Generalized mean field approximation for parallel dynamics of the Ising model

Hamed Mahmoudi and David Saad

Non-linearity and Complexity Research Group, Aston University, Birmingham B4 7ET, United Kingdom

The dynamics of non-equilibrium Ising model with parallel updates is investigated using a generalized mean field approximation that incorporates multiple two-site correlations at any two time steps, which can be obtained recursively. The proposed method shows significant improvement in predicting local system properties compared to other mean field approximation techniques, particularly in systems with symmetric interactions. Results are also evaluated against those obtained from Monte Carlo simulations. The method is also employed to obtain parameter values for the kinetic inverse Ising modeling problem, where couplings and local fields values of a fully connected spin system are inferred from data.

I. INTRODUCTION

Statistical physics of infinite-range Ising models has been extensively studied during the last four decades [1–3]. Various methods such as the replica and cavity methods have been devised to study the macroscopic properties of Ising systems with long range random interactions at equilibrium. Other mean field approximations that have been successfully used to study the infinite-range Ising model are naive mean field and the TAP approximation [2, 4], both are also applicable to broad range of problems [5, 6].

While equilibrium properties of Ising systems are well-understood, analysing their dynamics is still challenging and is not completely solved. The root of the problem is the multi-time statistical dependencies between time steps manifested in the corresponding correlation and response functions. Early attempts to analyze the dynamics of spherical spin models via mean field methods were based on the Langvin equations [3, 7] while a path integral analysis of the Glauber dynamics in spin glass systems was developed in [8, 9]. More recently, motivated by the need to infer parameter values in the kinetic inverse Ising modeling problem, several naive mean field approximations has been proposed for studying the dynamics of Ising system [10–12].

The Ising spin dynamics is usually described by a Markovian process whereby at every time step each spin σ_i is updated according to a probability that depends on effective field $h_i = \sum_j J_{ji}\sigma_j$. For the infinite range model, the interaction couplings J_{ij} are random variables sampled from a Gaussian distribution and the strength of couplings decreases with the system size. In symmetric networks *i.e.*, where $J_{ij} = J_{ji}$, the stationary state of system under Glauber dynamics can be described by Gibbs measure, a property that does not hold for asymmetric networks. In both cases, an *exact* transient analytical solution is unfeasible, and the evolution of dynamical quantities of interest such as magnetizations and correlations can be obtained only by time consuming numerical simulations.

It is therefore essential to develop accurate approximation methods for better understanding the dynamics of spin systems as well as for applications that require probabilistic inference of microscopic variables. One approach is to start from fully asymmetric networks which exhibit small correlations among spins at different times. This allows one to use central limit theorem and apply the Gaussian approximation for probability distributions of effective fields. Based on this property, an exact formalism was introduced for the dynamics of *fully asymmetric networks* in [10].

A central macroscopic quantity in studying the dynamics of spin glasses is the spin correlation function that describes the correlation between microscopic states at two time steps. In the present work we extend the mean field approximation introduced in [10] by incorporating information from the time dependent correlations. We still consider the effective fields to follow a Gaussian distribution but modify the standard deviation of the effective fields by considering the non-zero covariance between spins at two different time steps. We show how this modification improves the results, particularly for symmetric networks. Moreover, we provide an analytical recursive equation for calculating the covariance among spins at different time steps, a property which has been absent from other mean field methods.

To demonstrate the applicability of our approach as an inference tool for dynamical data we apply it for studying the network reconstruction problem. The latter is an inverse problem in which one assumes to have access to dynamical data from which time-dependent quantities such as magnetization and correlations could be calculated. The challenge is to infer model variable values that best describe the sample data; the model used in the current case is the Ising model, hence microscopic interaction and external field values should be inferred.

In the original equilibrium inverse Ising problem, the assumption is that the data is drawn from a Boltzmann-Gibbs distribution with an energy function that includes symmetric pairwise interactions [5, 13]. In the non-equilibrium variant of the problem [10, 12, 14] dynamical mean field methods are used to infer the best interaction values that

describe the data where samples are time-dependent. This problem has been actively studied in recent years, in particular in the context of neural network inference and protein-protein interactions [15, 16].

The paper is organized as follows. In Sec. II we briefly describe the model, introduce the macroscopic observables and the dynamic update rules. In Sec. III we use a Gaussian approximation which is exact for the fully asymmetric networks in the infinite-range model, but incorporate covariance values between spins at different time steps; we also derive a method for recursively calculating covariance values for spins at all times. In Sec. IV we introduce the kinetic inverse Ising problem and the corresponding application of our method as an inference tool. Section V focuses on the numerical results obtained for the two problems: (1) the forward calculation of macroscopic properties; and (2) inferring model values for the inverse Ising problem. We compare our results to other mean field approximation for both cases as well as against direct numerical simulation observed in Monte Carlo simulations. Section(VI) introduces concluding remarks and offers future research directions.

II. MODEL

We consider an Ising spin system comprising N binary variables $\vec{\sigma} = \{\sigma_1, .., \sigma_N\}$ which are connected through random couplings J_{ij}. The interactions can be asymmetric ($J_{ij} \neq J_{ji}$) or symmetric ($J_{ij} = J_{ji}$). Following [3, 7] we introduce a parameter k which measures the degree of symmetry in the couplings

$$J_{ij} = J_{ij}^{\text{s}} + k J_{ij}^{\text{as}} \ , \tag{1}$$

where J^{s} and J^{as} are symmetric and antisymmetric matrices, respectively. The off-diagonal elements of $J_{ij} \ \forall i, j$ are random Gaussian variables with zero means and variance

$$\left[(J_{ij}^{\text{s}} - J_0/N)^2 \right]_J = \left[(J_{ij}^{\text{as}})^2 \right] = \frac{1}{N^2} \frac{1}{1+k^2} \ . \tag{2}$$

We define the self connectivity to be zero ($J_{ii} = 0$). The parameter $k = 0$ corresponds to fully symmetric networks whereas $k = 1$ represents fully asymmetric networks. The interactions among spins determine the dynamics of system; in the parallel update scheme, all spins are updated simultaneously

$$\sigma_i(t+1) = \begin{cases} +1 & \text{with probability} \ \frac{1}{1+e^{-2\beta(h_i(t+1)+\theta_i(t+1))}} \\ -1 & \text{with probability} \ \frac{1}{1+e^{2\beta(h_i(t+1)+\theta_i(t+1))}} \end{cases} \ , \tag{3}$$

where $h_i(t)$ is the effective local field acting on spin i at time step t

$$h_i(t) = \sum_{j \in \partial i} J_{ji} \, \sigma_j(t-1) + \theta_i(t) \quad , \tag{4}$$

and the parameter β, analogous to inverse temperature, is a measure of the overall strength of the interactions. The notation ∂i stands for the set of spins neighboring target spin i with $J_{ji} \neq 0$. In this paper we only study the fully connected networks; however, this notation is use for consistency. This particular choice of sequential update [Eq. (3)] is known as *Glauber dynamics*.

In this paper we focus only on synchronous updates and aim to derive local system properties such as magnetizations and correlations directly from the Glauber dynamics. The joint probability distribution over all the spin histories $p(\vec{\sigma}(0), \ldots, \vec{\sigma}(t))$ has the following simple Markovian form

$$p(\vec{\sigma}(0) \ldots, \vec{\sigma}(t)) = \prod_{s=1}^{t} W[\, \vec{\sigma}(s) \,|\, \vec{h}(s)] \, p(\vec{\sigma}(0)) \ , \tag{5}$$

where W is the transition matrix, whose elements are defined by Eq. (3). The evolution of a a single spin is (trivially) defined by summing over the histories of all spins except that of spin i,

$$p_i(\sigma_i(0), \ldots, \sigma_i(t)) = \sum_{\vec{\sigma}_{\backslash i}(0), \ldots, \vec{\sigma}_{\backslash i}(t)} p(\vec{\sigma}(0), \ldots, \vec{\sigma}(t)) \ , \tag{6}$$

and similarly for pairwise joint probability of the histories of two spins one omits the corresponding trajectories of σ_i and σ_j at all times

$$p_{ij}(\sigma_i(0), \ldots \sigma_i(t), \sigma_j(0), \ldots, \sigma_j(t')) = \sum_{\vec{\sigma}_{\backslash i,j}(0), \ldots, \vec{\sigma}_{\backslash i,j}(t)} p(\vec{\sigma}(0), \ldots, \vec{\sigma}(t)) \quad . \tag{7}$$

It can be verified that marginal distribution of spin trajectories [Eq. (6)] can be expressed in terms of the joint probability distribution of neighboring spins

$$p_i(\sigma_i(0),...,\sigma_i(t)) = p_i(\sigma_i(0)) \sum_{\vec{\sigma}_{\partial i}(0)...\vec{\sigma}_{\partial i}(t-1)} p^{(i)}(\vec{\sigma}_{\partial i}(0),...,\vec{\sigma}_{\partial i}(t) / \sigma_i(0),...,\sigma_i(t))$$

$$\prod_{s=1}^{t} \frac{\exp\left(\beta\sigma_i(s)(\sum_{j\in\partial i} J_{ji}\sigma_j(s-1) + \theta_i(s))\right)}{2\cosh\left(\beta(\sum_{j\in\partial i} J_{ji}\sigma_j(s-1) + \theta_i(s))\right)} \qquad (8)$$

The intuition comes from the concept of cavity in statistical mechanics of disordered systems. The last term on the right hand side is the conditional probability of observing the trajectory $\{\sigma_i(0),...,\sigma_i(t)\}$ when the configuration of neighboring spins is given. The second term, representing the joint probability distribution of neighboring spins, must be evaluated on the cavity graph. Therefore, we denote it by $p^{(i)}(\vec{\sigma}_{\partial i}(0),...,\vec{\sigma}_{\partial i}(t) / \sigma_i(0),...,\sigma_i(t))$. The first term in Eq. (8) represents the initial conditions for the process. The site magnetization at time t can be derived from Eq. (8)

$$m_i(t) = p_i(\sigma_i(0)) \sum_{\sigma_i(0\to t)} \sigma_i(t) \sum_{\vec{\sigma}_{\partial i}(0\to t-1)} p^{(i)}(\vec{\sigma}_{\partial i}(0\to t) / \sigma_i(0\to t))$$

$$\prod_{s=1}^{t} \frac{\exp\left(\beta\sigma_i(s)(\sum_{j\in\partial i} J_{ji}\sigma_j(s-1) + \theta_i(s))\right)}{2\cosh\left(\beta(\sum_{j\in\partial i} J_{ji}\sigma_j(s-1) + \theta_i(s))\right)} , \qquad (9)$$

where we use $(s \to t)$ to indicate the time path through $\{s, s+1, ..., t\}$. Clearly, computing marginal distribution [Eq. (6)] and consequently the magnetization [Eq. (9)] is intractable. It requires a summation of order $\mathcal{O}(2^{T\|c_i\|})$ where c_i is the number of neighboring spins for target site i. We will show that even in the infinite-range spin glass systems, this prohibitive complexity remains.

III.   INFINITE-RANGE MODEL

Here, we focus on infinite-range spin glass systems, where the number of neighbors per spin is large $c_i \propto \mathcal{O}(N)$, and in turn interaction couplings scale inversely to the system size; such that in the thermodynamic limit $(N \to \infty)$ spins become weakly correlated. In the old statistical mechanics literature this was the basic idea behind naive mean field theory for the Sherrington-Kirkpatrick (SK) model. In order to implement similar assumption in non-equilibrium system, we introduce the probability of observing trajectories (time path) for the effective fields [Eq. (4)]

$$p_i(h_i(1 \to t)) = \sum_{\vec{\sigma}_{\partial i}(0\to t-1)} p(\vec{\sigma}_{\partial i}(0 \to t-1)) \prod_{s=1}^{t} \delta(h_i(s) - \sum_{j\in\partial i} J_{ji}\sigma_j(s-1)) . \qquad (10)$$

Note that introducing this probability will not facilitate our calculations even in the infinite-range model. In fact, computing $p_i(h_i(1 \to t))$ is as difficult as computing the joint probability distribution of neighboring spins in Eq. (9). This is due to the fact that dynamics of neighboring spins are correlated through the target spin i. In other words, the target spin i affects its neighboring spins and consequently they become correlated.

However, for fully asymmetric networks one can assume that the effective field $h_i(t)$ admits a Gaussian distribution, due to the small correlation between different sites at successive steps [10]. Consequently, the time single-spin magnetization is given by

$$m_i(t) = \int Dx \ \tanh\left[\beta\left(\sqrt{\Delta_i(t)} \ x + \sum_{j\in\partial i} J_{ji}m_j(t-1) + \theta_i(t)\right)\right] , \qquad (11)$$

where $Dx = \frac{dx}{\sqrt{2\pi}}e^{-x^2/2}$ is the Gaussian probability density and $\Delta_i(t) = \sum_{j\in i} J_{ji}^2(1 - m_j^2(t-1))$. Note that Eq. (11) is derived under the assumption of vanishing non-diagonal elements of the covariance matrix ($C_{ij} = 0$ $i \neq j$). Under this assumption the time dependent magnetization admits a Morkovian equation where at each time only information from the previous time step is required.

III   INFINITE-RANGE MODEL

The results are exact in the thermodynamic limit ($N \to \infty$) for any set of couplings and external fields, as it relies only on the central limit theorem. Equation 11 should be compared with the naive mean field approximations introduced in [17]

$$m_i(t) = \tanh\left[\beta\left(\sum_{j \in \partial i} J_{ji} m_j(t-1) + \theta_i(t)\right)\right] . \tag{12}$$

The naive mean field equation can be derived from Eq. (11) by an expansion in interaction strength. In [18] it was shown that such expansion, in the first order, will give rise to Eq. (12) and in second order will produce a *TAP-like* equation [12, 18]. Here, we will improve on the analysis of the dynamics by taking into account temporal correlations. The idea is to assume a Gaussian distribution for the effective fields but consider a non-zero covariance matrix. The approach is similar to the one introduced by Amari and Maginu [19], where they studied non-equilibrium dynamics of autocorrelation associative memory. Assuming the Gaussian distribution for the effective field, one can write a general description for the time dependent magnetization

$$m_i(t) = \int Dx \, \tanh\left[\beta\left(\sqrt{V_i(t)}\, x + \sum_{j \in \partial i} J_{ji} m_j(t-1) + \theta_i(t)\right)\right] , \tag{13}$$

where $V_i(t) = \langle h_i^2(t)\rangle - \langle h_i(t)\rangle^2$ is the autocovariance of the effective field at site i. In order to perform this integral we first need to compute V_i. From the definition of h_i we can substitute the effective field by its explicit expression in terms of spin values at an earlier time step

$$V_i(t) = \sum_{j \in \partial i} J_{ji} J_{ki} \, \langle \sigma_j(t-1)\sigma_k(t-1) - m_j(t-1)m_k(t-1)\rangle . \tag{14}$$

If connectivities are exactly evenly distributed or if we assume non-diagonal elements of the covariance matrix to be zero, the above equations reduces to Eq. (11). However, these conditions are not satisfied in general and the non-diagonal elements must be taken into account. In what follows we calculate $V_i(t)$ in terms of earlier correlation functions taking into account correlations between spins and effective fields at earlier times. We will show how using more correlations improves our predictions. By definition, for $V_{i,j}(t, s+1)$ we have

$$\begin{aligned} V_{i,k}(t, s+1) &= \langle h_i(t) h_k(s+1)\rangle - \langle h_i(t)\rangle \langle h_k(s+1)\rangle \\ &= \sum_{j \neq k} J_{jk} \, (\langle \sigma_j(s) h_i(t)\rangle - m_j(s)\langle h_i(t)\rangle) . \end{aligned} \tag{15}$$

The key idea here is to substitute $\langle h_j(s)\sigma_k(t-1)\rangle$ by early time correlations. Without loss of generality we assume that $s \leq t$ and

$$\langle h_i(t)\sigma_j(s)\rangle = \sum_{\sigma_j(0 \to s)} \int dh_i(t) \prod_{s'=1}^{s} dh_j(s') \, \sigma_j(s)\, h_i(t)\, p(\sigma_j(0 \to s), h_j(1 \to s), h_i(t)) , \tag{16}$$

where the summation is taken over all possible configurations of σ_j at all time steps from 0 to s. The two random variable $h_i(t)$ and $\sigma_j(s)$ are correlated through $h_j(1 \to s)$ in time and the joint probability distribution is denoted by $p(\sigma_j(0 \to s), h_j(1 \to s), h_i(t))$. Now we use Bayes theorem to express the joint probability distribution in a more appropriate form

$$\langle h_i(t)\sigma_j(s)\rangle = \sum_{\sigma_j(0 \to s)} \int dh_i(t) \prod_{s'=1}^{s} dh_j(s') \, \sigma_j(s)\, h_i(t)\, p(\sigma_j(0 \to s)|h_j(1 \to s))\, p(h_i^{(j)}, h_j^{(j)}) . \tag{17}$$

Here $h_i^{(j)}$ stands for the *cavity* effective field acting on spin i when the time path for spin j is fixed. Since the cavity effective fields h_i and h_j are random variables their joint probability distribution is represented by a multivariate Gaussian

$$p(h_i^{(j)}, h_j^{(j)}) = \frac{1}{\sqrt{2\pi V_i V_j (1-x)}} \exp\left(-\frac{(h_i^{(j)} - \langle h_i^{(j)}\rangle)^2}{2V_i\,(1-x)} - \frac{(h_j^{(j)} - \langle h_j^{(j)}\rangle)^2}{2V_j\,(1-x)} + V_{ij}\frac{(h_i - \langle h_i^{(j)}\rangle)(h_j^{(j)} - \langle h_j^{(j)}\rangle)}{V_j V_i\,(1-x)}\right) , \tag{18}$$

where $x = V_{ij}^2/(V_i V_j)$ and is small but non-zero in the thermodynamic limit. In deriving equation (18) we assume that the cavity covariance function is equal to the normal covariance function *i.e.* $V_{ij}^{(j)} = V_{ij}$. This turns out to be correct in the thermodynamic limit $N \to \infty$ for the SK model since the couplings are drawn independently and at random. The conditional probability $p(\sigma_j(0 \to s)|h_j(1 \to s))$ follows the Glauber dynamics rule defined in Eq. (3).

In order to perform the integral we introduce the explicit expression of the field $h_i(t) = h_i^{(j)}(t) + J_{ji}\sigma_j(t-1)$. We then have

$$\langle h_i(t)\sigma_j(s)\rangle = \frac{1}{\sqrt{2\pi V_i V_j(1-x)}} \int dh_i(t) \prod_{s'=1}^{s} dh_j(s') \left(h_i^{(j)}(t) + J_{ji}\sigma_j(t-1) \right) \sigma_j(s)$$
$$\prod_{s'=1}^{s} \frac{e^{\beta\sigma_j(s')\left(h_j(s')+\theta_j(s')\right)}}{2\cosh(\beta\left(h_j(s')+\theta_j(s')\right))} \; p(h_i^{(j)}(t), h_j^{(j)}(s)) \; . \tag{19}$$

By performing integration by part in Eq. (19) we get

$$\langle h_i(t)\sigma_j(s)\rangle = V_{ij}(t,s) \, G_j(s) + m_j(s) \langle h_i^{(j)}\rangle(t) + J_{ji}\langle\sigma_j(t-1)\sigma_j(s)\rangle \; , \tag{20}$$

where $G_i(s)$ is the generalized response function

$$G_i(s) = \int \; Dx \; \left[1 - \tanh^2 \left(\beta(\sqrt{V_j(s)} \; x + \sum_{k \in \partial j} J_{kj}m_k(s-1) + \theta_j(s)) \right) \right] \; . \tag{21}$$

Equation (20) is a remarkable result for dynamics of SK model as it relates the correlations between a single spin with the effective fields of its neighbors to the earlier time correlations. Substituting Eq. (20) into Eq. (19) we get

$$V_{i,j}(t,s) = \sum_{k \neq i} J_{ki} \left(G_k(t-1) V_{ik}(t-1,s) + J_{ik}C_k(t-1,s-1) \right) \; , \tag{22}$$

or in matrix notation

$$V(t,s) = J^\mathsf{T} G(t-1)V(t-1,s) + J^\mathsf{T}C(t-1,s-1)J \; . \tag{23}$$

where the matrix C on the right hand side of (22) and (23) is the auto-covariance function $C_{ik}(t,s) = \delta_{ik}\left(\langle\sigma_k(t)\sigma_k(s)\rangle - m_k(t)m_k(s)\right)$. A particular use of Eq. (23) is to find an estimate for the auto-covariance function $V_i(t)$ as it is required for evaluating the magnetizations. We therefore express it in terms of covariance function values at different times as

$$V_i(t) = \sum_{j \neq i, k \neq i} J_{ki}J_{ji}, G_j(t-1)G_k(t-1) \, V_{jk}(t-1,t-1) + \sum_k J_{ki}^2(1-m_k^2(t-1)) \; . \tag{24}$$

Higher order of covariance function can be computed using Eq. (20)

$$V(t,t+n) = V(t,t) \prod_{s=1}^{n-1} [\, G(t+n-s)\,J\,] + \sum_{s=1}^{n-1} J^\mathsf{T} C(t+n-s,t-1)\,J \prod_{s'=1}^{s-1} [\, G(t+n-s)\,J\,] \; , \tag{25}$$

where for the special case of $n = 1$ we have

$$V(t,t+1) = V(t,t)G(t)J + J^\mathsf{T}C(t-1,t)\,J \; . \tag{26}$$

Note that V is the covariance function for effective fields and is related to the spin covariance $V = J^\mathsf{T}\mathcal{C}J$ where $\mathcal{C}$ is the full covariance matrix. Assuming that J^{-1} exists we have

$$\mathcal{C}(t-1,t) = \mathcal{C}(t-1,t-1)JG(t-1) + C(t-1,t) \; , \tag{27}$$

where $C(t-1,t)$ is the the autocorrelation ($C_{ij} = 0$ for $i \neq j$). This equation provides us with an approximation for the non diagonal elements of covariance function between pairs of spins in SK model at different times.

$$\mathcal{C}_{ij}(t-1,t) = \sum_k \mathcal{C}_{ik}(t-1,t-1)J_{kj}G_j(t-1) \qquad \text{for i} \neq \text{j} \; , \tag{28}$$

Diagonal terms of covariance function must be computed through mean field equations. Since the effective fields are randomly distributed (drawn from a Gaussian distribution) the covariance function follows the bivariate Gaussian distribution

$$C_i(t-1,t) = \frac{1}{\sqrt{det(2\pi V_i)}} \int dh_i(t) \int dh_i(t-1) \tanh(\beta h_i(t) + \beta\theta_i) \tanh(\beta h_i(t-1) + \beta\theta_i) \tag{29}$$

$$\exp\left(-\frac{(h_i(t) - \langle h_i\rangle(t))^2}{2V_i(t)(1-x)} - \frac{(h_i(t-1) - \langle h_i(t-1)\rangle)^2}{2V_i(t-1)(1-x)} + V_i(t,t-1)\frac{(h_i(t) - \langle h_i(t)\rangle)(h_i(t-1) - \langle h_i(t-1)\rangle)}{V_i(t)V_i(t-1)(1-x)}\right).$$

We like to point out a major difference between equations (27) and (30) for the covariance matrix and the one used in the mean field approximation of fully asymmetric networks [10]. For the fully asymmetric mean field approximation, the covariance matrix of spins at two successive time is fully determined by equation (27). However, here the auto-covariance elements must be calculated separately according to equation (30).

The Gaussian assumption for the probability distribution of effective fields is correct only in fully asymmetric networks but it can be used as an approximation in networks with symmetric or partially symmetric connectivities. We can improve the approximation by taking into account the non diagonal elements of covariance function within the Gaussian approximation for the effective fields. This approximation is similar to [19] and [20] introduced in the context of statistical neuroscience.

IV. KINETIC INVERSE ISING PROBLEM

Our proposed mean field approximation can be used to solve the non-equilibrium inference problems in densely connected networks. To examine the efficacy of the method we apply it to a network reconstruction inference task - the kinetic inverse Ising problem. The task is to infer coupling interactions J_{ij} and external fields h_i of an Ising model given time dependent magnetizations $m_i(t)$ and covariances $C_{ij}(t,s)$ observed from data. In the setup used we sample the magnetizations and correlations of a system evolving according to the Glauber dynamics [21].

In the current example, we focus on inferring coupling interactions from $C(t,t-1)$ and $m(t)$. However, the framework used enables one to refine the inferred values by considering a broader range of covariance matrices of various time differences. We start from Eq. (27); the connectivity matrix appears explicitly on the right hand side and implicitly in the vector G. Following [10] we obtain the connectivity matrix by inverting equation (27). Introducing $B(t) = JG(t-1)$ we have

$$B(t-1) = \mathcal{C}^{-1}(t-1,t-1)\left[(1-\delta)\right]\mathcal{C}(t,t-1), \tag{30}$$

where $[(1-\delta)]\mathcal{C}(t,t-1)$ contains *only non-diagonal terms* of the covariance matrix $\mathcal{C}(t,t-1)$. The right hand side is entirely given by data, while the left hand side depends directly on the couplings. In order to compute the connectivities we use an iterative procedure proposed in [10]. We express $V(t)$ in terms of known quantities; by definition, $V_i(t)$ is the covariance matrix of the effective fields whereas $\mathcal{C}$ stands for covariance matrix of the spins. These two matrices are clearly related as $V(t) = J^{\mathsf{T}}\,\mathcal{C}(t-1)\,J$. Multiplying the connectivity matrix J by an identity matrix GG^{-1} from both sides simplifies the equation since JG is given by empirical data [see Eq. (30)]. The final expression for the covariance matrix V is

$$V(t) = G^{-1}(t)\left[GJ^{\mathsf{T}}\,\mathcal{C}(t-1)\,JG\right]G^{-1}(t). \tag{31}$$

The *diagonal elements* of $V(t)$ appear from computing magnetizations and correlations on the right hand side. At each time step we use the following three set of mean field equations

$$m_i(t) = \int Dx\, \tanh\left(\beta(\sqrt{V_i(t)}\,x + \sum_{j\in\partial i} J_{ji}m_j(t-1) + \theta_i(t)\right)$$

$$G_i(t) = \int Dx\,\left[1 - \tanh^2\left(\beta\left(\sqrt{V_i(t)}\,x + \sum_{j\in\partial i} J_{ji}m_j(t-1) + \theta_i(t)\right)\right)\right]$$

$$V_i(t) = \frac{\mathcal{D}_i}{G_i^2(t)},$$

where $\mathcal{D}_i = \sum_{j,k} B_{ij}(t)C_{jk}(t-1)B_{ki}(t)$ is calculated directly from the data as mentioned before. We start from an initial value for $\Delta_i(t) \equiv \sqrt{V_i(t)}$. Given $m_i(t)$, from the empirical data, we solve the first equation to find

$\sum_{j\in\partial i} J_{ji}m_j(t-1) + \theta_i(t)$. We then substitute $\sum_{j\in\partial i} J_{ji}m_j(t-1) + \theta_i(t)$ into the second equation to compute $G_i(t)$. The third equation updates $\Delta_i(t)$. Once the iterative process has converged we use $G_i(t)$ and $V_i(t)$ to extract connectivities and external fields values from Eq. (30).

Note that the only difference between our generalized mean field theory and the one proposed in [10] is the third equation. In our method we use Eq. (30) to compute the data, whereas for the mean field method of fully asymmetric networks [10] $V_i(t)$ follows a simpler equation $V_i(t) = \sum_{j\in i} J_{ji}^2(1 - m_j^2(t-1))$. For the naive mean field approximation [12] the reconstructed connectivities are given by

$$J = A\, C(t,t-1)\, C^{-1}(t-1,t-1)\ , \tag{32}$$

where $A = \sum_{j\in i} J_{ji}^2(1 - m_j^2(t-1))$.

V.   RESULTS

We investigate the performance of generalized mean field approximation both for calculating macroscopic quantities (forward problem) and inferring coupling values from data (inverse problem). For the former, given connectivities and external fields we compute time-dependent magnetization, auto-covariance and covariance matrix for fully connected asymmetric networks evolving via parallel updates. We compare our results to the mean field approximation introduced for the fully asymmetric networks [10] and the naive mean field approximation [17]. We evaluate results obtained by the different methods against numerical simulations based on the Glauber dynamics defined in Sec. II.

For the kinetic inverse Ising problem we sample a network instance at random, and generate correlations and magnetizations by heavily sampled Glauber dynamics. We then infer the connectivities values based on the different mean field methods described above.

A.   Forward problem

In this case network connectivities are known and time dependent magnetizations are computed according to equations (11),(12) and (13) for the mean field, naive mean field and generalized mean field methods, respectively. To compute the time dependent covariance matrix we use Eq. (27) in the generalized mean field method. Note that Eq. (27) provides results only for non diagonal terms. The auto-covariance functions can be estimated by Eq. (30). Note also that we can not compute the covariance function by the naive mean field method and mean field method for the fully asymmetric networks. Higher order of covariance functions can be also computed by Eq. (25).

We examine the performance of different methods by computing the mean squared error defined as

$$\delta(t) \;=\; \frac{1}{N}\sum_{i=1}^{N}(m_i^{\text{GD}}(t) - m_i^{\text{estimate}}(t))^2\ ,$$

$$\Delta(t) \;=\; \frac{1}{N(N-1)}\sum_{i,j}^{N}(C_{ij}^{\text{GD}}(t,t-1) - C_{ij}^{\text{GMF}}(t,t-1))^2\ , \tag{33}$$

where $\delta(t)$ is the mean squared error of time-dependent magnetizations obtained by the different methods and $\Delta(t)$ the corresponding mean squared error for the covariance matrices obtained. Here m_i^{GD} and C_{ij}^{GD} stand for magnetization and covariance obtained by numerical simulations of systems evolving via Glauber dynamics, respectively. Note that for $\Delta(t)$ only generalized mean field approximation is considered as covariance matrices cannot be calculated by the other two methods. Fig. 1 shows results for mean squared errors δ and Δ averaged over T time steps (here $T = 10$). Left panel compares averaged mean squared errors for magnetization $\bar{\delta} = 1/T\sum_{t=1}^{T}\delta(t)$ computed by the different methods for fully symmetric networks. Clearly, generalized mean field approximation performs better for the various β values. The mean field approximation [Eq. (11)] estimates correctly the magnetizations for small β values but the error grows for lower temperatures. The right panel presents mean squared error results for the averaged covariance function at two successive times $\bar{\Delta} = 1/T\sum_{t=1}^{T}\Delta(t)$ computed by the generalized mean field approximation for both fully asymmetric and fully symmetric networks. For small β values, the covariance matrices predicted by our method are similar to those obtained by numerical simulations. As β increases it fails in predicting time dependent covariance for fully symmetric networks.

Figure 2 shows the performance of the generalized mean field method in calculating the auto-covariance function $C_i(t,t-L)$ for various values of L in fully symmetric networks. Due to the parallel nature of dynamics, even values of L give rise to higher auto-covariance values and consequently higher errors.
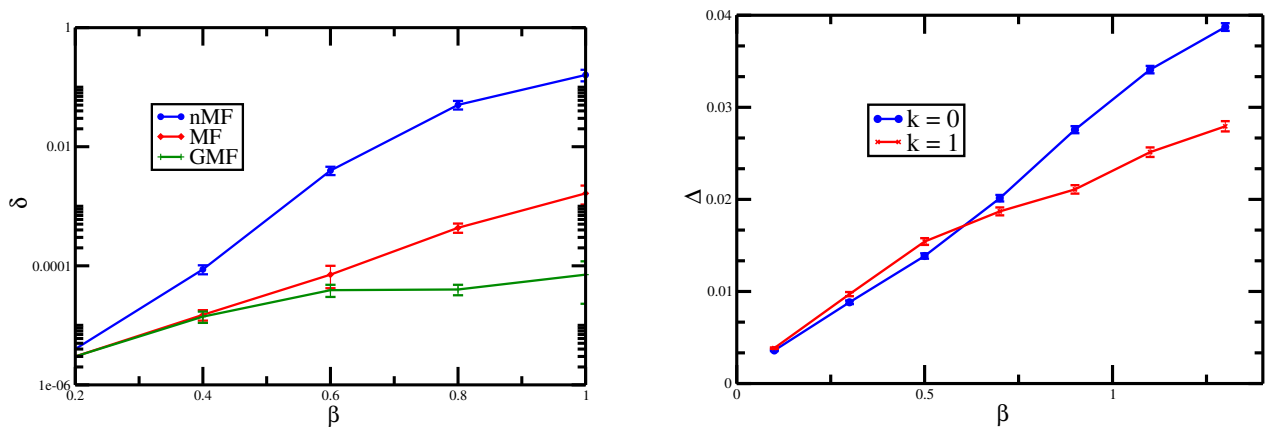
V RESULTS

FIG. 1: Left panel: Squared deviation of spin averages $\bar{\bar{\delta}}$ evaluated by different methods for a fully symmetric network ($k = 0$). The results are averaged over 10 time steps. Right panel: Squared deviation of covariance function computed by generalized mean field (GMF) method. Results are shown for both symmetric and asymmetric networks. In all simulations, $N = 100$, $T = 10$ and the sample size for the numerical simulations is 10^6. Results are averaged over 10 network instances, both mean values and standard deviations are shown (color online).

The left panel in Fig.2 shows the mean squared error results of the calculated auto-covariance function. As β increase the prediction becomes less accurate. The right hand side panel shows the point-wise auto-covariance function for different L values. Since spins are correlated in a parallel manner, the auto-covariance between two successive time steps is small. However for even L values we observe non-zero auto-covariance values. The results are compared to the numerical simulations (red points). We observe that the generalized mean field method almost always over estimates the auto-covariance function values. To highlight a qualitative comparison between different methods we show a
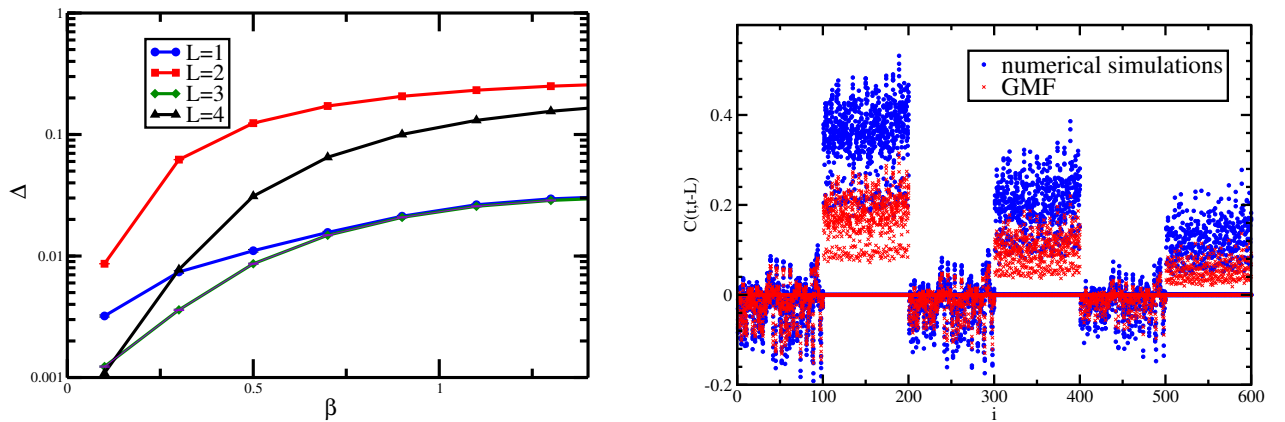

FIG. 2: Left panel: Squared deviation of spins auto-covariance as a function of β. Results are obtained by the generalized mean field method for symmetric networks and are averaged over 10 network samples and 10 time steps. Error bars are smaller than the symbols size. System size is 100 and external fields are set to zero. Right panel: Point-wise auto-covariance function $C_i(t, t - L)$ for different values of L in one single network sample of symmetric networks. The system size is $N = 100$ and $\beta = 0.7$. Calculated values are shown in red while values from numerical simulations are shown in blue (color online).

scatter plots for the local magnetizations obtained by the different methods versus those obtained from numerical simulations in Fig. 3. The results are for different values of $\beta = \{0.5, 0.7, 0.9, 1.1\}$. For small β values all methods predict the local magnetizations accurately. As we increase the value of β, naive mean field starts failing in its prediction. In all cases, generalized mean field theory outperforms the other two methods.
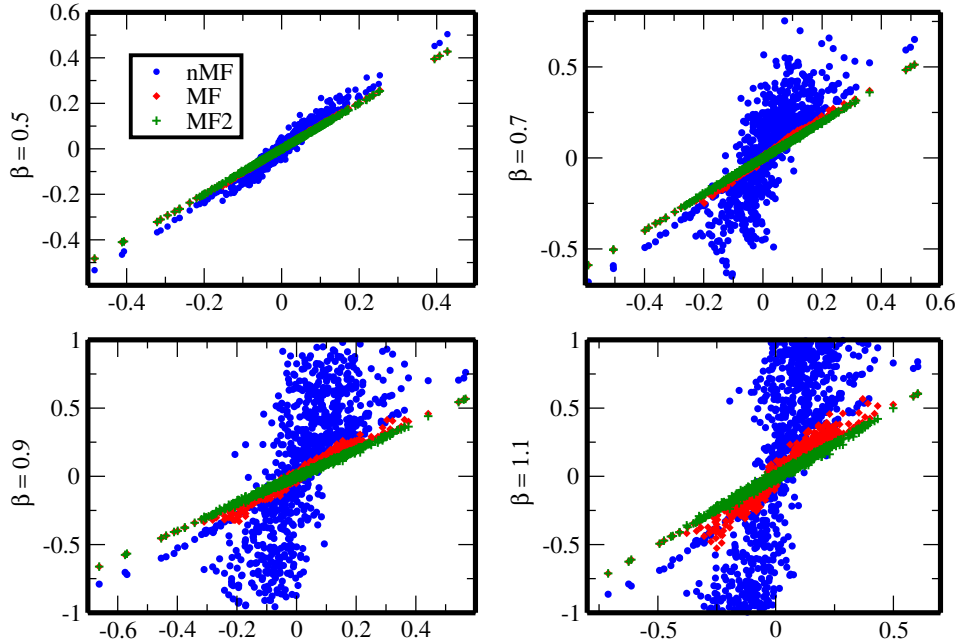
FIG. 3: Scatter plot of local magnetizations for fully symmetric networks($N = 100$) obtained by various methods, plotted against values obtained from simulation (x-axis). Green points represent results obtained by the generalized mean field method, blue points by naive mean field and red points by the mean field method (marked as MF2, color online). Results are for $\beta = 0.5, 0.7, 0.9, 1.1$ respectively.

B. Inverse problem

We now examine the performance of the different methods for the inverse Ising inference problem. We start with a fully connected spin glass systems with zero external fields and randomly sampled interaction values from Gaussian distribution of zero mean and unit variance; and then generate time dependent magnetization and covariance functions by applying Glauber dynamics. We examine both cases of symmetric and asymmetric couplings. The connectivities can be inferred by using Eqs. (32) and (32). Similar to the previous section we define mean squared error measure for the estimated error in the inferred connectivities

$$\mathcal{I} = \frac{1}{N(N-1)} \sum_{ij} \left(J_{ij} - J_{ij}^{\text{rec}} \right)^2 \tag{34}$$

where J_{ij}^{rec} are the estimated (reconstructed) connectivities inferred by the corresponding mean field method and J_{ij} are the original couplings.

The results, presented in Fig.4, show that the generalized mean field approximation outperforms the other two methods for different β values in symmetric networks. The results are averaged 10 time steps and over 10 networks samples. Indeed, it provides accurate results (compared to other methods) even for considerably high β values. To observe the comparison in a more qualitative manner, we show the scatter plot of reconstructed connectivities versus the true connectivities in Fig. 5. The generalized mean field method infers perfectly the connectivity couplings for fully symmetric network even at high temperature, especially compared to mean field and naive mean field results.

VI. CONCLUSION

In this paper, we studied the dynamics of spin glass systems and devised an inference method for the inverse Ising model. We focused on the Sherrington-Kirkpatrick model and generalized the mean field approximation, originally developed for the fully asymmetric networks, by taking into account information from the earlier time steps.

A key feature of fully asymmetric networks is to have negligible correlations between non-identical spins at different time steps. In our proposed method, the correlation between microscopic states at two time steps have been incorporated into the dynamics. The modification to the mean field approximation shows a great improvement in
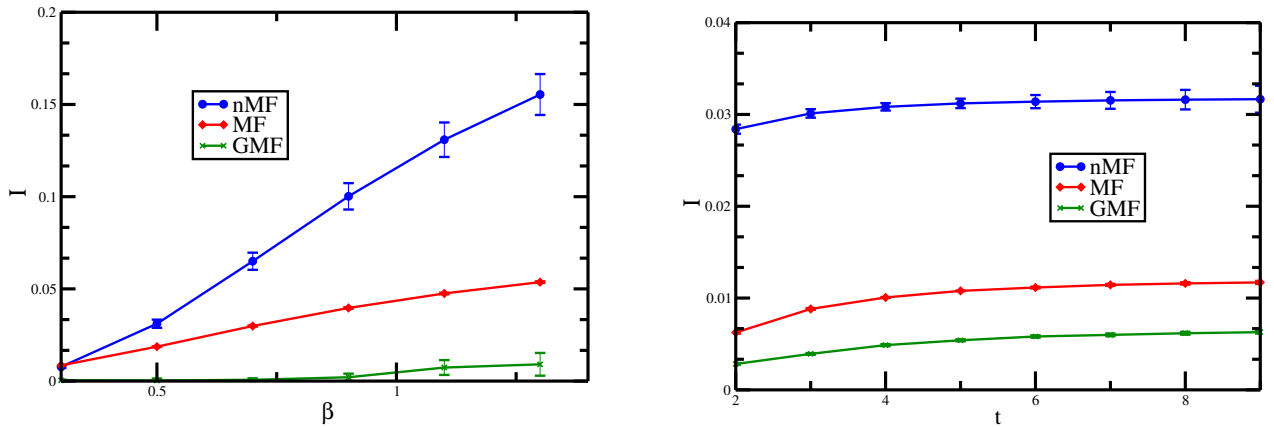
FIG. 4: Mean squared error for the reconstructed connectivities - $\mathcal{I}$ in symmetric networks with $N = 100$ averaged over 10 network instances. Left panel shows the performance of different methods for various β values averaged over 10 consecutive time steps. Right panel is the mean squared error for $\beta = 0.7$ in the different time steps. For results obtained by mean field method and generalized mean field method error bars are smaller than the symbols size (color online).
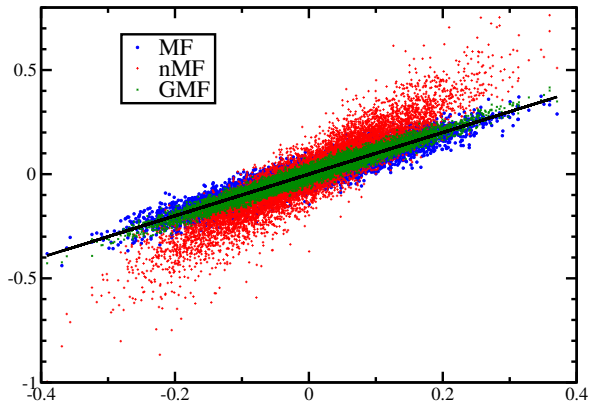

FIG. 5: Scatter plot of reconstructed connectivities in fully symmetric networks with $N = 100$ and $\beta = 0.7$. Green points represent reconstructed connectivities obtained by the generalized mean field method, blue points by mean field method of fully asymmetric networks and red points by the naive mean field method (color online).

predicting local properties of the system in both forward and inverseproblems; the former aims at predicting macroscopic properties dynamically, while the latter aims at inferring microscopic values from data in specific instances. Another advantage of our proposed method is that it enables one to compute time dependent correlations at different time steps, a quantity that cannot be studied by other mean field approximations. The performance of our method is evaluated by extensive numerical simulations.

Our numerical experiments show that for the fully asymmetric networks the improvement due to the generalization, over the mean field approximation, is modest as the correlation among the spins at different time steps is small. However, for symmetric networks our method shows significant improvement over other methods and an excellent agreement with data obtained numerically using Glauber dynamics, even at high temperature regime.

This technique can be used to include correlations at any depth, time differences, in a systematic approach to improve the prediction of dynamically evolving macroscopic quantities and inference of microscopic model parameters from data; this is expected as correlations at different times do carry valuable information for both tasks. Extending the calculation to devise improved approximations for symmetric densely connected spin systems is underway as well as the generalization of the method for sparse systems. Both research directions are highly promising and would impact on both theoretical research an practical applications.

VI CONCLUSION

————————————————

[1] G. Parisi, J. Phys. A **13**, 1101 (1980).
[2] M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[3] A. Crisanti and H. Sompolinsky, Phys. Rev. A **37** (1988).
[4] Y. Kabashima and D. Saad, Europhys. Lett. **44**, 668 (1998).
[5] Y. Roudi, J. A. Hertz, and E. Aurell, Front. Comput. Neurosci. **3:22** (2009).
[6] M. Opper and D. Saad, *Advanced mean field methods: Theory and practice* (the MIT press, 2001).
[7] H. Sompolinsky and A. Zippelius, Phys. Rev. B **25**, 6860 (1982).
[8] H. J. Sommers, Phys. Rev. Lett. **58**, 1268 (1987).
[9] G. Biroli, Jou. Phys. A **32**, 8365 (1999).
[10] M. Mezard and J. Sakellariou, Jour. Stat. Mech. **2011**, L07001 (2011).
[11] J. Sakellariou, Y. Roudi, M. Mezard, and J. Hertz, Phil. Mag. **92**, 272 (2012).
[12] J. Hertz and Y. Roudi, Phys. Rev. Lett. **106**, 048702 (2011).
[13] E. Schneidman, M. Berry, R. Segev, and W. Bialek, Nature **440**, 1007 (2006).
[14] H.-L. Zeng, E. Aurell, M. Alava, and H. Mahmoudi, Phys. Rev. E **83**, 041135 (2011).
[15] Y. Roudi, S. Nirenberg, and L. P. E., PLoS Comp. Biol. p. In press (2009), URL `preprintarXiv:0811.0903v1[q-bio.QM]`.
[16] M. Weigt, R. A. White, H. Szurmant, J. A. Hoch, and T. Hwa, PNAS **106**, 67 (2009).
[17] Y. Roudi and J. Hertz, J. Stat. Mech. p. P03031 (2011).
[18] E. Aurell and H. Mahmoudi, Phys. Rev. E **85**, 031119 (2012).
[19] S. Amari and K. Maginu, Neural Networks **1**, 63 (1988).
[20] M. Okada, Neural Networks **8**, 833 (1995).
[21] R. J. Glauber, Jour. of Math. Phys. **4**, 294 (1963).