Dottorato di Ricerca in Fisica, XXX Ciclo

Università degli studi di Roma la Sapienza

PhD research project

April 4th, 2016

On the optimal use of the Bethe approximation for models on graphs with loops

Candidate: Gabriele Perugini

Supervisor: Prof. Federico Ricci-Tersenghi

Introduction

Among mean field approximations the Bethe approximation (BA)[1] has attracted a renewed attention in the last two decades. Under the name of cavity method[2,3] it was succesfully applied to many different problems, such as spin glasses, error correcting codes and constraint satisfaction problems[4].

The Bethe approximation is exact for models defined on trees, and most of the interesting results were obtained in models defined on locally tree-like graphs, e.g. random graphs where loops mean size diverges in the thermodynamic limit.

Beyond random graphs, i.e. on graphs with short loops, very little is known about how good is the BA. In recent years some efforts have been done for computing the finite size corrections to disordered systems defined on locally tree-like graphs[5,6] and it has been found that these corrections are related to the density of short loops. However in the general case it is unclear how wide is the application range of the BA, and in particular whether it can be of any practical use for models defined on regular lattices.

Despite the presence of short loops is known to be one of the main limitations to a wider use of the BA, we strongly believe that BA can provide some interesting insight also on models where it is expected not to be exact.

On the algorithmic side, it is worth noting that the BA self-consistency equations can be conveniently solved[7] via an iterative message passing algorithm called Belief Propagation (BP)[8], which is very fast and efficient in many practical applications (and thus often used even in absence of any proof of convergence or exactness).

Besides the enormous empirical success of the BP algorithm over the last years (some nice examples are[9–11]) the analytic results supporting its correctness are somehow limited. The exactness of the BP algorithm was demonstrated only on trees and graphs with up to one single loop[12,13]. Moreover its application is limited to models where replica symmetry holds, a conjecture[14] that was rigorously proved only recently[15]. For some specific models on locally tree like graphs the exactness of BP was rigorously proved under the assumption on the uniqueness of the Gibbs state[16,17]. What the behavior of the BP algorithm should be away from straightforward applications is still unclear.

The main purpose of my Ph.D. will be to characterize the Bethe approximation, going beyond the standard field of its application. Indeed, as preliminary results suggest, much information can be extracted both from analytical and computational grounds, even on models defined on graphs with loops, or in models where the uniqueness of the Gibbs measure does not hold. Moreover, even when it is known that BP fails (e.g. it does not converge) one can modify BP and construct fast and robust algorithm for useful tasks such as optimization.

Present work

At the present time we are performing a detailed study of the Bethe approximation on two disordered Ising spin models: the random field Ising model (RFIM)[18] and the Ising spin glass (SG) model[4]. These two models represent an interesting challenge for the BA: on one side the RFIM presents many different BP fixed points, possibly breaking the assumption of the uniqueness of the Gibbs state in the thermodynamic limit, on the other side bypassing the lack of convergence of BP on SG problems due to the replica symmetry breaking would be of great importance for the development of heuristic optimization algorithms.

Bethe approximation and metastable states in the zero temperature RFIM

Only recently it was shown[19] that the global minimum of the Bethe free energy is the maximum likelihood solution of the zero temperature RFIM, irrespective of the graph on which is defined on (so even on graphs with loops). This motivated us to perform a detailed analysis of the performance of the BP algorithm, whose fixed points correspond to the local minima of the Bethe free energy.

The RFIM is one of the simplest prototypes of disordered systems with quenched disorder and still presents many unanswered questions (for a recent review on some open problem see[20]). Even if the presence of replica symmetry breaking scenario was excluded only recently[21], the failure of dimensional reduction[22] is still an open problem, and some aspects of the phase transition still need to be clarified, as well as the anomalous slow dynamics[23]. Renormalization group arguments show that the model can be studied in the zero temperature limit. This represent a great advantage since at $T = 0$ an exact algorithm exists with polynomial time complexity, the min-cut/max-flow algorithm[24], thus providing a benchmark for numerical studies.

We are considering the RFIM defined on two very different topologies: (i) random regular graphs (RRG), i.e. an ensemble of locally tree like graphs with fixed connectivity/degree, and (ii) 3-dimensional lattices. On RRG the BA is expected to be exact in the thermodynamic limit, but the presence of loops in finite graphs could induce some non-perturbative effects, while the 3D case is expected to be a much worst case scenario, due to the presence of many short loops.
At present we have some preliminary results suggesting that BA can provide some useful insight even on models with short loops:

- It is known that the zero temperature RFIM present a large number of 1-spin-flip-stable fixed points inside the hysteresis loop. They have been studied extensively in recent works[25,26] and analytical characterization has been obtained for the lower and upper branch of the hysteresis curve (see for example[27,28]). These states can be reached with a standard zero temperature Glauber dynamics[29] or in general with a unidirectional dynamics scheme[30]. However we verified that these states are unstable against small thermal fluctuations, and they can be seen as a byproduct of the $T \to 0$ limit in a finite model.
  In contrast, the BP fixed points are obtained from a set of self consistency equations derived as the zero temperature limit of the model in the thermodynamic limit[31] and they are thermally stable as numerical simulations suggest. Indeed it was shown that the so called "maximal neighborhood" stability property[32] holds for BP fixed points, which let the algorithm avoid the (exponential number of) trivial fixed points present in the model.

- We found by simple inequalities that two initial conditions exist that provide maximal solutions, the ones with largest and smallest magnetization. Every fixed point configuration is bounded by these maximal solutions.

In the critical region we found a percolation phenomenon for the unfrozen variables, i.e. those variables taking different values in the two extremal configurations. We located the percolation threshold of the unfrozen variables and found that it coincides with the thermodynamic phase transition on the RRG ensemble, while in the 3D case it precedes the phase transition. Interestingly the largest computational complexity is found close to the percolation threshold, rather than at the thermodynamical critical point. We believe this can have direct consequences on the RFIM dynamics.

- We designed a modified BP algorithm which is able to identify many different local minima of the Bethe free energy, by initializing BP with a convex combination of the already found fixed points. We observed that the number of stable BP fixed points (local minima of the Bethe free-energy) increases with the system size at the critical point. The algorithm is always convergent and every configuration is reached in almost linear time. Furthermore this modified version of BP is able to find the true ground state of the model in a very competitive time, at most on random graphs, and it is a much more efficient way of exploring the Bethe free energy landscape, since few initial condition are needed to find most of the solution (in contrast to the widespread method of randomly initialize BP with a lot of initial condition[33]). An analysis of the structure of these stable states seems to suggests the possibility to work out a reweighted mean field theory, that until now was tested only in the fully connected topology[34].

What we learned and observed up to now on the RFIM raised several interesting questions:

- In general it seems that, at least in the critical region, the number of stable, non trivial BP fixed points increase with the system size. So that what one should expect in the thermodynamic limit? All the recent effort to numerically analyze the RFIM are based on the exact ground state computation with the min-cut algorithm[20,35]. Can it be the case that only the ground state is not enough for a full description of the model?

- On random graphs, despite the number of states grows in the critical region, the population dynamics algorithm[4] is unable to detect more than one state. Is the model homogeneity assumption wrong for the RFIM?

- A lot of exact results on the BP algorithm rely on the uniqueness of the BP fixed point. Is it possible to extend some of these results in the case the model displays more than one fixed point? Can we characterize the basin of attraction of these fixed points at least in a specific model such as the RFIM?

BP as an heuristic for SG optimization problems

One of the best known heuristics for the spin glass (SG) optimization problem[24,36] consists in exactly optimizing a subset of nodes that form an unfrustrated cluster with the min-cut algorithm[37]. Compared to the min-cut algorithm, the BP algorithm (although not provably exact in presence of loops) is very fast and effective: e.g., it can easily found the ground state even in the presence of a small amount of frustration. Optimizing clusters with a controlled number of frustrated links seems to be a much more powerful tool to find the global optimum. This could be a great application of the loopy BP, since at present numerical simulations of very large systems seem to be the only way for solving some open questions on the SG nature.

Another promising approach that we are testing consist in randomly pinning (i.e. fixing) a fraction of variables and iterate BP on a modified graph where the pinned variables act as an external field. Despite its simplicity, this scheme reveals some interesting features. As expected, the convergence of BP is guaranteed only above a certain fraction of pinned variables. However we observed that as the fraction of pinned spins is increased from 0 to 1, the energy monotonically decreases. The average energy obtained in this way seems to be consistent with recent numerical simulations on the $3D$ EA model[38], proposing this modified BP as a valid, almost-linear alternative for SG optimization. Furthermore the behavior of this algorithm is telling us that BP outputs meaningful marginals even when it does not converge. Indeed the low energy state is reached by initializing the BP equation with the fixed point found with lower fraction of pinned variables (randomly initialize BP while lowering the free spins will result in a trivial fixed point). Hopefully a detailed characterization of this process will support our conjecture about the usefulness of BP beyond safe grounds applications.

Future work

Many research directions can naturally follow these ongoing projects.
A standard application will be the development of new algorithms for solving computationally hard problem (e.g. constraint satisfaction problems) via modified BP algorithms. Our main interest will be on disordered Ising spin models, but hopefully we will be able to extend the present study to the

case of continuous spin variables, which may find an even broader field of application, for example in solving maximum likelihood problems via convex optimization[39].

---

[1] H. A. Bethe. Statistical Theory of Superlattices. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 150(871):552–575, jul 1935.

[2] M. Mézard and G. Parisi. The Bethe lattice spin glass revisited. *The European Physical Journal B*, 20(2):217–233, mar 2001.

[3] Marc Mézard and Giorgio Parisi. The Cavity Method at Zero Temperature. *Journal of Statistical Physics*, 111(1-2):1–34, 2003.

[4] Marc Mézard and Andrea Montanari. *Information, Physics, and Computation.* Oxford University Press, 2009.

[5] U. Ferrari, C. Lucibello, F. Morone, G. Parisi, F. Ricci-Tersenghi, and T. Rizzo. Finite-size corrections to disordered systems on Erdös-Rényi random graphs. *Physical Review B*, 88(18):184201, nov 2013.

[6] C Lucibello, F Morone, G Parisi, F Ricci-Tersenghi, and Tommaso Rizzo. Finite-size corrections to disordered Ising models on random regular graphs. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 90(1):012146, jul 2014.

[7] J.S. Yedidia, W.T. Freeman, and Y. Weiss. Constructing Free-Energy Approximations and Generalized Belief Propagation Algorithms. *IEEE Transactions on Information Theory*, 51(7):2282–2312, jul 2005.

[8] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann, San Francisco, 1988.

[9] F Ricci-Tersenghi and G Semerjian. On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms. *Journal of Statistical Mechanics: . . .*, P09001, 2009.

[10] T Rogers, IP Castillo, R Kühn, and K Takeda. Cavity approach to the spectral density of sparse symmetric random matrices. *Physical Review E*, 2008.

[11] Florent Krzakala, Cristopher Moore, Elchanan Mossel, Joe Neeman, Allan Sly, Lenka Zdeborová, and Pan Zhang. Spectral redemption in clustering sparse networks. *Proceedings of the National Academy of Sciences of the United States of America*, 110(52):20935–40, dec 2013.

[12] Y Weiss. Belief propagation and revision in networks with loops. *MIT AI Lab., Tech Rep*, 1616, 1997.

[13] Yair Weiss. Correctness of Local Probability Propagation in Graphical Models with Loops. *Neural Computation*, 12(1):1–41, jan 2000.

[14] Florent Krzakala, Andrea Montanari, Federico Ricci-Tersenghi, Guilhem Semerjian, and Lenka Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25):10318–23, jun 2007.

[15] A Coja-Oghlan and W Perkins. Belief Propagation on replica symmetric random factor graph models.

arXiv preprint arXiv:1603.08191, 2016.

[16] A Dembo, A Montanari, and N Sun. Factor models on locally tree-like graphs. *The Annals of Probability*, 2013.

[17] A Dembo, A Montanari, A Sly, and N Sun. The replica symmetric solution for Potts models on d-regular graphs. *Communications in Mathematical ...*, 327(2):551–575, 2014.

[18] T. Nattermann. Theory of the Random Field Ising Model. In *Spin Glasses and Random Fields, ed. by A. P. Young, World Scientific*. 1997.

[19] Michael Chertkov. Exactness of belief propagation for some graphical models with loops. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10016, oct 2008.

[20] NG Fytas and V Martín-Mayor. Universality in the three-dimensional random-field Ising model. *Physical review letters*, 2013.

[21] F Krzakala, F Ricci-Tersenghi, and L Zdeborová. Elusive spin-glass phase in the random field Ising model. *Physical review letters*, 2010.

[22] G. Parisi and N. Sourlas. Random Magnetic Fields, Supersymmetry, and Negative Dimensions. *Physical Review Letters*, 43(11):744–745, sep 1979.

[23] S von Ohr, M Manssen, and AK Hartmann. Aging in the three-dimensional Random Field Ising Model. *arXiv preprint arXiv:1601.02455*, 2016.

[24] AK Hartmann and H Rieger. *Optimization algorithms in physics*. Wiley, 2002.

[25] F. Detcheverry, M. L. Rosinberg, and G. Tarjus. Metastable states and T=0 hysteresis in the random-field Ising model on random graphs. *The European Physical Journal B*, 44(3):327–343, apr 2005.

[26] M L Rosinberg, G Tarjus, and F J Pérez-Reche. The T = 0 random-field Ising model on a Bethe lattice with large coordination number: hysteresis and metastable states. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(03):P03003, mar 2009.

[27] Deepak Dhar, Prabodh Shukla, and James P Sethna. Zero-temperature hysteresis in the random-field Ising model on a Bethe lattice. *Journal of Physics A: Mathematical and General*, 30(15):5259–5267, aug 1997.

[28] T P Handford, F-J Perez-Reche, and S N Taraskin. Exact spin-spin correlation function for the zero-temperature random-field Ising model. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(01):P01001, jan 2012.

[29] H Ohta and S Sasa. A universal form of slow dynamics in zero-temperature random-field Ising model. *EPL (Europhysics Letters)*, 2010.

[30] AY Lokhov, M Mézard, and L Zdeborová. Dynamic message-passing equations for models with unidirectional dynamics. *Physical Review E*, 2015.

[31] F Morone, G Parisi, and F Ricci-Tersenghi. Large deviations of correlation functions in random magnets. *Physical Review B*, 2014.

[32] Y. Weiss and W.T. Freeman. On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs. *IEEE Transactions on Information Theory*, 47(2):736–744, 2001.

[33] E Dominguez, A Lage-Castellanos, and R Mulet. Random Field Ising Model in two dimensions: Bethe approximation, Cluster Variational Method and message passing algorithms. *arXiv preprint arXiv: . . .*, 2015.

[34] D Lancaster, E Marinari, and G Parisi. Weighted mean-field theory for the random field Ising model. *Journal of Physics A: Mathematical and General*, 28(14):3959–3973, 1995.

[35] PE Theodorakis and NG Fytas. Random-field Ising model: Insight from zero-temperature simulations. *arXiv preprint arXiv:1501.02338*, 2015.

[36] AK Hartmann and H Rieger. *New optimization algorithms in physics*. Wiley, 2006.

[37] AK Hartmann. Cluster-exact approximation of spin glass groundstates. *Physica A: Statistical Mechanics and its Applications*, 1996.

[38] F Romá and S Risau-Gusman. The ground state energy of the Edwards Anderson spin glass model with a parallel tempering Monte Carlo algorithm. *Physica A: Statistical . . .*, 388(14):2821–2838, 2009.

[39] Adel Javanmard, Andrea Montanari, and Federico Ricci-Tersenghi. Phase Transitions in Semidefinite Relaxations. page 71, nov 2015.