari A 2009 *Information, Physics and Computation* (Cambridge: Oxford University Press)
[31] Takeda K, Sasamoto T and Nishimori H 2005 *J. Phys. A: Math. Gen.* **38** 3751
[32] Nishimori H and Ohzeki M 2006 *J. Phys. Soc. Japan* **75** 034004
[33] Poulter J and Blackman J 2001 *J. Phys. A: Math. Gen.* **34** 7527
[34] Zhou H, Wang C, Xiao J-Q and Bi Z 2011 *J. Stat. Mech.* L12001
[35] Griffiths R B 1969 *Phys. Rev. Lett.* **23** 17
[36] Vojta T 2006 *J. Phys. A: Math. Gen.* **39** R143
[37] Thomas C K and Middleton A A 2009 *Phys. Rev.* E **80** 046708