Inferring hidden states in a random kinetic Ising model: replica analysis

Ludovica Bachschmid Romano and Manfred Opper

Department of Artificial Intelligence, Technische Universität Berlin, Marchstraße 23, Berlin 10587, Germany

E-mail: `ludovica.bachschmidromano@tu-berlin.de` and `manfred.opper@tu-berlin.de`

Abstract. We consider the problem of predicting the spin states in a kinetic Ising model when spin trajectories are observed for only a finite fraction of sites. In a Bayesian setting, where the probabilistic model of the spin dynamics is assumed to be known, the optimal prediction can be computed from the conditional (posterior) distribution of unobserved spins given the observed ones. Using the replica method, we compute the error of the Bayes optimal predictor for parallel discrete time dynamics in a fully connected spin system with non symmetric random couplings. The results, exact in the thermodynamic limit, agree very well with simulations of finite spin systems.

1. Introduction

The problem of statistical inference in kinetic Ising models has recently attracted considerable interest in the statistical physics community, see e.g. [1–5]. These systems can be viewed as simple models of networks of spiking neurons and provide a prototype model for which a reconstruction of the network from dynamical data can be studied. Based on a temporal sequence of observed spin variables, a major goal is to estimate the couplings between sites. This task gets more complicated when at some sites the spin trajectories are not observed. Besides the problem of inferring the couplings it is then also interesting to predict the states of the non observed spins when the couplings are known. In fact, an iterative solution to the maximum likelihood problem for estimating the couplings is the *Expectation Maximization* (EM) algorithm [6] which would iterate between estimating hidden spin states (given the last estimate of the couplings) and reestimating the couplings. Unfortunately, exact inference of hidden states is not tractable for large networks, but algorithms which are based on statistical physics approximations have recently been discussed [7,8]. Hence, it will be interesting and important to study a scenario for which the theoretically optimal performance for predicting hidden spins can be computed exactly. In this paper, we will show that such a solution can be found in the thermodynamic limit of an infinitely large network when the couplings are random. Our approach will be based on the replica method of disordered systems which enables us to compute quenched averages over the random

couplings for thermodynamic quantities of the model. These thermodynamic quantities are themselves functions of posterior averages (e.g. local magnetizations) of the hidden spins. The replica approach has been successfully applied in the past to a large variety of statistical learning problems for *static* network models (for a summary see [9–11]). We will restrict ourselves to a model where the couplings are mutually independent random variables, i.e. where no symmetry between in-and outgoing connections are assumed. For such type of models (without the observations) various exact solutions for the non equilibrium dynamics have been computed, see e.g. [1, 5] and [12, 13] for soft spin models. From the point of view of equilibrium statistical physics the case of symmetric couplings might be interesting. Such a spin model would obey detailed balance and allow for a stationary Gibbs distribution. Unfortunately, for the Ising case, the exact computation of time dependent correlation functions which are necessary for our analysis seems not possible. On the other hand, from a point of view of neural modeling, the assumption of symmetric couplings is not realistic [1, 14], as synaptic connections in biological networks are known to be strongly asymmetric. Hence, we believe that our restriction to asymmetric couplings is justified both from a modeling and a computational perspective.

2. The model and Bayes optimal inference

We will consider a model with N Ising spins which are divided into two groups: a group of spins $s_i(t)$ at sites i=1,..., $N_{\text{obs}} = \lambda N$ which are observed during a time interval of T time steps, and a group of hidden, i.e. unobserved spins, denoted by $\sigma_a(t)$ at sites $a = 1, \ldots, N_{\text{hid}} = (1 - \lambda)N$. We assume parallel Markovian dynamics for the entire spin system, which is governed by the transition probability

$$P[\{s, \sigma\}(t + 1)|\{s, \sigma\}(t)] = \prod_i \frac{e^{s_i(t+1)g_i(t)}}{2 \cosh[g_i(t)]} \prod_a \frac{e^{\sigma_a(t+1)g_a(t)}}{2 \cosh[g_a(t)]}, \tag{1}$$

where the fields are defined as

$$g_i(t) = \sum_j J_{ij}s_j(t) + \sum_b J_{ib}\sigma_b(t), \qquad g_a(t) = \sum_j J_{aj}s_j(t) + \sum_b J_{ab}\sigma_b(t), \tag{2}$$

in terms of the couplings J and $\{s, \sigma\}$ denotes all the possible spin vector configurations; when the time index is not specified we are considering the whole time series, $t = 0...T$. The total probability for a spin trajectory is given by

$$P(\{s, \sigma\}) = \frac{1}{2^N} \prod_{t=0}^{T-1} P[\{s, \sigma\}(t + 1)|\{s, \sigma\}(t)], \tag{3}$$

where we have considered completely random initial condition $P_0[\{s, \sigma\}(0)] = 1/2^N$.

To make predictions on the unobserved spins $\sigma_a(t)$, we assume that the model given by the couplings J is perfectly known and the posterior, i.e. conditional probability of

the hidden spins defined by

$$P(\{\sigma\}|\{s\}) = \frac{P(\{s,\sigma\})}{P(\{s\})} , \qquad (4)$$

gives the complete information for an optimal inference of hidden spins. Based on this probabilistic information, the best possible prediction $\sigma_a^{opt}(t)$ for the hidden spin at site a and at time t is computed by

$$\sigma_a^{opt}(t) = \text{sign}[m_a(t)] , \qquad (5)$$

where the local magnetization is defined as the posterior expectation

$$m_a(t) = \sum_{\{\sigma\}} \sigma_a(t) P(\{\sigma\}|\{s\}) . \qquad (6)$$

Note that this does not correspond to the most likely spin *configuration* $\{\sigma\}$, because we have averaged out the configurations of spins $\sigma_b(t')$ for $b \neq a$ and $t' \neq t$.

Given a true 'teacher' sequence $\{\sigma^*\}$ of unobserved spins, we are interested in the total quality of the Bayes optimal prediction, i.e. in the expected probability of *wrongly* predicting a spin at site a and time t, given by the Bayes error

$$\varepsilon = \sum_{\{s,\sigma^*\}} P(\{s,\sigma^*\})\Theta(-\sigma_a^*(t)m_a(t)) = \sum_{\{s\}} P(\{s\}) \sum_{\{\sigma^*\}} P(\{\sigma^*\}|\{s\})\Theta(-\sigma_a^*(t)m_a(t)) , \qquad (7)$$

where the step function $\Theta(x) = 1$ for $x > 0$ and 0 else. In the next section we will use the replica method to compute the error in the thermodynamic limit $N \to \infty$, when the couplings J are assumed to be mutually independent Gaussian random variables, with zero mean and variance of the order $1/N$.

3. Replica analysis

The posterior statistics of the hidden spins can be obtained from the following partition function

$$P(\{s\}) = \frac{1}{2^N} \sum_{\{\sigma\}} \prod_t P[\{s,\sigma\}(t+1)|\{s,\sigma\}(t)], \qquad (8)$$

which equals the total probability of the observed spin configurations and is also the normalizer of the posterior probability. Typical performance in the thermodynamic limit for random couplings are then computed from the quenched average of the free energy $F = -\langle \ln P(\{s\})\rangle_{J,s}$, where the average is taken over the the couplings J and over the observed spin configurations with their weights $P(\{s\})$. Hence, the averaged free energy is given by

$$F = -\sum_{\{s\}} \langle P(\{s\}) \log P(\{s\})\rangle_J . \qquad (9)$$

This average can be computed by the replica trick [9–11] in the following way:

$$F = -\lim_{n \to 1} \frac{d}{dn} \log \sum_{\{s\}} \langle P^n(\{s\}) \rangle. \tag{10}$$

For integer n, we have

$$\sum_{\{s\}} \langle P^n(\{s\}) \rangle_J = \frac{1}{2^{nN}} \sum_{\{s\}} \sum_{\{\sigma^{(1)}\}} \cdots \sum_{\{\sigma^{(n)}\}} \left\langle \left[\prod_{\alpha=1}^{n} \exp \left\{ \sum_{it} s_i(t+1) g_i^{\alpha}(t) \right. \right. \right.$$

$$\left. \left. \left. + \sum_{at} \sigma_a^{\alpha}(t+1) g_a^{\alpha}(t) - \sum_{it} \log 2 \cosh[g_i^{\alpha}(t)] - \sum_{at} \log 2 \cosh[g_a^{\alpha}(t)] \right\} \right] \right\rangle_J, \tag{11}$$

with

$$g_i^{\alpha}(t) = \sum_j J_{ij} s_j(t) + \sum_b J_{ib} \sigma_b^{\alpha}(t), \qquad g_a^{\alpha}(t) = \sum_j J_{aj} s_j(t) + \sum_b J_{ab} \sigma_b^{\alpha}(t). \tag{12}$$

To perform the average over the couplings J_{ij}, J_{ib}, J_{aj} and J_{ab}, which are assumed to be mutually independent Gaussian random variables with zero mean and variance k^2/N, we note that the fields $g_i^{\alpha}(t)$ and $g_a^{\alpha}(t)$ are also Gaussian, which are independent for different sites i and a, but will be dependent for different replica index α and β and also possibly for different times. This yields

$$\left\langle g_i^{\alpha}(t) g_i^{\beta}(t') \right\rangle = \left\langle g_a^{\alpha}(t) g_a^{\beta}(t') \right\rangle = k^2 \left(\lambda S(t,t') + (1-\lambda) Q^{\alpha\beta}(t,t') \right),$$

$$\langle g_i^{\alpha}(t) g_i^{\alpha}(t') \rangle = \langle g_a^{\alpha}(t) g_a^{\alpha}(t') \rangle = k^2 \left(\lambda S(t,t') + (1-\lambda) C^{\alpha}(t,t') \right), \tag{13}$$

where we have defined the following order parameters

$$C^{\alpha}(t,t') = \frac{1}{N_{\text{hid}}} \sum_a \sigma_a^{\alpha}(t) \sigma_a^{\alpha}(t') \text{ for } t < t', \tag{14}$$

$$Q^{\alpha\beta}(t,t') = \frac{1}{N_{\text{hid}}} \sum_a \sigma_a^{\alpha}(t) \sigma_a^{\beta}(t') \text{ for } \alpha < \beta,$$

$$S(t,t') = \frac{1}{N_{\text{obs}}} \sum_i s_i(t) s_i(t') \text{ for } t < t'. \tag{15}$$

Introducing these definitions within δ functions and expressing the δ functions using conjugate (hatted) integration parameters, we get the following expression:

$$
\sum_{\{s\}} \langle P^n(\mathbf{s}) \rangle_J = \frac{c}{2^{nN}} \int \prod_{t,t'} \prod_{\alpha<\beta} \left(dQ^{\alpha\beta}(t,t') d\hat{Q}^{\alpha\beta}(t,t') \right)
$$

$$
\prod_{t<t'} \prod_{\alpha} \left(dC^{\alpha}(t,t') d\hat{C}^{\alpha}(t,t') \right) \prod_{t<t'} \left(dS(t,t') d\hat{S}(t,t') \right)
$$

$$
\exp \left(iN_{\text{hid}} \sum_{\alpha} \sum_{t<t'} C^{\alpha}(t,t')\hat{C}^{\alpha}(t,t') + iN_{\text{hid}} \sum_{\alpha<\beta} \sum_{tt'} Q^{\alpha\beta}(t,t')\hat{Q}^{\alpha\beta}(t,t') \right.
$$

$$
+ iN_{\text{obs}} \sum_{t<t'} S(t,t')\hat{S}(t,t') + N_{\text{obs}} \log \mathcal{E}_{\text{obs}}(C,Q)
$$

$$
\left. + N_{\text{hid}} \log \mathcal{E}_{\text{hid}}(C,\hat{C},Q,\hat{Q}) \right),
$$

(16)

where c is a trivial constant non depending on N,

$$
\mathcal{E}_{\text{obs}}(C,Q) = \sum_{\{s\}} \left\langle \exp \left(\sum_{t\alpha} s(t+1)g^{\alpha}(t) - \sum_{t\alpha} \log 2\cosh[g^{\alpha}(t)] \right. \right.
$$

$$
\left. \left. -i \sum_{t<t'} \hat{S}(t,t')s(t)s(t') \right) \right\rangle_g
$$

(17)

$$
\mathcal{E}_{\text{hid}}(C,\hat{C},Q,\hat{Q}) = \sum_{\{\sigma^{(1)}\}} ... \sum_{\{\sigma^{(n)}\}} \left\langle \exp \left(\sum_{t} \sigma^{\alpha}(t+1)g^{\alpha}(t) - \sum_{t\alpha} \log 2\cosh[g^{\alpha}(t)] \right. \right.
$$

$$
\left. \left. -i \sum_{\alpha} \sum_{t<t'} \hat{C}^{\alpha}(t,t')\sigma^{\alpha}(t)\sigma^{\alpha}(t') - i \sum_{\alpha<\beta} \sum_{tt'} \hat{Q}^{\alpha\beta}(t,t')\sigma^{\alpha}(t)\sigma^{\beta}(t') \right) \right\rangle_g,
$$

and the average is over the Gaussian fields with statistics given by (13). In the limit $N \to \infty$, keeping the ratio $\lambda = N_{\text{obs}}/N$ fixed, the integrals over the order parameters can be performed using the saddle point method, where we assume replica symmetry, i.e. $C^{\alpha}(t,t') = C(t,t') \quad \forall \alpha, t < t'$ and, $Q^{\alpha\beta}(t,t') = Q(t,t') \quad \forall \alpha < \beta, t, t'$. We get

$$
\lim_{N\to\infty} \frac{1}{N} \log \sum_{\{s\}} \langle P^n(\{s\}) \rangle = \text{Extr } f_n(C,S,\dots),
$$

where we have to take the extremum with respect to the order parameters in the expression

$$f_n(C, S, \ldots) = i(1 - \lambda)n \sum_{t<t'} C(t,t')\hat{C}(t,t') + i(1-\lambda)\frac{(n^2 - n)}{2}\sum_{tt'} Q(t,t')\hat{Q}(t,t')$$

$$+ i\lambda \sum_{t<t'} S(t,t')\hat{S}(t,t') + \lambda \log \sum_{\{s\}} \left\langle\!\!\left\langle \left\langle \prod_t V(t) \right\rangle_\zeta^n e^{\sum_t s(t)\nu(t)} \right\rangle_{\psi,\nu}\right.$$ (18)

$$+ (1 - \lambda)\log \left\langle\!\!\left\langle \left\langle \Gamma_0 \prod_t Z(t) \right\rangle_{\xi,\zeta}^n \right\rangle_{\phi,\psi} - n\log 2,$$

where we have introduced

$$V(t) = \frac{e^{s(t+1)(\psi(t)+\zeta(t))}}{2\cosh(\psi(t)+\zeta(t))}, \qquad Z(t) = \frac{\cosh[\psi(t)+\zeta(t)+\phi(t+1)+\xi(t+1)]}{\cosh(\psi(t)+\zeta(t))}, \quad (19)$$

in terms of Gaussian independent random fields $\psi(t)$, $\zeta(t)$, $\nu(t)$, $\xi(t)$ and $\phi(t)$, with zero mean and covariances given by the following set of equations:

$$\langle\psi(t)\psi(t')\rangle = k^2\left(\lambda S(t,t') + (1-\lambda)Q(t,t')\right), \quad (20)$$

$$\langle\zeta(t)\zeta(t')\rangle = k^2(1-\lambda)\left(C(t,t') - Q(t,t')\right), \quad (21)$$

$$\langle\nu(t)\nu(t')\rangle = -i\hat{S}(t,t'), \quad (22)$$

$$\langle\xi(t)\xi(t')\rangle = -i(\hat{C}(t,t') - \hat{Q}(t,t')), \quad (23)$$

$$\langle\phi(t)\phi(t')\rangle = -i\hat{Q}(t,t'), \quad (24)$$

for $t' \neq t$ and

$$\langle\psi(t)\psi(t)\rangle = k^2\left(\lambda + (1-\lambda)Q(t,t)\right), \qquad \langle\zeta(t)\zeta(t)\rangle = k^2(1-\lambda)\left(1 - Q(t,t)\right), \quad (25)$$

$$\langle\nu(t)\nu(t)\rangle = 0, \qquad\qquad\qquad \langle\xi(t)\xi(t)\rangle = i\hat{Q}(t,t), \quad (26)$$

$$\langle\phi(t)\phi(t)\rangle = -i\hat{Q}(t,t), \quad (27)$$

for $t' = t$. The term Γ_0 contains the initial condition for the fields ϕ, ξ (Appendix B). The 3 sets of Gaussian variables in (23,24,26,27) have been introduced to linearize the quadratic forms in equation (17). We can now perform the continuation to noninteger n and obtain the free energy per spin $\lim_{N\to\infty} F/N$ as the stationary value of

$$f(C, S, \ldots) = -i(1-\lambda)\sum_{t<t'} C(t,t')\hat{C}(t,t') - i\frac{(1-\lambda)}{2}\sum_{tt'} Q(t,t')\hat{Q}(t,t')$$

$$- \lambda\frac{\sum_{\{s\}} \left\langle \langle\prod_t V(t)\rangle_\zeta \log\langle\prod_t V(t)\rangle_\zeta \, e^{-i\sum_t s(t)\nu(t)}\right\rangle_{\psi,\nu}}{\sum_{\{s\}} \left\langle\langle\prod_t V(t)\rangle_\zeta \, e^{\sum_t s(t)\nu(t)}\right\rangle_{\psi,\nu}}$$ (28)

$$- (1-\lambda)\frac{\left\langle\langle\Gamma_0\prod_t Z(t)\rangle_{\xi,\zeta} \log\langle\Gamma_0\prod_t Z(t)\rangle_{\xi,\zeta}\right\rangle_{\phi,\psi}}{\langle\Gamma_0\prod_t Z(t)\rangle_{\xi,\zeta,\phi,\psi}}.$$

From equation (28) we can compute the self-averageing values of the order parameters and their conjugates. Previous studies [1,5,15] of spin models with asymmetric couplings have shown that spin correlations $S(t,t')$ decay after one time step. Hence, we expect that also for our model the other two time order parameters are zero for $t \neq t'$. Indeed, we can show (for an example, see Appendix A) that the results

$$C(t,t') = Q(t,t') = \hat{C}(t,t') = \hat{Q}(t,t') = \hat{S}(t,t') = 0$$

are self-consistent solutions of the order parameter equations for $t' \neq t$ and this solution is also supported by simulations. In this case, only the terms with $t' = t$ give non-zero contribution in equetion (16) and the free energy of the system simplifies to

$$
\begin{aligned}
f(Q,\hat{Q}) = {}& -\frac{i}{2}(1-\lambda)\sum_{t=0}^{T} Q(t)\hat{Q}(t) - \frac{i}{2}(1-\lambda)\sum_{t=0}^{T}\hat{Q}(t) \\
& - \lambda \sum_{t=0}^{T-1}\sum_{\{s\}(t+1)} \left\langle \langle V(t)\rangle_{\zeta_t} \log \langle V(t)\rangle_{\zeta_t}\right\rangle_{\psi_t} \\
& - (1-\lambda)\sum_{t=0}^{T-1} \frac{\left\langle \left\langle \tilde{Z}(t)\right\rangle_{\zeta_t} \log \left\langle \tilde{Z}(t)\right\rangle_{\zeta_t}\right\rangle_{\phi_{t+1},\psi_t}}{\left\langle \tilde{Z}(t)\right\rangle_{\zeta_t,\phi_{t+1},\psi_t}} - (1-\lambda)\tilde{\Gamma}_0,
\end{aligned}
\tag{29}
$$

where

$$\tilde{Z}(t) = \frac{\cosh\left[\psi(t) + \zeta(t) + \phi(t+1)\right]}{\cosh\left(\psi(t) + \zeta(t)\right)},$$

$Q(t) \equiv Q(t,t)$ and the initial condition $\tilde{\Gamma}_0$ is given in Appendix B. The order parameter $Q(t)$ gives the typical overlap of two independent spin configurations at time t drawn at random from the posterior distribution. By symmetry, it also describes the expected overlap of the hidden spins drawn from the posterior with the true 'teacher' spins of the model from which the observation data were generated. Hence the limit $Q(t) = 0$ describes a situations where the posterior gives no information on the hidden spins. On the other hand, $Q(t) = 1$, means that we can predict the hidden spins perfectly. We obtain the following equations for the order parameters:

$$
Q(t) = \frac{1}{\left\langle \tilde{Z}(t-1)\right\rangle_{\zeta_{t-1},\phi_t,\psi_{t-1}}} \left\langle \frac{\left\langle \tanh \tilde{A}(t-1)\tilde{Z}(t-1)\right\rangle_{\zeta_{t-1}}^2}{\left\langle \tilde{Z}(t-1)\right\rangle_{\zeta_{t-1}}}\right\rangle_{\phi_t,\psi_{t-1}}, \qquad t = 1...T
\tag{30}
$$

$$\hat{Q}(t) = \frac{ik^2(1-\lambda)}{\left\langle \tilde{Z}(t) \right\rangle_{\zeta_t,\phi_{t+1},\psi_t}} \left\langle \frac{\left\langle [\tanh \tilde{A}(t) - \tanh \tilde{B}(t)]\tilde{Z}(t) \right\rangle^2_{\zeta_t}}{\left\langle \tilde{Z}(t) \right\rangle_{\zeta_t}} \right\rangle_{\phi_{t+1},\psi_t}$$

$$+ ik^2\lambda \sum_{\{s\}(t+1)} \left\langle \frac{\left\langle [s(t+1) - \tanh \tilde{B}(t)]V(t) \right\rangle^2_{\zeta_t}}{\left\langle V(t) \right\rangle_{\zeta_t}} \right\rangle_{\psi_t}, \qquad t = 0...T-1 \tag{31}$$

where

$$\tilde{A}(t) = \psi(t) + \zeta(t) + \phi(t+1), \qquad\qquad \tilde{B}(t) = \psi(t) + \zeta(t).$$

The equations for the initial and final conditions, $Q(0)$ and $\hat{Q}(T)$, are given in in Appendix B.

4. Distribution of local magnetization

It is easy to extend the replica approach to the computation of other thermodynamic quantities such as functions of the local magnetizations. We find that, in the thermodynamic limit, hidden spins can be viewed as mutually independent random variables which are coupled to random fields. The spins have local magnetizations

$$m(t|\psi,\phi) = \frac{\left\langle \tanh A(t-1)\tilde{Z}(t-1) \right\rangle_{\zeta_{t-1}}}{\left\langle \tilde{Z}(t-1) \right\rangle_{\zeta_{t-1}}}, \tag{32}$$

where the 'inner' averages over ζ reflect the averaging out of the other spins. The magnetizations depend on the random fields $\psi(t-1), \phi(t)$. These Gaussian fields reflect the disorder originating from the random couplings. In computing expectations they get an extra statistical weight given by

$$w(\psi,\phi) = \frac{\left\langle \tilde{Z}(t-1) \right\rangle_{\zeta_{t-1}}}{\left\langle \tilde{Z}(t-1) \right\rangle_{\zeta_{t-1},\psi_{t-1},\phi_t}} \tag{33}$$

in the 'outer' average. Hence, the distribution of local magnetizations at an arbitrary site and at time t is given by

$$p_t(m) = \left\langle w(\psi,\phi)\delta(m - m(t|\psi,\phi)) \right\rangle_{\psi_{t-1},\phi_t}, \tag{34}$$

from which the overlap Q is recovered as $Q(t) = \int_{-1}^{1} p_t(m)m^2 dm$. Finally, to get the Bayes error we note that the (posterior) probability of a spin σ equals $\frac{1}{2}(m\sigma+1)$. Hence eq (7) is translated into

$$\varepsilon = \frac{1}{2} \sum_{\sigma=\pm 1} \int_{-1}^{1} p_t(m)(m\sigma + 1)\Theta(-\sigma m)dm = \frac{1}{2}\left(1 - \int_{-1}^{1} p_t(m)\,|m|\,dm\right), \tag{35}$$

where the last equality follows easily from the fact that $p_t(m) = p_t(-m)$.

5. Results

We have solved the order parameter equations (30) and (31) by iterating equations (25, 27, 30, 31) for different values of the load parameter λ and coupling strength k (for an example see figure 1). We start the recursion from the prior initial condition $Q(0) = 0$ and then iterate the equations forward and backward, updating the boundary conditions at each iteration according to equation (42). The overlap is smallest at the boundary $t = 0$ and $t = T$, because there the information flow is only from one direction and is also expected to decay over the time T.

When the length T of the spin trajectories grows, the order parameters $Q(t)$ and $\hat{Q}(t)$ for times t away from the boundaries, i.e. $0 \ll t \ll T$, converge to stationary values Q_{stat} and $\hat{Q}_{stat}$. These can be directly computed from eqs. (30-31) by setting $Q(t) = Q(t-1) = Q_{stat}$ and $\hat{Q}(t) = \hat{Q}(t+1) = \hat{Q}_{stat}$. For given stationary order parameters we have then computed the distribution of local magnetizations and the Bayes error. The Bayes error ε is shown in figure (2) as a function of the load factor λ. In the limit of no observations, $\lambda = 0$, the prediction on the the state of hidden spins is completely random and the error has the trivial value $\varepsilon = 0.5$. The error rapidly decreases as λ gets larger, but remains nonzero for $\lambda = 1$, indicating the presence of a residual error in almost fully observed systems due to the stochasticity of the Markov process. Since the couplings are responsible for the propagation of information between spin sites, the Bayes error decreases as the coupling strength increases; in particular we find that $\varepsilon \to 0$ for $k \to \infty$. This behaviour is illustrated in figure (3), where the distribution $p(m)$ of the local magnetization (eq. 34) is shown. For small k the distribution is close to a Gaussian centered at zero, with vanishing variance as $k \to 0$, meaning (see eq. 35) that nontrivial prediction on the magnetization can be made. As k grows larger the distribution broadens and above a critical value the curve becomes bimodal. For large k, the distribution $p(m)$ concentrates at $m = \pm 1$, allowing for a perfect prediction of hidden spins.

Our analytical results agree very well with simulations of spin systems with relative small number of spins. For these systems we could compute local magnetizations $m_a(t)$ exactly by enumeration. The Markovian spin dynamics facilitated these computations with the use of a *forward–backward* algorithm [16] well known for hidden Markov models (Appendix C). We compute $Q(t)$ using

$$Q(t) = \frac{1}{N_{\text{hid}}} \sum_{a=1}^{N_{\text{hid}}} E_{s,J} \, m_a^2(t) \,,$$

where $E_{s,J}$ denotes the expectation over all possible observed spins and over the set of random couplings.
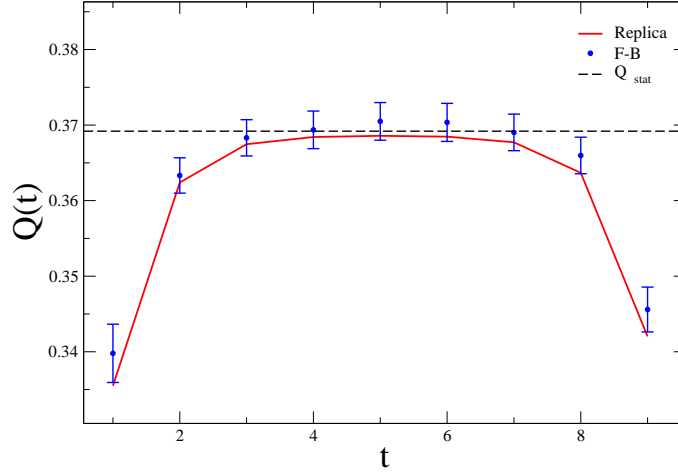
Figure 1: Order parameter Q as a function of time, for a system with $\lambda = 0.4$ and $k = 1$. Red line: solution of the order parameter equations. Black dashed line: stationary value Q_{stat} of the order parameter. Blue points: Q from numerical simulation of a system with $N_{\text{hid}} = 10$ hidden spins, averaged over 10000 samples; the error bars represent the standard deviation.
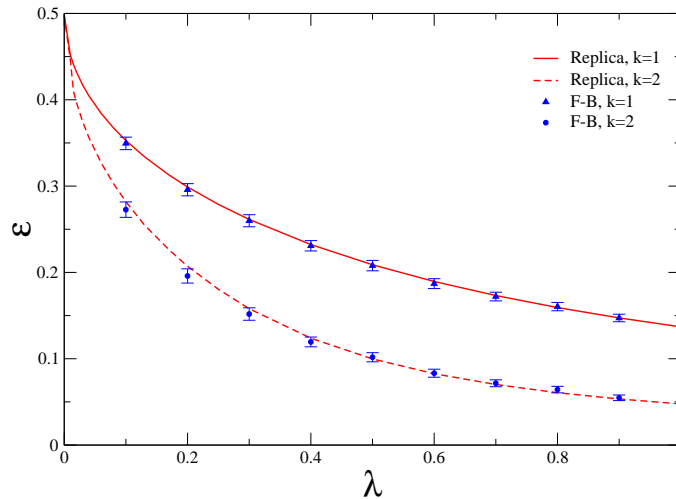

Figure 2: Bayes error as a function of the load factor for $k = 1$ (solid red line, blue triangles) and $k = 2$ (dashed red line, blue circles). Red lines: replica result, computed with the stationary values of the order parameters. Blue points: numerical simulation of a system with $N_{\text{hid}} = 8$ hidden spins, averaged over 2500 samples; the error bars represent the standard deviation; the Bayes error is computed at time $t = T/2$, with $T = 20$.
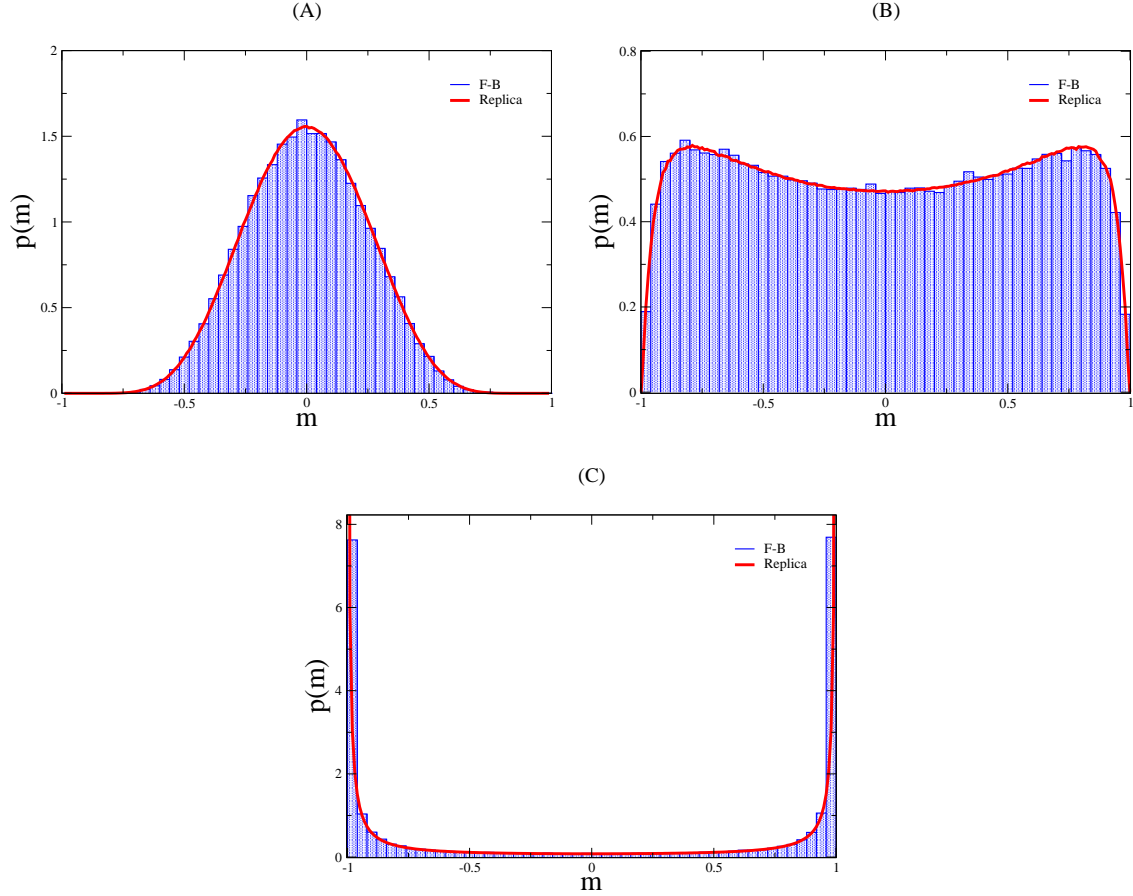
(A)

(B)

(C)

Figure 3: Distribution of local magnetization for load factor $\lambda = 0.8$ and coupling strengths $k = 0.2(A)$, $k = 0.6(B)$ and $k = 2(C)$. Red line: analytical result (eq. 34) assuming stationary values of the order parameters. Blue histogram: numerical simulations averaged over 80000 samples for a system with 8 hidden spins. The magnetization is computed at time $t = T/2$, with $T = 20$.

6. A comment on symmetric networks

From the point of view of equilibrium statistical physics a corresponding analysis for symmetric couplings $J_{ij} = J_{ji}$ might be of interest. In this case our approach would lead to additional order parameters (e.g. response functions). More important, order parameters would be usually non–zero for $t \neq t'$. Take for example the order parameter

$$
\begin{aligned}
C(t,t') =& E_J \left[\sum_{\{s\}} P(\{s\}) \sum_{\{\sigma\}} P(\{\sigma\}|\{s\})\sigma_a(t)\sigma_a(t') \right] \\
=& E_J \left[\sum_{\{s\},\{\sigma\}} P(\{\sigma\},\{s\})\sigma_a(t)\sigma_a(t') \right],
\end{aligned}
\tag{36}
$$

where the last line follows from Bayes theorem. Hence, $C(t, t')$ equals the usual spin correlation in a system of $N_{\mathrm{hid}} + N_{\mathrm{obs}} = N$ spins, where there is no difference between hidden and observed spins (because there is *no conditioning* on the latter ones). Unfortunately, even for this simpler, more standard type of spin–glass model (studied extensively in the 1990s), exact analytical results for two time correlations (except for the case of uncorrelated couplings and Gaussian or spherical spin models) were not possible. A Monte Carlo approach to the effective non–Markovian single spin dynamics [15,17] could be adapted to our model but it would require extensive nontrivial numerical simulations with an increasing complexity when the time window T grows. Moreover, this method cannot be easily extended to the stationary case.

To circumvent this problem, one might be tempted to resort to equilibrium techniques instead. In fact, for the case of symmetric couplings, the Markovian dynamics of the joint system of s and σ spins has a well known stationary equilibrium distribution. This static distribution is usually known as the Little model [18–20] and was frequently discussed in in the framework of Hopfield type neural networks with parallel dynamics. On might then calculate learning properties of the static model by using again the replica approach. While this should indeed be feasible (when replica symmetry breaking effects are neglected), one should note that this approach would consider a quite different statistical ensemble. The equilibrium case would deal with the probability $P(\sigma(t)|s(t))$ of spins at fixed large time t, whereas our dynamic ensemble is concerned with $P(\{\sigma\}|\{s\})$ with a conditioning on information $\{s\}$ from the *time history* of past and future observations.

Hence, the problem of solving the model with symmetric couplings is far different from the asymmetric case studied in this paper and will be postponed to future work.

7. Outlook

In this paper we have presented a first step in analyzing optimal Bayesian inference for kinetic Ising models with observed and unobserved spins valid for large random systems. The replica analysis revealed a fairly simple statistical picture of the posterior trajectories of hidden spins. Spins at different time steps (and sites) are statistically independent, but their local magnetizations depend on the propagation of information from past and future spins which is expressed through order parameters.

One can expect that this simple picture derived for the disorder averaged system can be translated into equations for the local magnetizations of hidden spins which are valid for a typical *single* system with fixed couplings and observations. In fact, such mean field equations generalizing the results of [1] to the case of observations can be derived from cavity arguments and could be used as an efficient algorithm for the computation of local magnetizations in large random networks. This could then be used as an approximation in the E-Step of an EM algorithm [6] which aims at computing the maximum likelihood estimator of the network couplings J_{ij}, averaging out unobserved spins. We will discuss such an approach in a forthcoming paper.

It will be interesting to extend this replica approach to other dynamical models. As long as we restrict ourselves to asymmetric random couplings one can expect that the case of continuous time (at least for the stationary limit) models could be treated. This would include e.g. continuous time Glauber dynamics and coupled stochastic differential equations (soft spin models).

Acknowledgements

A. Self consistent solution for the two time order parameters

Let us consider, as an example, the stationary value of the order parameter Q. From the saddle point equation $\frac{\partial f}{\partial Q} = 0$ we find:

$$
Q(t,t') = \frac{1}{\langle \Gamma_0 \prod_\tau Z(\tau) \rangle_{\xi,\zeta,\phi,\psi}}
\left\langle \frac{\langle \Gamma_0 \tanh A(t-1) \prod_\tau Z(\tau) \rangle_{\xi,\zeta} \, \langle \Gamma_0 \tanh A(t'-1) \prod_\tau Z(\tau) \rangle_{\xi,\zeta}}{\langle \Gamma_0 \prod_\tau Z(\tau) \rangle_{\xi,\zeta}} \right\rangle_{\phi,\psi},
\tag{37}
$$

where

$$
A(t) = \psi(t) + \zeta(t) + \phi(t+1) + \xi(t+1).
\tag{38}
$$

We want to show that $Q(t,t') = 0$ for $t \neq t'$ is a self consistent solution. If our assumption holds for the order parameters on the right hand side of equation (37), the averages over the gaussian fields factorize over time, yielding:

$$
Q(t,t') = \frac{\langle \tanh A(t-1) Z(t-1) \rangle_{\xi_{t-1},\zeta_{t-1},\phi_t,\psi_t}}{\langle Z(t-1) \rangle_{\xi_{t-1},\zeta_{t-1},\phi_t,\psi_t}} \frac{\langle \tanh A(t'-1) Z(t'-1) \rangle_{\xi_{t'-1},\zeta_{t'-1},\phi_{t'},\psi_{t'}}}{\langle Z(t'-1) \rangle_{\xi_{t'-1},\zeta_{t'-1},\phi_{t'},\psi_{t'}}}.
\tag{39}
$$

The first two terms in the numerator of the above equation can be written in terms of the independent random variables $x = \psi(t-1) + \zeta(t-1)$ and $y = \phi(t) + \xi(t)$ as

$$
\left\langle \frac{\sinh(x+y)}{\cosh(x)} \right\rangle_{x,y} = \left\langle \frac{\sinh(x)\cosh(y) + \cosh(x)\sinh(y)}{\cosh(x)} \right\rangle_{x,y}
$$
$$
= \langle \tanh(x) \rangle_x \langle \cosh(y) \rangle_y + \langle \sinh(y) \rangle_y = 0.
\tag{40}
$$

Using a similar procedure, this argument can be extended to all the other order parameters.

B. Boundary conditions

The parameters Γ_0 and $\tilde{\Gamma}_0$ containing the initial conditions have the following expression:

$$\Gamma_0 = 2\cosh[\phi(0) + \xi(0)], \qquad \tilde{\Gamma}_0 = \frac{\langle\cosh(\phi_0)\log(2\cosh(\phi_0))\rangle_{\phi_0}}{\langle\cosh(\phi_0)\rangle_{\phi_0}}. \qquad (41)$$

The initial and final condition for the order parameter are:

$$Q(0) = \frac{\langle\tanh(\phi_0)\sinh(\phi_0)\rangle_{\phi_0}}{\langle\cosh(\phi_0)\rangle_{\phi_0}}, \qquad \hat{Q}(T) = 0. \qquad (42)$$

C. Forward-backward algorithm

In order to compute the local magnetizations of hidden spins at each time t, we need the posterior distribution $P[\{\sigma\}(t)|\{s\}]\, 1 \leq t < T$ of the hidden spins at time t, given the obserserved spins at *all times*.

It is convenient to divide the computation of $P[\{\sigma\}(t)|\{s\}]$ in two parts, one involving the spins up to time $t+1$, the other the spins from $t+2$ to T:

$$\begin{aligned}
P[\{\sigma\}(t)|\{s\}] &= P[\{\sigma\}(t)|\{s\}_{1:t+1}, \{s\}_{t+2:T}] \\
&\propto P[\{\sigma\}(t)|\{s\}_{1:t+1}]\, P[\{s\}_{t+2:T}|\{\sigma\}(t), \{s\}(t+1)],
\end{aligned} \qquad (43)$$

where the last line follows from Bayes' rule and the conditional independence of $\{s\}_{t+2:T}$ and $\{s\}_{1:t}$ given $\{s\}(t+1)$ and $\{\sigma\}(t)$. The two terms in the right hand side of eq. (43) can be computed by recursion through time. In particular, it can be shown [16] that the first term, referred to as the "forward message", $\mathrm{fm}[\{\sigma\}(t)] = P[\{\sigma\}(t)|\{s\}_{1:t+1}]$, is obtained by a forward recursion form 1 to t governed by the following equation:

$$\mathrm{fm}[\{\sigma\}(t)] \propto P[\{s\}(t+1)|\{\sigma, s\}(t)] \sum_{\{\sigma\}(t-1)} P[\{\sigma\}(t)|\sigma(t-1), \{s\}(t-1)]\, \mathrm{fm}[\{\sigma\}(t-1)]. \qquad (44)$$

The second term, or "backward message" $\mathrm{bm}[\{\sigma\}(t+1)] = P[\{s\}_{t+2:T}|\{\sigma\}(t), \{s\}(t+1)]$, is obtained by a backward recursion, running from T to $t+1$ and obeying:

$$\mathrm{bm}[\{\sigma\}(t+1)] = \sum_{\{\sigma\}(t+1)} P[\{s\}(t+2)|\{\sigma, s\}(t+1)]\mathrm{bm}[\{\sigma\}(t+2)]\, P[\{\sigma\}(t+1)|\{\sigma, s\}(t)]. \qquad (45)$$

[1] Mézard, M. and Sakellariou, J.: Exact mean-field inference in asymmetric kinetic Ising systems. *J Stat Mech* 2011, L07001 (2011)
[2] Roudi, Y. and Hertz, J.: Mean Field Theory for Nonequilibrium Network Reconstruction. *Phys Rev Lett* 106, 048702 (2011)
[3] Zeng, H.-L., Aurell, E., Alava, M., and Mahmoudi, H.: Network inference using asynchronously updated kinetic Ising model. *Phys Rev E* 83, 041135 (2011)

[4] Sakellariou, J., Roudi, Y., Mézard, M., and Hertz, J.: Effect of coupling asymmetry on mean-field solutions of the direct and inverse SherringtonKirkpatrick model. *Philosophical Magazine* 92, 272–279 (2012)

[5] Huang, H. and Kabashima, Y.: Dynamics of asymmetric kinetic Ising systems revisited. *arXiv:13105003* (2013)

[6] Dempster, A. P., Laird, N. M., and Rubin, D. B.: Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society Series B* 39, 1–38 (1977)

[7] Dunn, B. and Roudi, Y.: Learning and inference in a nonequilibrium Ising model with hidden nodes. *Phys Rev E* 87, 022127 (2013)

[8] Tyrcha, J. and Hertz, J.: Network inference with hidden units. *MBE* 11, 149 (2014)

[9] Opper, M. and Kinzel, W.: Statistical Mechanics of Generalization. In Domany, E., van Hemmen, J., and Schulten, K., eds., *Models of Neural Networks III.* Springer-Verlag (1996)

[10] Nishimori, H.: *Statistical Physics of Spin Glasses and Information Processing: An Introduction.* Oxford University Press (2001)

[11] Engel, A. and Van den Broeck, C.: *Statistical Mechanics of Learning.* Cambridge University Press (2001)

[12] Crisanti, A. and Sompolinsky, H.: Dynamics of spin systems with randomly asymmetric bonds: Langevin dynamics and a spherical model. *Phys Rev A* 36, 4922–4939 (1987)

[13] Sompolinsky, H., Crisanti, A., and Sommers, H. J.: Chaos in Random Neural Networks. *Phys Rev Lett* 61, 259–262 (1988)

[14] Parisi, G.: Asymmetric neural networks and the process of learning. *Journal of Physics A: Mathematical and General* 19, L675 (1986)

[15] Eissfeller, H. and Opper, M.: Mean-field Monte Carlo approach to the Sherrington-Kirkpatrick model with asymmetric couplings. *Phys Rev E* 50, 709–720 (1994)

[16] Russel, S. and Norvig, P.: *Artificial Intelligence: A Modern Approach.* Pearson Education (1995)

[17] Scharnagl, A., Opper, M., and Kinzel, W.: On the relaxation of infinite-range spin glasses. *Journal of Physics A: Mathematical and General* 28, 5721 (1995)

[18] Little, W.: The existence of persistent states in the brain. *Mathematical Biosciences* 19, 101–120 (1974)

[19] Little, W. and Shaw, G. L.: A statistical theory of short and long term memory. *Behavioral Biology* 14, 115–133 (1975)

[20] Little, W. and Shaw, G. L.: Analytic study of the memory storage capacity of a neural network. *Mathematical Biosciences* 39, 281–290 (1978)