Mean Field Theory for Nonequilibrium Network Reconstruction

Yasser Roudi[1] and John Hertz[1,2]
[1]NORDITA, 10691 Stockholm, Sweden
[2]The Niels Bohr Institute, 2100 Copenhagen, Denmark
(Received 29 September 2010; published 27 January 2011)

There has been recent progress on inferring the structure of interactions in complex networks when they are in stationary states satisfying detailed balance, but little has been done for nonequilibrium systems. Here we introduce an approach to this problem, considering, as an example, the question of recovering the interactions in an asymmetrically coupled, synchronously updated Sherrington-Kirkpatrick model. We derive an exact iterative inversion algorithm and develop efficient approximations based on dynamical mean-field and Thouless-Anderson-Palmer equations that express the interactions in terms of equal-time and one-time-step-delayed correlation functions.

DOI: 10.1103/PhysRevLett.106.048702 PACS numbers: 89.75.Hc, 02.50.Tt, 05.10.−a, 75.10.Nr

Introduction.—Finding the connectivity in complex networks is crucial for understanding how they operate. Gene and multielectrode microarrays have recently made the type of data required for this purpose available. What is needed now is appropriate theoretical tools for analyzing these data and extracting the connectivity.

In much recent work on this subject [1–3], the problem has been posed as that of inferring the parameters of a stationary Gibbs distribution modeling the system. While satisfied in many applications, the assumption of Gibbs equilibrium is unlikely to hold for many biological systems since they are usually driven by time-dependent external fields, their interactions may not satisfy detailed balance, or they may only be observed while the transients dominate the dynamics. Applying the equilibrium approach to such cases usually yields effective interactions that do not bear an obvious relationship to the real ones [3]. Kinetic and nonequilibrium models provide a much richer platform for studying such systems [4–6].

Whereas for equilibrium models the development of systematic mean-field inference methods [7] has led to great practical and conceptual advancements, a mean-field theory for nonequilibrium network reconstruction is still lacking. In this Letter, we show how a mean-field theory for inference can also be developed for a nonequilibrium system. We consider this problem for a particular simple nonequilibrium model: a kinetic Ising model with random asymmetric interactions (J_{ji} independent of J_{ij}), in an external field which may be time dependent. This is a discrete-time, synchronously updated model composed of N spins $s_i = \pm 1$ with transition probability

$$\Pr(\boldsymbol{s}(t+1)|\boldsymbol{s}(t)) = \prod_i \frac{\exp[s_i(t+1)\theta_i(t)]}{2\cosh(\theta_i(t))}, \qquad (1)$$

where $\theta_i(t) = h_i(t) + \sum_j J_{ij} s_j(t)$. The couplings J_{ij} are independent Gaussian variables with variance g^2/N. This model can be readily applied to time-binned neural data, where t labels the bins, and $s_i(t) = \pm 1$ represents a spike or no spike by neuron i in bin t [1]. The temperature has been set equal to 1, since any effects of changing the temperature can be realized by changing the coupling parameter g and the field strengths. Even for time-independent field and in a steady state, this system is not in a Gibbs equilibrium [8]. However, we show that, like its equilibrium counterpart, the nonequilibrium inverse problem for this model can be solved using a gradient descent method and also via systematic approximate inferences derived using dynamical versions of naive mean-field (nMF) and Thouless-Anderson-Palmer (TAP) equations. We show that for both the stationary and nonstationary systems these methods provide efficient reconstruction of interactions. We also analytically quantify their errors.

Exact, nMF and TAP learning.—Suppose we have observed R realizations of duration L time steps of the process in (1). We denote the observed state of the system at time t of realization r by $\boldsymbol{s}^r(t) = \{s_1^r(t), \cdots, s_N^r(t)\}$. To find the couplings and external fields, we maximize the likelihood of the observed states under the model (1). This can be done using an iterative algorithm, analogous to Boltzmann learning for the equilibrium model: starting from an initial set of couplings and fields, one adjusts them iteratively by steps of sizes $\delta h_i = \eta_h \frac{\partial \mathcal{L}}{\partial h_i}$ and $\delta J_{ij} = \eta_J \frac{\partial \mathcal{L}}{\partial J_{ij}}$, $\mathcal{L}$ being the log-likelihood. The learning steps thus are

$$\delta h_i(t) = \eta_h\{\langle s_i(t+1)\rangle_r - \langle \tanh[\theta_i(t))]\rangle_r]\}, \qquad (2a)$$

$$\delta J_{ij} = \eta_J\{\langle s_i(t+1)s_j(t)\rangle - \langle \tanh[\theta_i(t)]s_j(t)\rangle\}, \qquad (2b)$$

where η_h and η_J are learning rates. Here and in what follows $\langle\cdots\rangle_r$ and $\langle\cdots\rangle$ represent averaging over repeats, and both repeats and time, respectively. An overline, instead, will indicate averaging over the spins. One can think of Eq. (2b) as performing a logistic regression to explain one-step separated correlations. This is similar to what is

proposed in [9] as an approximation for inferring the connectivity in an equilibrium Ising model.

Since performing the steps in this algorithm does not require Monte Carlo runs, it is faster than the equilibrium Boltzmann learning. However, two factors still make this algorithm slow for large systems and/or data sets, warranting the development of fast approximations. First, (2) is still an iterative algorithm which could take a long time to converge if not provided with a good initial condition and learning rates. Second, at each step the averages on the right-hand side of (2) should be calculated from the data *de novo*, given the adjusted parameters.

Two fast approximations, nMF and TAP learning rules, are derived and studied below. To implement them in the stationary case, one first uses the data to calculate the one-step-delayed and equal-time correlations, $D_{ij} = \langle \delta s_i(t+1)\delta s_j(t)\rangle$ and $C_{ij} = \langle \delta s_i(t)\delta s_j(t)\rangle$, where $m_i = \langle s_i\rangle$ and $\delta s_i = s_i - m_i$. The approximations are

$$\mathsf{J}^{\mathrm{nMF/TAP}} = \mathsf{A}^{\mathrm{nMF/TAP}^{-1}}\mathsf{D}\mathsf{C}^{-1} \tag{3}$$

where $A_{ij}^{\mathrm{nMF}} = (1 - m_i^2)\delta_{ij}$, $A_{ij}^{\mathrm{TAP}} = A_{ij}^{\mathrm{nMF}}(1 - F_i)$, and F_i is the root of the cubic equation (6) below. In the nonstationary case too, similar learning rules can be derived as shown later in the Letter.

Derivation of nMF and TAP inversion.—For simplicity, we consider first the stationary case, for which the sequence index r is superfluous, as averaging over time and repeats would be equivalent. We start with the maximum likelihood conditions, i.e., $\delta h_i = \delta J_{ij} = 0$ in (2). Using the nMF equations $m_i = \tanh(h_i + \sum_j J_{ik}^{\mathrm{nMF}} m_k)$, and writing the s_i in (2) as $m_i + \delta s_i$, we expand the tanh in the δs_i. The first nonzero term gives

$$\langle \delta s_i(t+1)\delta s_j(t)\rangle = (1 - m_i^2)\sum_k J_{ik}^{\mathrm{nMF}}\langle \delta s_k(t)\delta s_j(t)\rangle, \tag{4}$$

which can be written as (3) for the nMF case.

To get the TAP inversion formula, we start instead by assuming that the m_i satisfy the TAP equations $m_i = \tanh[h_i + \sum_k J_{ik}^{\mathrm{TAP}} m_k - m_i \sum_k (J^{\mathrm{TAP}})_{ik}^2 (1 - m_k^2)]$, which take into account the Onsager reaction term. The TAP equations, although usually derived for the equilibrium (symmetric-J) SK model, also hold for the asynchronously updated, asymmetric-J model in a stationary state [10]. We have verified that they are also valid in our synchronously updated model [11]. We again write $s_i = m_i + \delta s_i$, expanding the tanh to third order in powers of $\sum_k J_{ik}^{\mathrm{TAP}}\delta s_k + m_i \sum_k (J^{\mathrm{TAP}})_{ik}^2 (1 - m_k^2)$. Keeping terms up to order g^3 leads to $\mathsf{D} = \mathsf{A}^{\mathrm{TAP}}\mathsf{J}^{\mathrm{TAP}}\mathsf{C}$, where

$$A_{ij}^{\mathrm{TAP}} = A_{ij}^{\mathrm{nMF}}\left[1 - (1 - m_i^2)\sum_l (J^{\mathrm{TAP}})_{il}^2 (1 - m_l^2)\right]. \tag{5}$$

These equations cannot be solved directly as in the nMF case because $\mathsf{A}^{\mathrm{TAP}}$ depends on $\mathsf{J}^{\mathrm{TAP}}$. However, one can derive a cubic equation for the quantities $F_i = (1 - m_i^2)\sum_l (J^{\mathrm{TAP}})_{il}^2 (1 - m_l^2)$:

$$F_i(1 - F_i)^2 = (1 - m_i^2)\sum_j (J^{\mathrm{nMF}})_{ij}^2 (1 - m_j^2). \tag{6}$$

This determines $A_{ij}^{\mathrm{TAP}} = A^{\mathrm{nMF}}(1 - F_i)$, yielding (3) for the TAP case. The relevant root of (6) is the smallest one (the one approaching zero as $g \to 0$). This root cannot exceed $1/3$, restricting this technique to weak couplings.

For both nMF and TAP reconstruction, the external fields h_i can also be found by solving the respective magnetization equations after the J_{ij} have been obtained, just as in the equilibrium problem [7].

Performance of the algorithms.—We have verified that the algorithm (2) recovers the couplings of an asymmetric SK model exactly in the limit of $L \to \infty$, for a wide range of coupling strengths g, external fields, and system sizes. The mean square error $\epsilon_{\mathrm{exact}}$ is in general proportional to $1/L$, and in the weak-coupling limit a quadratic expansion of log-likelihood yields

$$\epsilon_{\mathrm{exact}} = \overline{\delta J_{ij}^2} \equiv \overline{(J_{ij} - J_{ij}^{\mathrm{true}})^2} = \frac{1}{(1 - m_i^2)L}, \tag{7}$$

where $J_{ij}(J_{ij}^{\mathrm{true}})$ are the inferred (true) couplings.

We find that the nMF algorithm leads to an error ϵ_{MF} of the form $\epsilon_{\mathrm{exact}} + \epsilon_{\mathrm{nMF}}^{\infty}$, where $\epsilon_{\mathrm{nMF}}^{\infty}$ is independent of L and proportional to $1/N$. Thus, for data sets of length $L \ll L^* = 1/\epsilon_{\mathrm{nMF}}^{\infty} \propto N$, nMF does almost as well as the exact algorithm. Furthermore, the larger the network, the better nMF does. The errors for the exact and nMF algorithms vs L are shown in Fig. 1(a).

For weak coupling, we can calculate the asymptotic nMF error $\epsilon_{\mathrm{nMF}}^{\infty}$ analytically as follows. We present the zero-field case here for simplicity. We expand the tanh in the max-likelihood equation to third order, giving

$$D_{\mathrm{in}} = \sum_k J_{ik}\langle s_k s_n\rangle - \frac{1}{3}\sum_{klm} J_{ik}J_{il}J_{im}\langle s_k s_l s_m s_n\rangle + \cdots. \tag{8}$$

Correlations here are at equal times, except for D_{in}. The dominant contributions in the sum over k, l, m are those with $k = l$, $l = m$, and $m = k$. Multiplying on the right by $(C^{-1})_{nj}$, summing over n and using (3) for nMF, yields

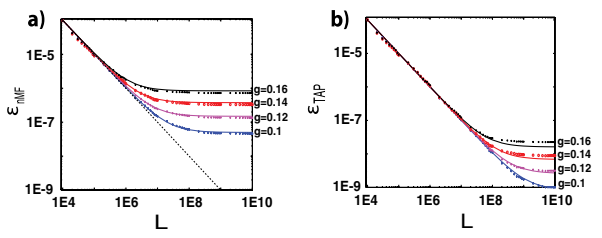

FIG. 1 (color online). Performance of the algorithms. Exact and nMF (a). and the TAP (b) errors are shown vs data length L for $g = 0.1$ (blue stars), 0.12 (magenta crosses), 0.14 (red circles), and 0.16 (black $\times$), all for $N = 20$ and zero external field. Theoretical predictions are the solid lines.

$$J_{ij}^{\mathrm{nMF}} = J_{ij} - \sum_k J_{ik}^2 J_{ij}, \tag{9}$$

with corrections of relative order $1/N$. Equation (9) also yields the TAP-approximation couplings found above, showing that the TAP reconstruction indeed corrects the leading MF errors. To leading order the sum on k is just g^2, and the asymptotic mean square MF error is

$$\epsilon_{\mathrm{nMF}}^{\infty} = \overline{(J_{ij} - J_{ij}^{\mathrm{nMF}})^2} = \frac{g^6}{N}. \tag{10}$$

The solid curves in Fig. 1(a) are $1/L + g^6/N$; the fit is evidently good. As shown in Fig. S1 [12], nMF exhibits a systematic error by underestimating the magnitude of the couplings. The factor $1 - F_i$ in TAP formula corrects for this to relative order g^2. Thus, when one is interested only in the presence or absence of connections, there would be little difference between nMF and TAP.

The error for the TAP reconstruction is much lower than that of the nMF one and reaches its minimum at much larger L: for $N = 20$ and the coupling strengths we studied, we had to go to $L \sim 10^9$ to see the error flatten [Fig. 1(b)]. To calculate the asymptotic reconstruction error for TAP, we expand the tanh to 5th order and proceed to evaluate the averages as we did for nMF. The nMF error terms analyzed above are compensated for by the TAP equations, as $N \to \infty$, leading to an asymptotic $\epsilon_{\mathrm{TAP}}^{\infty} = 4g^{10}/N$. For $N \gg 1/g^2$ this is the leading term in the asymptotic TAP error. Outside this regime, a finite-size effect should also be taken into account, because for the TAP correction, the term in (8) with $k = l = m$ has been counted 3 times in obtaining (9) instead of once. The resulting error is $(2/3)^2 \overline{J_{ij}^6} = (20g^6)/(3N^3)$ and should be added to the $4g^{10}/N$ term.

Nonstationary case.—The magnetizations, $m_i(t) = \langle s_i^r(t) \rangle_r$, are now time dependent and, for TAP, solve [11]

$$m_i(t+1) = \tanh\Big[h_i(t) + \sum_j J_{ij}^{\mathrm{TAP}} m_j(t) - m_i(t+1) \sum_j (J^{\mathrm{TAP}})_{ij}^2 (1 - m_j^2(t)) \Big]. \tag{11}$$

For nMF the final term inside the brackets is not present.

For inversion, we start by defining time-dependent correlation matrices $D_{ij}(t) \equiv \langle \delta s_i^r(t+1) \delta s_j^r(t) \rangle_r$ and $C_{ij}(t) \equiv \langle \delta s_i^r(t) \delta s_j^r(t) \rangle_r$. For nMF, using the same procedure that lead to (4), we find

$$\langle D_{ij}(t) \rangle_t = \sum_k J_{ik}^{\mathrm{nMF}} \langle (1 - m_i^2(t+1)) C_{kj}(t) \rangle_t. \tag{12}$$

One can still solve for J by simple matrix algebra:

$$J_{ij}^{\mathrm{nMF}} = \sum_k \langle D_{ik}(t) \rangle_t [(\mathsf{B}^{(i)})^{-1}]_{kj}, \tag{13}$$

where $B_{kj}^{(i)} = \langle (1 - m_i^2(t+1)) C_{kj}(t) \rangle_t$. The problem is more complex than the stationary one only because one has to invert a different matrix $\mathsf{B}^{(i)}$ for each i.

For TAP, analogously to the stationary case, the $\mathsf{B}^{(i)}$ acquire an extra factor inside the time average:

$$B_{kj}^{(i)} = \langle (1 - m_i^2(t+1))(1 - F_i(t)) C_{kj}(t) \rangle_t, \tag{14a}$$

$$F_i(t) = (1 - m_i^2(t+1)) \sum_l (J^{\mathrm{TAP}})_{il}^2 (1 - m_l^2(t)). \tag{14b}$$

Exact TAP inversion requires iterative solution of (13), with J_{ij}^{TAP} instead of J_{ij}^{nMF}, together with (14). We have found, however, that effective reconstruction is still possible under the simplifying approximation that $F_i(t)$ in Eq. (14a) can be represented by its temporal mean. In this case, $F_i \equiv \langle F_i(t) \rangle_t$ solves the cubic equation

$$F_i(1 - F_i)^2 = \sum_j (J^{\mathrm{nMF}})_{ij}^2 \langle (1 - m_i^2(t+1))(1 - m_j^2(t)) \rangle_t.$$

Solving it and using it in Eq. (14a), one can calculate $J_{ij}^{\mathrm{TAP}} = J_{ij}^{\mathrm{nMF}}/(1 - F_i)$. Similar to the stationary case, after inferring the couplings, one can use the forward dynamical nMF and TAP equations (11) to infer the time-varying external field. The result of reconstructing the couplings of a network driven by a common sinusoidal external field to all spins is shown in Fig. 2. Figure 2(a) shows how well the couplings are inferred by nonstationary nMF using $L = 10^5$ and $R = 100$. Nonstationary TAP couplings (not shown) have a lower mean squared error: 6.7×10^{-7} versus 10^{-6} for nMF. In Fig. 2(b), we also plot the couplings inferred using stationary nMF inversion for each of the 100 repeats and averaging over them. Not surprisingly, the stationary nMF performs poorly on this nonstationary data. Importantly, there is a systematic overestimation of the couplings in this case, because the stationary method accounts for correlations induced by the common, time-varying external field through adjusting the couplings. Correspondingly, if one uses the couplings inferred by stationary nMF to infer $h_i(t)$, the amplitude of this field is underestimated, while the use of nonstationary nMF couplings yields a very good reconstruction of $h_i(t)$; see Fig. 2(c).

Discussion.—We have shown how to infer interactions in a simple but nontrivial nonequilibrium system: a kinetic Ising model with random and potentially asymmetric interactions. The model is the maximum-entropy model for each time step, given mean magnetizations and one-step separated correlations. We have described both an exact iterative algorithm and two approximate ones, based on dynamical nMF and TAP equations, which are correct up to corrections of order $1/N$. We calculated analytically the errors of these approximations for weak coupling. The method shows particular promise when applied to nonstationary states, where it separates true interactions from the

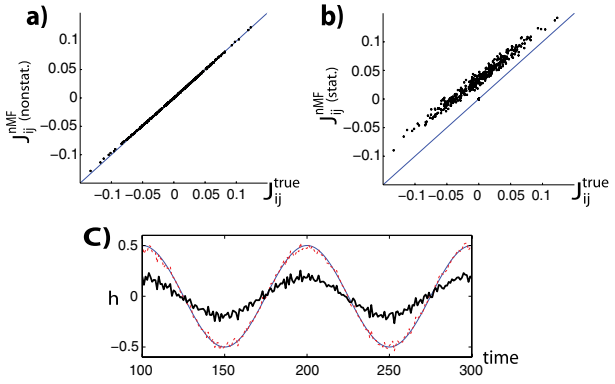

FIG. 2 (color online). Inference in the nonstationary case. (a) Couplings of a network of $N = 20$ driven by a sinusoidal external field inferred using the nonstationary nMF, and (b) the stationary nMF. (c) Two periods of the external field (thin blue full curve) and its reconstruction using the nonstationary nMF couplings (red dashed curve) and stationary nMF (thick black full curve).

apparent ones found by applying a stationary theory to a nonstationary state.

A kinetic Ising model will show an intrinsic error when applied to data from a different kind of system. However, even when applied to data from a realistic network, the simple approximate learning rules developed here identify the connections much better than their equilibrium counterparts. Figure 3 shows the distribution of couplings found by applying the nonequilibrium TAP learning to data from a simulated model cortical column with inhibitory and excitatory neurons [13]. The connections in the model were dilute with 10% probability of connection. When there is no synapse from neuron j to i, the inferred J_{ij} follows a zero mean distribution, while if there is an excitatory (inhibitory) synapse, it follows a positive (negative) mean distribution, well separated from the first one. One can thus easily use the distribution of inferred couplings to infer the presence, absence, and sign of the

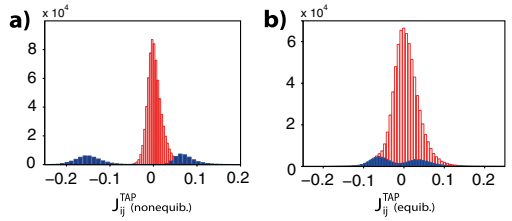

FIG. 3 (color online). Finding connections in a cortical network model. (a) The histogram of the couplings inferred using the stationary nonequilibrium TAP for pairs of neurons that were connected (blue full bars), and those that were not (red empty bars). The separation between the histograms shows that one can use the TAP approximation to separate connected and disconnected pairs. (b) same as (a) for equilibrium TAP.

connections; see [12]. However, this is not the case when equilibrium TAP learning is used. When using a model like (1) to infer connectivity in a system with a different dynamics, or when faced with data limitation, including prior knowledge about the network could be very beneficial. In particular, taking into account sparsity of the connections via a l-1 regularizer added to the likelihood has been shown to be very useful [9]. It is easy to show that adding an l-1 regularizer to the likelihood of the data under (1) would modify (3) by adding a term proportional to $\mathsf{A}^{\mathrm{nMF/TAP}^{-1}}\mathrm{sgn}(\mathsf{J})\mathsf{C}^{-1}$ to the right-hand side. How this improves inferring connections in biological networks will be discussed elsewhere.

A simple extension of (1) is its continuous time version. As shown in [14], for this model, too, a mean-field theory can be developed using the approach presented here. In other recent kinetic approaches to problems like this, the equilibrium maximum-entropy approach [1] is extended to include non-equal-time correlations [5] and an approximate scheme for fitting an integrate-and-fire network to data was developed in [4]. There has also been work [6], closely connected to (1), in which $s_i(t+1)$ depends on linear combinations of $h(t')$ and $s(t')$, for $t' \leq t$. Given the advantage of these nonequilibrium models over the equilibrium ones for describing spike train statistics, a mean-field theory for inferring their parameters would be of great benefit. For such models, we expect one can use the techniques in [10] or [8,15] to derive dynamical nMF and TAP equations. Employing the approach developed here one can then build approximate mean-field inversion techniques based on these equations.

[1] E. Schneidman *et al.*, Nature (London) **440**, 1007 (2006); J. Shlens *et al.*, J. Neurosci. **26**, 8254 (2006).
[2] M. Weigt *et al.*, Proc. Natl. Acad. Sci. U.S.A. **106**, 67 (2009); T. Lezon *et al.*, *ibid.* **103**, 19 033 (2006).
[3] Y. Roudi, J. Tyrcha, and J. A. Hertz, Phys. Rev. E **79**, 051915 (2009).
[4] S. Cocco *et al.*, Proc. Natl. Acad. Sci. U.S.A. **106**, 14 058 (2009).
[5] O. Marre *et al.*, Phys. Rev. Lett. **102**, 138101 (2009).
[6] W. Truccolo *et al.*, J. Neurophysiol. **93**, 1074 (2005).
[7] H. J. Kappen and F. B. Rodriguez, Neural Comput. **10**, 1137 (1998); T. Tanaka, Phys. Rev. E **58**, 2302 (1998).
[8] A. C. C. Coolen, arXiv:cond-mat, 0006011.
[9] P. Ravikumar *et al.*, Ann. Stat. **38**, 1287 (2010).
[10] H. J. Kappen and J. J. Spanjers, Phys. Rev. E **61**, 5658 (2000).
[11] Y. Roudi and J. Hertz (to be published).
[12] See supplemental material at http://link.aps.org/supplemental/10.1103/PhysRevLett.106.048702 for details about the computer simulations and Fig. S1.
[13] J. A. Hertz, Neural Comput., **22**, 427 (2010).
[14] H.-L Zeng *et al.*, arXiv:1011.6216v1.
[15] G. Biroli, J. Phys. A **32**, 8365 (1999).