Loop corrections in spin models through density consistency
Supplementary Information

Alfredo Braunstein
Politecnico di Torino, Corso Duca Degli Abruzzi 24, Torino, Italy
Italian Institute for Genomic Medicine, Via Nizza 52, Torino, Italy and
Collegio Carlo Alberto, Via Real Collegio 1, Moncalieri,
Italy & INFN Sezione di Torino, Via P. Giuria 1, I-10125 Torino, Italy

Giovanni Catania
Politecnico di Torino, Corso Duca Degli Abruzzi 24, Torino, Italy

Luca Dall'Asta
Politecnico di Torino, Corso Duca Degli Abruzzi 24, Torino, Italy and
Collegio Carlo Alberto, Via Real Collegio 1, Moncalieri,
Italy & INFN Sezione di Torino, Via P. Giuria 1, I-10125 Torino, Italy

I. GENERAL PROPERTIES OF DC SCHEME

For completeness, the Density Consistency (DC) relations are

$$\frac{\mu_i}{\Sigma_{ii}} = \tanh^{-1} \langle x_i \rangle_{g^{(a)}} \tag{1}$$

$$\mu_i = \langle x_i \rangle_{g^{(a)}} \tag{2}$$

$$\frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}} = \rho \mathrm{corr}_{g^{(a)}} (x_i, x_j) \tag{3}$$

where

$$g(\boldsymbol{x}) = \frac{1}{z_g} \prod_a \phi_a (\boldsymbol{x}_a) = \frac{1}{z'} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}.$$

$$g^{(a)}(\boldsymbol{x}) = \frac{1}{z_a} g(\boldsymbol{x}) \sum_{\hat{\boldsymbol{x}}_a \in X_a} \delta (\boldsymbol{x}_a - \hat{\boldsymbol{x}}_a) \frac{\psi_a (\hat{\boldsymbol{x}}_a)}{\phi_a (\hat{\boldsymbol{x}}_a)}$$

are the full gaussian and the tilted distribution, respectively.

A. Relation with the Bethe Approximation (BP)

On acyclic graphs, both the DC scheme and BP are exact and thus they must coincide on their computation of marginals. However, a deeper connection can be pointed out. BP fixed point equations are

$$m_{ai}(x_i) \propto \sum_{\boldsymbol{x}_a} \psi_a (x_a) \prod_{j \in a \setminus i} m_{ja} (x_j) \tag{4}$$

$$m_{ia}(x_i) \propto \prod_{b \in i \setminus a} m_{bi} (x_i) \tag{5}$$

$$m_i(x_i) \propto \prod_{b \in i} m_{bi} (x_i) \tag{6}$$

Theorem 1. *If (H1) the DC scheme applies zero covariances or (H2) the factor graph is acyclic, $m_{ia}(x_i) \propto g^{-a}(x_i)$ satisfies (4)-(5). Moreover, the updates follow dynamically BP updates. In particular, if equations converge, approximate marginals $g(x_i)$ are proportional to belief magnetizations (6).*

Proof. In the either hypothesis (H1 or H2) , $g^{-a}(x_a) \propto \prod_{j \in \partial a} m_{ja}(x_j)$. Define $m_{ai}(x_i) \propto \frac{g(x_i)}{m_{ia}(x_i)}$. We obtain

$$\begin{aligned} m_{ai}(x_i) &\propto \frac{1}{m_{ia}(x_i)} \int dx_{a \setminus i} g(x_a) \\ &\propto \int dx_{a \setminus i} \prod_{j \in \partial a \setminus i} m_{ja}(x_j) \phi_a (x_a) \end{aligned} \tag{7}$$

Thanks to (1), $g(x_i) \propto g^{(a)}(x_i)$ when $x_i \in \{-1, 1\}$ (and this is precisely the purpose of (1)). In particular, for $x_i \in \{-1, 1\}$ we get also

$$\begin{aligned} m_{ai}(x_i) &\propto \frac{1}{m_{ia}(x_i)} g^{(a)}(x_i) \\ &= \frac{1}{m_{ia}(x_i)} \sum_{x_{a \setminus i}} g^{-a}(x_a) \psi_a (x_a) \\ &= \sum_{x_{a \setminus i}} \prod_{j \in \partial a \setminus i} m_{ja}(x_j) \psi_a (x_a) \end{aligned} \tag{8}$$

which is Eq. (4). Eq. (5) is also verified in either hypothesis:

1. Factorized case: if $\phi_a(x_a) = \prod_{i\in\partial a}\phi_a(x_i)$, then clearly $m_{ai}(x_i) \propto \phi_a(x_i)$ and $m_{ia}(x_i) \propto \prod_{b\in\partial i\setminus a} m_{bi}(x_i)$

2. Acyclic case: if T_b denotes the set of factors in the connected component of b once i is removed, we get

$$\begin{aligned} m_{ia}(x_i) &\propto g^{-a}(x_i) \\ &\propto \int dx_{-i} \prod_{b\in\partial i\setminus a}\phi_b(x_b) \prod_{c\in T_b\setminus b}\phi_c(x_c) \\ &\propto \prod_{b\in\partial i\setminus a}\int dx_{b\setminus i}\phi_b(x_b) \prod_{j\in\partial b\setminus i} g^{-b}(x_j) \\ &\propto \prod_{b\in\partial i\setminus a}\int dx_{b\setminus i}\phi_b(x_b) \prod_{j\in\partial b\setminus i} m_{bj}(x_j) \\ &\propto \prod_{b\in\partial i\setminus a} m_{bi}(x_i) \end{aligned} \tag{9}$$

where the last line follows from (7).

□

B. Relation with EP

The DC scheme can be thought of a modified Gaussian EP scheme for factors [1]

$$\hat{\psi}_{ij}(x_i,x_j) = \psi_{ij}(x_i,x_j)\left(\delta(x_i+1)+\delta(x_i-1)\right)\left(\delta(x_i+1)+\delta(x_i-1)\right)$$

Classic EP equations in this context can be obtained by replacing $\operatorname{atanh}\langle x_i\rangle_{g^{(a)}}$ in the RHS of (1) by the qualitatively similar function $\frac{\langle x_i\rangle_{g^{(a)}}}{1-\langle x_i\rangle^2_{g^{(a)}}}$, but this of course invalidates Theorems 1-2 and turns out to give a much worse approximation in general.

C. Weight gauge

One interesting property common to both DC and EP scheme concerns the possibility to move freely gaussian densities in and out the exact factors $\psi_a(x_a)$. Let $\rho_a(x_a)$ be Gaussian densities;

$$\begin{aligned} p(x) &\propto g(x)\prod_a \psi_a(x_a) \\ q(x) &\propto g(x)\prod_a \phi_a(x_a) \end{aligned}$$

and q a Gaussian EP or DC approximation. We have

$$p^{(a)}(x_a) \propto \psi_a(x_a)\int dx_{-a}\frac{g(x)\prod_b\phi_b(x_b)}{\phi_a(x_a)} \tag{10}$$

$$\propto \frac{\psi_a(x_a)}{\rho_a(x_a)}\int dx_{-a}\frac{\left[g(x)\prod_b\rho_b(x_b)\right]\prod_b\phi_b(x_b)/\rho_b(x_b)}{\phi_a(x_a)/\rho_a(x_a)} \tag{11}$$

$$q(x_a) = \int dx_{-a} g(x)\prod_b\phi_b(x_b) \tag{12}$$

$$= \int dx_{-a}\left[g(x)\prod_b\rho_b(x_b)\right]\prod_b\phi_b(x_b)/\rho_b(x_b) \tag{13}$$

As DC and EP algorithms impose constraints between $p^{(a)}(x_a)$ and $q(x_a)$, any approximating family $\{\phi_a\}$ for $(g,\{\psi_a\})$ leads to an equivalent family $\{\phi_a/\rho_a\}$ for $(g'=g\prod_b\rho_b, \{\psi'_a=\psi_a/\rho_a\})$ for arbitrary factors ρ_a.

D. Other closure equations

Eq. 1 is the only condition needed to make the approximation scheme exact on tree-graphs. In principle one could complement it with any other condition in order to obtain a well-determined system of equations and unknowns in the factor parameters. In this work we tried other complementary closure equations (including matching of covariance matrix, constrained Kullback-Leiber Divergence minimization, matching of off-diagonal covariances, in addition to 1). However, we found out that 2-3 were experimentally performing uniformly better on all the cases we analyzed.

II. HOMOGENEOUS ISING MODEL

Consider a homogeneous ferromagnetic Ising Model with hamiltonian $\mathcal{H} = -J\sum_{\langle i,j\rangle} x_i x_j - h^{ext}\sum_i x_i$ defined on a d-dimensional hypercubic lattice with periodic (toroidal) boundary condition: because of the translational invariance, all Gaussian factors ϕ_a are identical and the covariance matrix admits an analytic diagonalization.

At a given inverse temperature β, DC equations (1-3) are identical for all factors:

$$\begin{aligned} \sigma_0 &= \frac{m}{\text{atanh}m} \\ \sigma_1 &= \rho\frac{c-m^2}{1-m^2}\sigma_0 \\ y &= m\left(\gamma_0+\gamma_1\right) \end{aligned} \tag{14}$$

The DC solution is found by solving the above system of 3 fixed-point equations in the Gaussian parameters y, γ_0, γ_1 where $\sigma_0, \sigma_1, \gamma_0, \gamma_1$ equal respectively Σ_{ii}, Σ_{ij}, $(2d)^{-1}\left(\Sigma^{-1}\right)_{ii}$, $\left(\Sigma^{-1}\right)_{ij}$ for i,j two first lattice neighbors. Here $m = \langle x_i\rangle_{g^{(a)}}$ and $c = \langle x_i x_j\rangle_{g^{(a)}}$ are the moments computed under the distribution $g^{(a)}$:

$$\begin{aligned} m &= \tanh\left[z + \text{atanh}\left(\tanh\Gamma\tanh z\right)\right] \\ c &= \tanh\left[\Gamma + \text{atanh}\left(\tanh^2 z\right)\right] \end{aligned}$$

where

$$\begin{aligned} z &= \frac{\beta h^{ext}}{2d} + \left(\frac{1}{\sigma_0+\sigma_1}\frac{1}{\gamma_0+\gamma_1} - 1\right)y \\ \Gamma &= \beta J + \frac{\sigma_1}{\sigma_0^2-\sigma_1^2} + \gamma_1 \end{aligned}$$

The matrix Σ is the gaussian covariance matrix whose inverse is parametrized as follows:

$$\Sigma^{-1} = \mathcal{S}^{(d)} = 2d\gamma_0\mathbb{I}_{L^d} + \gamma_1\mathcal{A}^{(d)}$$

where $\mathcal{A}^{(d)}$ is the lattice adjacency matrix in dimension d, whose diagonalization is discussed in the next section.

A. Diagonalization of $A^{(d)}$

The hypercubic lattice in d dimensions can be regarded as the cartesian product of linear-chain graphs, one for each dimension. The adiacency matrix of the whole lattice can be thus expressed as function of the adiacency matrices of the single linear chains, by means of the Kronecker product (indicated by $\otimes$):

$$\mathcal{A}^{(d)} = \mathcal{A}^{(1)}\otimes\mathbb{I}_L\otimes ...\otimes\mathbb{I}_L + \mathbb{I}_L\otimes\mathcal{A}^{(1)}\otimes\mathbb{I}_L\otimes ...\mathbb{I}_L + ... + \mathbb{I}_L\otimes ...\otimes\mathbb{I}_L\otimes\mathcal{A}^{(1)}$$

where $\mathcal{A}^{(1)}$ is the adiacency matrix of a (closed) linear chain of size L:

$$\mathcal{A}^{(1)} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

The above expression allows to compute the spectral decomposition of $\mathcal{A}^{(d)}$ just by knowing the spectrum of the adiacency matrix of the linear chain. The matrix $\mathcal{A}^{(1)}$ is a special kind of circulant matrix and therefore it can be diagonalized exactly [2]. Its eigenvalues and eigenvectors are shown below:

$$\lambda_x^{(1)} = 2\cos\left(\frac{2\pi}{L}x\right) \qquad \nu_x^{(1)} = \frac{1}{\sqrt{L}}\left(1, w_x, w_x^2, \ldots, w_x^{L-1}\right)$$

where $x \in \{0, ..., L-1\}$ and $w_x = e^{i\frac{2\pi}{L}x}$.
The spectral decomposition of $\mathcal{A}^{(d)}$ reads:

$$\lambda^{(d)}_{(x_1,\ldots,x_d)} = \sum_{j=1}^{d} \lambda^{(1)}_{x_j} = 2\sum_{j=1}^{d} \cos\left(\frac{2\pi}{L}x_j\right) \tag{15}$$

$$\nu_{(x_1,\ldots x_d)} = \otimes_{j=1}^{d} \nu^{(1)}_{x_j} \tag{16}$$

We recall now the expression of the eigenvalues of $S^{(d)}$:

$$\lambda_{(x_1,\ldots,x_d)} = 2d\gamma_0 + 2\gamma_1 \sum_{j=1}^{d} \cos\left(\frac{2\pi}{L}x_j\right)$$

The inverse matrix elements Σ_{ii}, Σ_{ij} can be computed in a straightfoward way. In particular, in the thermodynamic limit ($L \to \infty$) their expressions read:

$$\sigma_0 = \frac{1}{\gamma_0} R\left(r\right) \tag{17}$$

$$\sigma_1 = \frac{1}{\gamma_0 r}\left[\frac{1}{2d} - R\left(r\right)\right] \tag{18}$$

where $r = \frac{\gamma_1}{\gamma_0}$ and $R_d\left(r\right) = \frac{1}{2}\int_0^\infty dt e^{-dt}\left[\mathcal{I}_0\left(rt\right)\right]^d$, where $\mathcal{I}_0$ is the modified Bessel function of the first kind of order 0.

B. Simplified DC equations

It is possible to simplify the original system (14) in order to get a fixed point equation for the magnetization m. By eliminating the variable y and setting $J = 1$ we get

$$\begin{aligned} z =& m\left(\frac{1}{\sigma_0 + \sigma_1} - (\gamma_0 + \gamma_1)\right) + \frac{\beta h^{ext}}{2d} \\ =& m\gamma_0\left(\frac{1}{R_d\left(r\right) + \frac{1}{r}\left[\frac{1}{2d} - R_d\left(r\right)\right]} - r - 1\right) + \frac{\beta h^{ext}}{2d} \end{aligned} \tag{19}$$

$$\begin{aligned} \Gamma =& \beta + \frac{\sigma_1}{\sigma_0^2 - \sigma_1^2} + r\gamma_0 \\ =& \beta + \gamma_0\left(\frac{\frac{1}{r}\left(\frac{1}{2d} - R_d\left(r\right)\right)}{R_d^2\left(r\right) - \frac{1}{r^2}\left[\frac{1}{2d} - R_d\left(r\right)\right]^2} + r\right) \end{aligned} \tag{20}$$

Now, putting together Eq.(14) with Eq.(19)-(20) and the definitions (17)-(18) we get the following system:

$$\beta = \text{atanh}\left[\frac{1}{\rho}k_r\left(1 - m^2\right) + m^2\right] - g_r\frac{\text{atanh}m}{m} - \text{atanh}\left[\tanh^2\left(f_r\text{atanh}m + \frac{\beta}{2d}h^{ext}\right)\right] \tag{21}$$

$$\begin{aligned} m =& \tanh\left[f_r\text{atanh}m + \frac{\beta}{2d}h^{ext} + \right. \\ &\left. + \text{atanh}\left(\tanh\left(\beta + g_r\frac{\text{atanh}m}{m}\right)\tanh\left(f_r\text{atanh}m + \frac{\beta}{2d}h^{ext}\right)\right)\right] \end{aligned} \tag{22}$$

where $k_r = \frac{1-2dR_d(r)}{2drR_d(r)}$, $g_r = \frac{k_r}{1-k_r^2} + rR_r$, $f_r = \frac{1}{1+k_r} - (r+1)R_d\left(r\right)$. Such equations can be solved at fixed r in the variables β, m. For $h = 0$ the system reduces to a single fixed point equation for $m = M\left(m\left(r\right), r\right)$ while β is fixed by (21).

Computation of β_p

For the paramagnetic solution $m = 0$ (with $h = 0$) we get the following equation for $\beta(r)$:

$$\beta = \text{atanh}\left(\frac{1}{\rho r}\left[\frac{1}{2dR_d(r)} - 1\right]\right) - g_r$$

For $d \geq 3$, the maximum value at which a paramagnetic solution exists corresponds to the point $r = -1$. Therefore, the value of the critical point β_p is computed by taking the $r \to -1$ limit of Eq.(21):

$$\beta_p = \text{atanh}\left(1 - \frac{1}{z}\right) - z\left(\frac{z-1}{2z-1}\right) + \frac{z}{2d} \tag{23}$$

with $z = 2dR_d(-1)$.

Computation of β_m

Eqs. (21)-(22) implicitly define a function $m(r)$ such that $M(m(r), r) = m$, and thus also $\beta(r) = \beta(m(r), r)$. We seek to find the point $m^* = m(r^*)$ and $\beta_m = \beta(m^*, r^*)$ such that $\frac{d\beta}{dr}(m(r^*), r^*) = 0$. Taking the total derivative of $\beta(m(r), r)$ we get the equation to be solved

$$0 = \frac{d\beta}{dr} = \frac{\partial\beta}{\partial r} + \frac{\partial\beta}{\partial m}\frac{dm}{dr}$$

To compute $\frac{dm}{dr}$ we use its implicit definition,

$$\begin{aligned} 0 &= \frac{d}{dr}\{M(m(r), r) - m(r)\} \\ &= \left(\frac{\partial M}{\partial m}(m(r), r) - 1\right)\frac{dm}{dr} + \frac{\partial M}{\partial r}(m(r), r) \\ \frac{dm}{dr} &= -\frac{\frac{\partial M}{\partial r}(m(r), r)}{\frac{\partial M}{\partial m}(m(r), r) - 1} \end{aligned}$$

to get finally the 2×2 system in variables m, r:

$$M(m, r) - m = 0 \tag{24}$$

$$\frac{\partial\beta}{\partial r}(m, r)\left(\frac{\partial M}{\partial m}(m, r) - 1\right) - \frac{\partial M}{\partial r}(m, r)\frac{\partial\beta}{\partial m}(m, r) = 0 \tag{25}$$

Stability

The stability of a fixed point $m^* = m(r^*)$ can be analyzed by computing $\left.\frac{dM}{dm}\right|_{m^*}$. In particular, starting from the system (21)-(22) where r is implicitly defined as $r = R(\beta, m)$,the instability occurs when $\left.\frac{dM}{dm}\right|_{m^*} = 1$. Writing the original system using the definition of r we get $m = M(m, R(\beta, m))$ and $\beta = B(m, R(\beta, m))$. The equation we want to solve is

$$1 = \frac{dM}{dm} = \frac{\partial M}{\partial m} + \frac{\partial M}{\partial r}\frac{\partial R}{\partial m}$$

To compute $\frac{\partial R}{\partial m}$ we use again its implicit definition:

$$\begin{aligned} 0 &= \frac{\partial B}{\partial m} + \frac{\partial B}{\partial r}\frac{\partial R}{\partial m} \\ \frac{\partial R}{\partial m} &= -\frac{\frac{\partial B}{\partial m}}{\frac{\partial B}{\partial r}} \end{aligned}$$

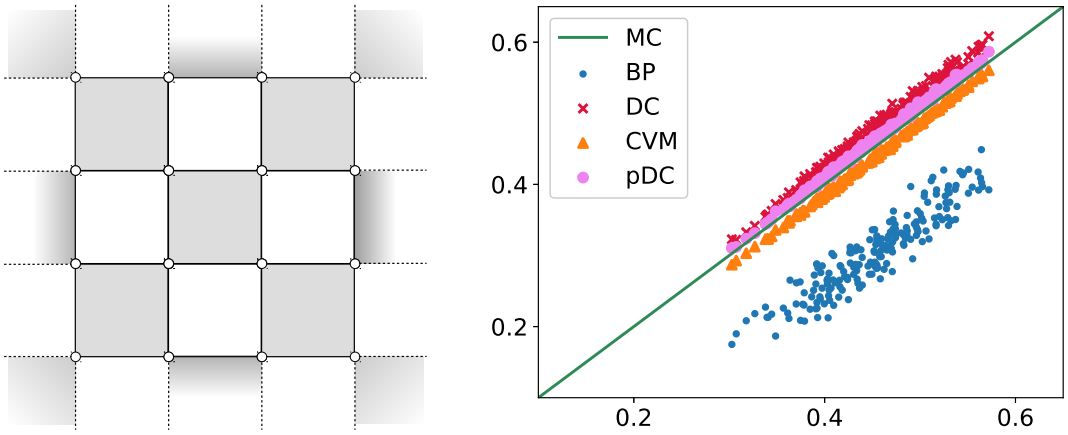

Figure 1. Left: distribution of non-overlapping plaquettes (in grey) on a 2-dimensional square lattice. Right: correlations on 2-dimensional Ising Model on square lattice of size $L = 10$ at $\beta = 0.36$, with 0 external field and couplings drawn from a uniform distribution in $(0.5, 1.5)$. Comparison of DC, pDC, BP and CVM

The final system to solve is

$$M(m, r) - m = 0 \tag{26}$$

$$\frac{\partial M}{\partial m} - \frac{\partial M}{\partial r} \frac{\frac{\partial B}{\partial m}}{\frac{\partial B}{\partial r}} = 1 \tag{27}$$

For $d \geq 3$, the solution becomes unstable exactly at the point (r_m, β_m) computed through 24-25.

C. D=2

On a 2-dimensional square lattice, the DC solution is qualitatively different w.r.t. $d \geq 3$ because the function $R(r)$ is logarithmically divergent for $r \to -1$. In such case the maximum value at which the paramagnetic solution exists ($\beta_p = 0.37693$) corresponds to the point $r_p = -0.994843$. The ferromagnetic solution turns out to be stable for $r_m < r < 0$ with $r_m = -0.99405$, corresponding to $\beta_m = 0.388448$ (the point (r_m, β_m) is found as a solution of Eq. 26-27). Therefore there exists a temperature interval $\beta_p < \beta < \beta_m$ in which no stable DC solution can be found.

For finite size lattices the DC solution can still be found numerically, showing similar performances with respect to CVM on both ferromagnetic and spin glass models (Fig. 1). However, especially on ferromagnetic systems DC solution is numerically unstable close to the transition β_p. One way to reduce numerical instability in such region is to decrease the interpolation parameter ρ, typically fixed to 1 for DC. Neverthless, the meaning of the DC(ρ) approximation in this case is not clear.

One possible way to improve the DC approximation is to take into account small loops explicitly. In particular, we consider a gaussian family of approximating distribution factorized over plaquettes of 2^dspins (d is the number of dimensions). Plaquettes are chosen in such a way that there is no overlap between links in the gaussian distribution. In this way, DC equation are exact on a plaquette tree with only site-overlaps. Results are shown in 1: plaquette-DC (pDC) is in general slighty better than standard DC and comparable to CVM.

D. Finite size corrections

In homogeneous models the gaussian covariance matrix can be diagonalized analytically even for a finite size lattice (of size L). Therefore we can compute finite size corrections to the DC solution at a fixed β, as shown in the following plot:

DC solution turns out to be in good agreement with MC results; on the other hand, BP does not take into account at all finite size corrections because of the local character of the approximation.

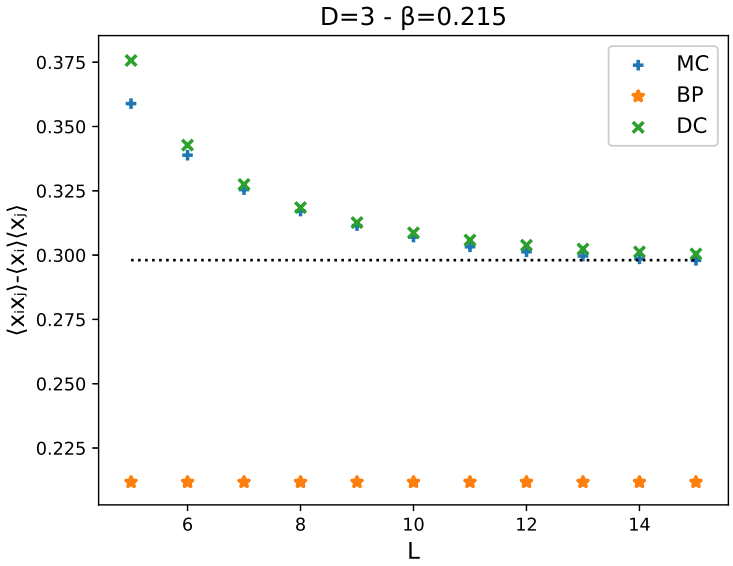

Figure 2. Finite size correction of equilibrium correlations at $\beta = 0.215$ on a 3-dimensional cubic lattice of size $L \in \{5, ..., 15\}$ with $h = 0, J = 1$. Comparison of BP and DC solutions with Monte-Carlo simulations.

E. Scaling of β_c in the high dimensional limit

Starting from the expression of the critical inverse temperature β_p it is possible to compute the $1/d$ expansion in the high-dimensional limit. We recall the expression of the critical temperature (23):

$$\beta_p = \text{atanh}\left(1 - \frac{1}{z}\right) - z\left(\frac{z-1}{2z-1}\right) + \frac{z}{2d}$$

where $z = 2dR_d(-1)$.

Defining $x = 1/d$ and expanding around $x = 0$ we get:

$$\frac{1}{2d\beta_p} = 1 - \frac{1}{2}d^{-1} - \frac{1}{3}d^{-2} - \frac{13}{24}d^{-3} - \frac{979}{720}d^{-4} - \frac{2039}{480}d^{-5} + O\left(d^{-6}\right).$$

This expansion is exact up to the d^{-4} order (the correct coefficient of d^{-5} is $-\frac{2009}{480}$) Fisher and Gaunt [3]. For comparison, Mean Field is exact up to the d^0 order, Bethe is exact up to the d^{-1} order, and Loop-Corrected Bethe and Plaquette-CVM are exact up to the d^{-2} order.

For the sake of completeness, we report the series expansion of $R_d(-1)$ around $x = 0$:

$$R_d(-1) = \frac{1}{2}d^{-1} + \frac{1}{4}d^{-2} + \frac{3}{8}d^{-3} + \frac{3}{4}d^{-4} + \frac{15}{8}d^{-5} + \frac{355}{64}d^{-6} + \frac{595}{32}d^{-7} + O\left(d^{-8}\right)$$

III. MULTISTATES VARIABLES

The method we presented is based on the possibility to fit the probability values of a discrete binary distribution with the density values of a univariate gaussian on the same support. When the model variables take $q > 2$ values there is no general way to fit single-node marginals with a univariate Gaussian distribution. One possible solution is to replace each q-state variable x_i with a vector of q (correlated) binary variables $\boldsymbol{s^i}$, where $s^i_\alpha \in \{-1, 1\}\, \forall \alpha = 1, .., q$, with the following constraint:

$$\sum_{\alpha=1}^{q} s^i_\alpha = 2 - q$$

In this way, for each node i, only configurations of the type $\boldsymbol{s}^i = \{1, -1, ..., -1\}$ (and its permutations) are allowed, in order to select just one of the q states for x_i. For each factor node a, such constraints can be implemented by adding a set of delta functions in the original probability distribution, which is now a function of the new binary variables $\boldsymbol{s}^i$. The correlations induced by these constraints on the spin components of each $\boldsymbol{s}^{\boldsymbol{i}}$ introduce short loops even when the original graph is a tree. Neverthless, it is still possible to write a set of matching equation similar to the 2-states case which is exact on trees.

---

[1] With the apparent ambiguity of the appearance of δ^k terms in $p(x)$, which is actually really not as problematic as it may seem as the distribution is defined up to a normalization factor; more precisely consider

$$p(x) = \lim_{\sigma \to 0} \frac{1}{Z_\sigma} \prod_{i \sim j} \hat{\psi}_{\sigma,ij}(x_i, x_j)$$

for factors $\hat{\psi}_{\sigma,ij}(x_i, x_j) = \psi_{ij}(x_i, x_j) (\mathcal{N}(x_i; 1, \sigma) + \mathcal{N}(x_i; -1, \sigma)) (\mathcal{N}(x_j; 1, \sigma) + \mathcal{N}(x_j; -1, \sigma))$.

[2] P. J. Davis, *Circulant Matrices* (1979).

[3] M. E. Fisher and D. S. Gaunt, Phys. Rev. **133**, A224 (1964).