Gaussian average method for the kinetic Ising model

Claudia Battistin & Yasser Roudi

1. Centre for Neural Computation, NTNU, Trondheim, Norway

2. Kavli Institute for Systems Neuroscience

NETADIS Statistical Physics Approaches to Networks Across Disciplines

NTNU

Kavli Institute for Systems Neuroscience
Centre for Neural Computation

Norwegian Centre of Excellence

THE KAVLI FOUNDATION

1. Introduction

Kinetic Ising model provides a useful platform for studying learning and inference in non-equilibrium models with potential applications to hight-throughput data analysis.

The aim of this work is to provide a variational approach for calculating magnetization and correlation functions of a kinetic Ising model and to extend the technique to network reconstruction in biology.

In particular we considered a network of N interacting binary units, with couplings J_{ij}, whose synchronous dynamics is defined by the transition probability:

$$\text{Prob}[\sigma(t)|\sigma(t-1)] = \prod_{i=1}^{N} \frac{e^{\sigma_i(t+1)H_i(t)}}{2\cosh(H_i(t))}$$

where $\sigma(t)=\{\sigma_i(t)\}_{i=1,...,N}$ is the state of the system at time t and $H_i(t)=h_i(t)+\sum_j J_{ij}\sigma_j(t)$ is the field (external+generated by the other spins) acting on spin i at time t.

2. The Variational Approximation

The generating functional associate to the dynamics in the interval [0,T] can be expressed as:

$$Z[h,\psi] = \frac{1}{(2\pi)^{NT}} \int d\mathcal{G} e^{-L[h,\psi,\mathcal{G}]}$$

where the trace has been performed by introducing a set of 2NT auxiliary variables $G = \{g(t), \hat{g}(t)\}_{t=0,...,T}$, $g(t)=\{g_i(t)\}_{i=1,...,N}$.
While it's well known that the Saddle Point approximation of the generating functional generates Naive Mean Field equation, we have also shown that taking into account one-loop corrections one gets TAP equations.
Following the derivation of Mühlschlegel and Zittartz[1] we expand -ln Z[h,ψ] to the first order around the quadratic form $L_s = \frac{1}{2}(\mathcal{G}-\overline{\alpha})^\top \mathbf{S}(\mathcal{G}-\overline{\alpha})$

$$-\ln(Z[h,\psi]) \simeq -\ln \int D\mathcal{G} + \frac{\int D\mathcal{G}\,(L-L_s)}{\int D\mathcal{G}} + NT\ln 2\pi$$

By optimizing the right-hand side of the expression above with respect to the parameters $\mathbf{S}, \overline{\alpha}$ in the limit of $\psi \to 0$ one gets the following set of equations:

Saddle Point Equations:

$$m_i(t) = \frac{\det \mathbf{S}^{1/2}}{(2\pi)^{NT}} \int D\mathcal{G}\ \tanh[g_i(t-1)+\alpha_i(t-1)]$$

$$\gamma_{ij}(t) = \sum_k J_{ik}J_{jk}\left(1-m_i^2(t)\right)$$

$$\lambda_{ij}(t,t+1) = iJ_{ji}(1-m_i^2(t+1))$$

$$\alpha_i(t) = h_i(t) + \sum_j J_{ij}m_j(t)$$

$$\hat{\alpha}_i(t) = 0$$

where $m_i(t)$ is the magnetization of the unit i at time t, while γ and λ are blocks of the matrix $\mathbf{S}$ that takes the following shape:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}(0,0) & \mathbf{S}(0,1) & 0 & 0 & 0 & \cdots \\ \mathbf{S}(1,0) & \mathbf{S}(1,1) & \mathbf{S}(1,2) & 0 & 0 & \cdots \\ 0 & \mathbf{S}(2,1) & \mathbf{S}(2,2) & \mathbf{S}(2,3) & 0 & \cdots \\ 0 & 0 & \mathbf{S}(3,2) & \mathbf{S}(3,3) & \mathbf{S}(3,4) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

with $\mathbf{S}(t,t) = \left[\begin{array}{c|c} 0 & -i\mathbb{I} \\ \hline -i\mathbb{I} & \gamma(t) \end{array}\right]$ and $\mathbf{S}(t,t+1) = \mathbf{S}(t+1,t)^\top = \left[\begin{array}{c|c} 0 & \lambda(t,t+1) \\ \hline 0 & 0 \end{array}\right]$

The optimization provides us a closed set of equation, though not explicit expression of $m_i(t)$ as function of $m_i(t-1)$.

How to solve it?

- According to the features of the distribution of the couplings and the couplings strength the integral for the magnetization can be dimensionally reduced (cf. Numerical Test). The averages under iid normally distributed random variables can be estimated by Monte Carlo sampling (future direction).
- The most general approach to the problem is the construction of an ADAPTIVE ALGORITHM, though the complex generating functional landscape rises issues on the convergence of the algorithm that has to be analysed carefully (future direction).

3. Numerical tests

We chose the couplings distribution to be parametrized as in [3]:

$$J_{ij} = J_{ij}^{\text{sym}} + kJ_{ij}^{\text{asym}}$$

$$J_{ij}^{\text{asym}} = -J_{ji}^{\text{asym}} \qquad J_{ij}^{\text{sym}} = J_{ji}^{\text{sym}}$$

and

$$\left\langle \left(J_{ij}^{\text{sym}}\right)^2 \right\rangle = \left\langle \left(J_{ij}^{\text{asym}}\right)^2 \right\rangle = \frac{g^2}{N(1+k^2)}$$

where J_{asym} and J_{sym} are normally distributed iid random variables with zero mean.

This choice for the couplings distribution allows to reduce the saddle point equation for the magnetization to a sum of 2 one-dimensional integrals:

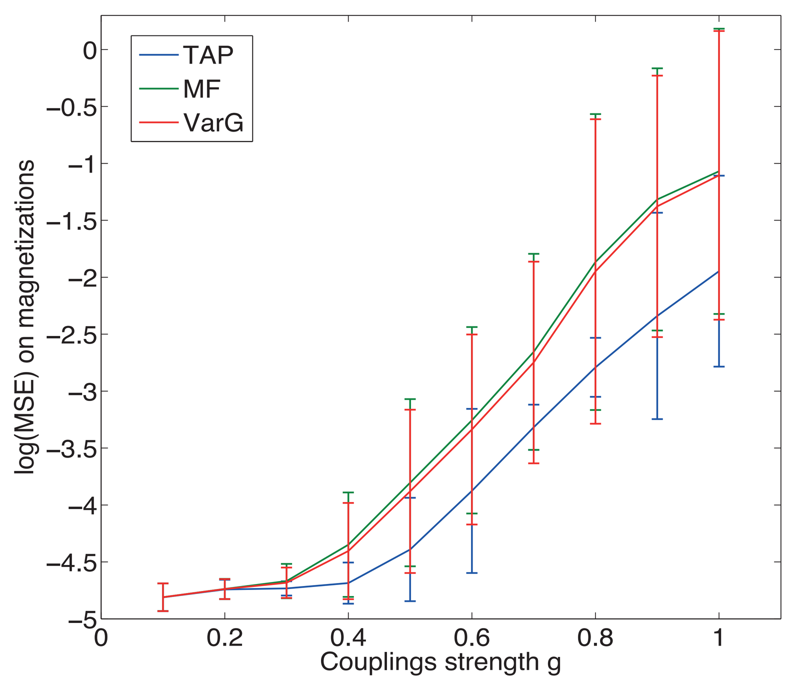
$$m_i(t) = \int DxR_i(x,t-1) - \sum_{j\neq i}\left\{\left(\gamma(t-1)^{1/2}\right)_{ij}\right\}^2 \int DxR_j(x,t-1)\left\{1-R_j(x,t-1)^2\right\} + O(J^4)$$

where $R_i(x,t) \equiv \tanh\left[\left(\gamma(t)^{1/2}\right)_{ii} x + \alpha_i(t)\right]$

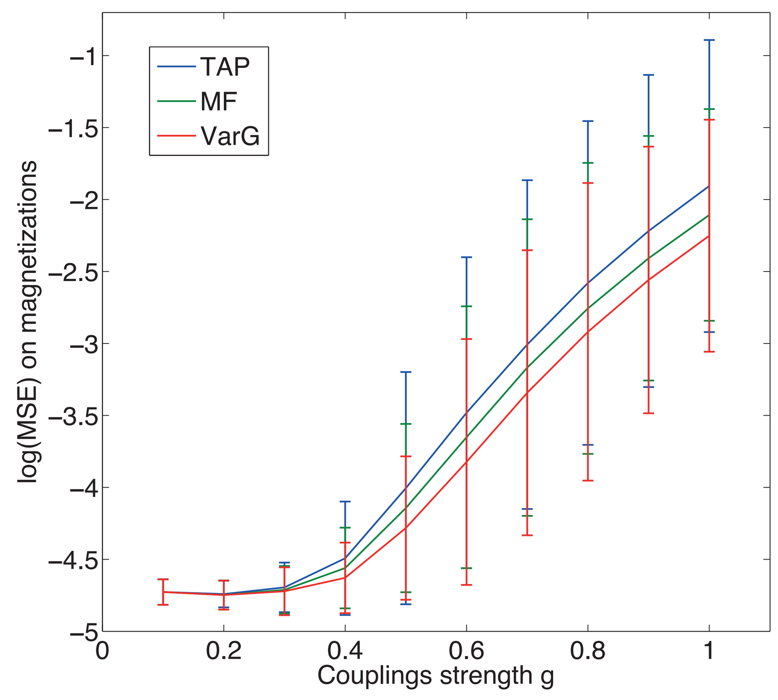
We studied the performances of TAP, MF and VarG in terms of Mean Square error on the magnetizations,

$$MSE = \frac{1}{T}\frac{1}{N}\left\langle \sum_{i=1}^{N}\sum_{t=1}^{T}\left(m_i(t) - m_i^{\exp}(t)\right)^2 \right\rangle_{Jrealizations}$$

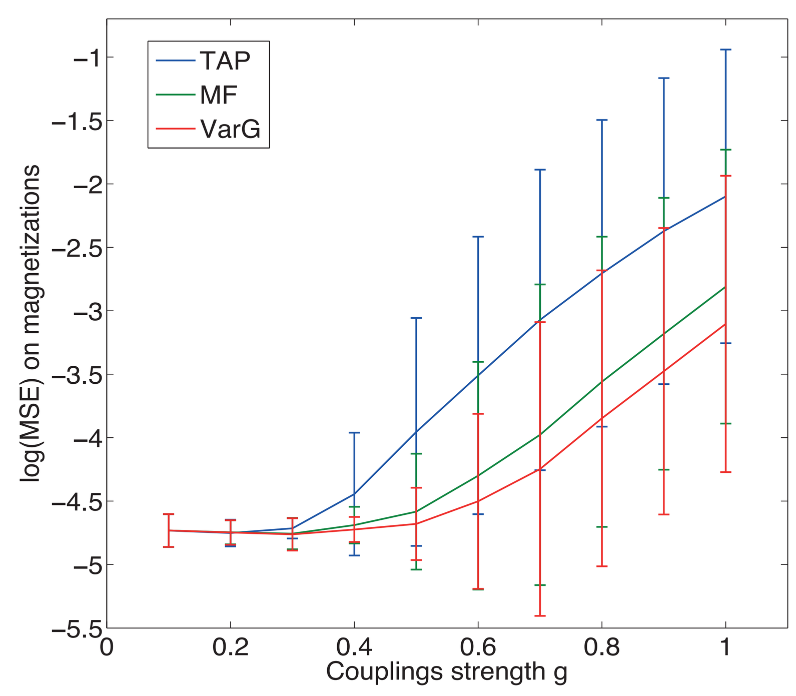
where the approximations predict the entire dynamics.



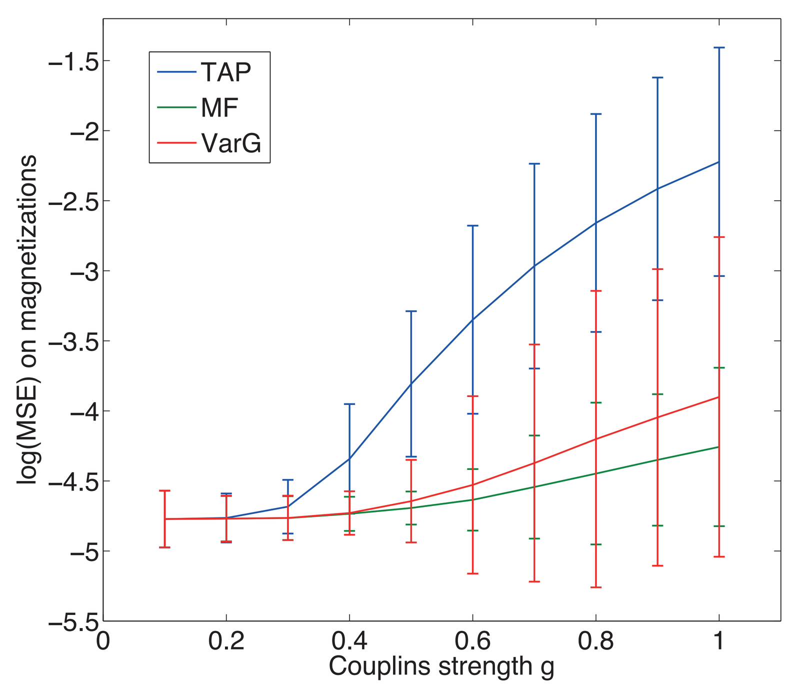
k=0



k=0.4



k=0.6



k=1

N=50, T=100, Jreal=10, nIter=50000

4. Weak couplings

For weak couplings one can neglect the λ matrix entries in the interaction matrix S and expand the tanh in the Saddle Point equation for the magnetizations around $\alpha_i(t-1)$. The resulting expression, up to the second order in $(\gamma^{1/2})_{ij}$ then will correspond to the first order expansion of :

$$m_i(t) = \tanh\left[\alpha_i(t-1) - m_i(t)\sum_j\left(\gamma^{1/2}\right)_{ij}^2\right]$$

By simple algebraic arguments one can show that this equations are the TAP equation:

$$m_i(t) = \tanh\left[h_i(t-1) + \sum_j J_{ij}m_j(t-1) - m_i(t)\sum_j J_{ij}^2\left(1-m_j(t-1)^2\right)\right]$$

Surprisingly the loops introduced by the out-of-diagonal terms of γ do not give any relevant contribution to the magnetizations in the weak coupling limit.

5. Fully asymmetric couplings

We consider the case of fully asymmetric couplings (fully asymmetric kinetic Sherrington-Kirkpatrick model):

$$J_{ij} \sim \mathcal{N}(0, 1/N^{1/2}) \quad \text{iid}$$

For large systems the Saddle point equations reduce to:

$$m_i(t) = \int \frac{dz}{(2\pi)^{1/2}} e^{-z^2/2} \tanh[z\sqrt{\Delta_i(t-1)} + h_i(t-1) + \sum_k J_{ik}m_k(t-1)]$$

$$\Delta_i(t) = \sum_j J_{ij}^2 \int \frac{dz}{(2\pi)^{1/2}} e^{-z^2/2}[1-\tanh^2[z\sqrt{\Delta_j(t-1)} + h_j(t-1) + \sum_k J_{jk}m_k(t-1)]]$$

that are the Exact Mean Field Equations for fully asymmetric couplings derived by Mezard and Sakellariou in [2].

6. Conclusions

1. We developed a Variational Gaussian Approach to the Kinetic Ising Model with synchronous update. The equations are derived without making any assumption on the couplings distribution and therefore are suitable for further simplification and numerical implementation tailored to features of the specific system.

2. We recover the mean field equations of [2] in the limit of asymmetric couplings and match TAP first order expansion for weak couplings.

3. The preliminary numerical results show that the Gaussian Average Method outperforms the mean field theory in [2] and TAP for coupling of intermediate asymmetry.

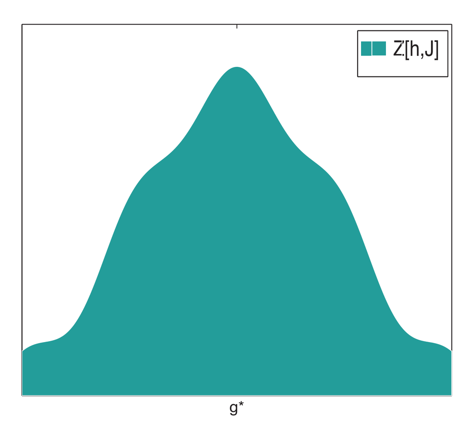
7. Future Directions

A. We are interested in studying how the perfomance of the Gaussian Average Method scales with the size of network and the length of the data.

B. We will apply our approximation to the inverse problem, focusing on sparse networks, the Generalized Linear Model and the hidden nodes case.

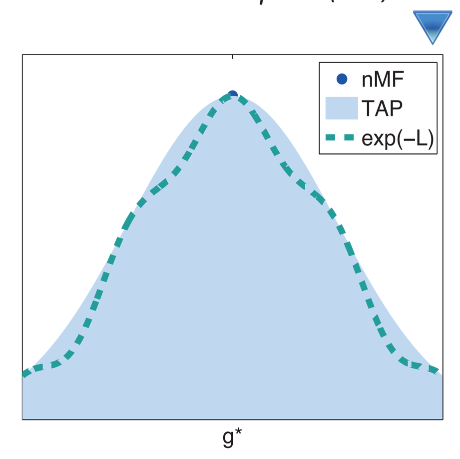
C. We will also study another variational method in which instead of taylor expanding the generating functional, we will use the KL divergence between a factorized distribution and the full history.

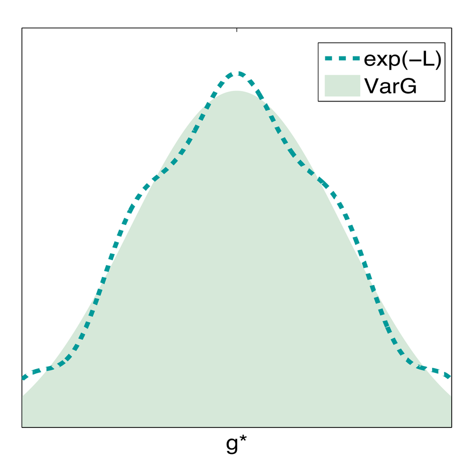
THE BASIC IDEA



Given the generating functional integral of a certain distribution in the space of the G's

one can approximate it with its saddle point value (nMF) or through a gaussian with the same curvature at the saddle point (TAP)





or by taking the optimal Gaussian integral (Gaussian Average Method)

[1] B. Mu hlschlegel, H. Zittartz , Gaussian Average Method in the Statistical Theory of the Ising model, *Zeitschrift fur Physik* (1963), Volume 175, Issue 5, pp 553-573

[2] M. Mezard and J. Sakellariou, Exact mean-field inference in asymmetric kinetic Ising systems, *J. Stat. Mech.* (2011) L07001 doi:10.1088/1742-5468/2011/07/L07001

[3] J. Sakellariou, Y. Roudi, M. Mezard & J. Hertz , Effect of coupling asymmetry on mean-field solutions of the direct and inverse Sherrington-Kirkpatrick model, *Philosophical Magazine*

Supported by the Norwegian Research Council Centre of Excellence Grant