The Critical Properties of Magnetic System Using Generalized Belief Propagation Technique

Tomasz M. Gwizdałła

University of Łódź, Dept. of Solid State Physics
Pomorska 149/153, 90-236 Łódź, Poland
email: tomgwizd@uni.lodz.pl

Abstract. Since few years the Belief Propagation [13, 14, 15] algorithm is reported as a very efficient tool to perform the optimization of systems which can be topologically transformed to the one of acceptable equivalent forms [9, 7]. The Ising system is often mentioned in these papers as a good example to present some basic foundations of BP. It is however rarely used as a tool to solve the Ising system itself. In this article we are going to present the analysis of critical properties, connected to the phase transition of magnetic system described by the Ising hamiltonian and the comparison of results to those obtained using evolutionary algorithm.

1 Belief Propagation Algorithm and Its Generalization

Consider the well known Ising spin system on the two-dimensional, square lattice. The spins are numbered with indices i, j and can take one of the two values from the set $\{-1, 1\}$. These values may be understood as an orientation of z-component of spin up or down. The energy of such a system is given by the formula 1.

$$H = -J \sum_{ij} S_i S_j - h \sum_i S_i. \tag{1}$$

The summation in the first part goes only over the nearest neighbours of selected spin. This part is responsible for the mutual magnetic interaction between spins. In the most simple model, proposed in 1920 by Lenz and solved later by Ising (for one dimension), there were the same spins spread over lattice and the exchange integral J is equal for all pairs. More generally its value can be different for different pairs and then must be included into summation. The second part is the interaction of single spin with the external magnetic field. In our attempt we will use the basic model without the external field.

The search for the minimum of total energy wouldn't certainly give us the information about the optimum state in the arbitrary temperature. In order to perform such optimization we have to take into account the value of so called Gibbs free energy, given by the formula:

$$G = U - TS \tag{2}$$

The crucial problem during performing calculations according to the presented formula is the determination of entropy. The usual way of doing it is to calculate the configurational entropy of system. Among the algorithm allowing this one can mention: well known Boltzmann/Gibbs/Shannon formula; Bukman's approach based on the cluster expansion [1]; Tsallis extensive entropy [12] or Meirovitch formula [6, 11]. Not all of these attempts (like Meirovitch one) can be directly used for single sample situation and the physical as well as algorithmical consequences of using other methods have been presented in earlier papers [3, 4].

Here we are going to use the other numerical approach which allows to minimize the Gibbs free energy - the belief propagation technique. Its detailed description, commonly with some mathematical proofs and derivations, may be found elsewhere [13, 14, 15] so here let us to present only the short sketch showing basics and main features. The starting point is the topological equivalence between three different types of structures: Bayesian Networks (BN), Markov Random Fields (MRF), and Tanner Factor Graphs (FG). Characteristic feature of Bayesian Network [10] is the possibility to pass the message between different nodes of network. These messages reflect the conditional probability of reaching the assumed final state of the given node. From the other hand, the magnetic Ising model can be well described by the MRF with the hidden nodes corresponding to the interaction with external field or by the FG with squares characterizing the spin-spin interaction.

The analysis we will start from the formula which is directly adapted from the statistical physics. Let x will be the set of states x1, x2, ..., xn of successively numbered nodes of system. Notice that we change the notation from $s_i \in \{-1, 1\}$ to $x_i \in \{0, 1\}$ in order to be consistent with standard notation. The probability that the system will be in the state $\{x\}$ is equal to:

$$p\Big(\{x\}\Big) = \frac{1}{Z} \prod_{ij} \Psi_{ij}(x_i, x_j) \prod_{i} \Phi_i(x_i) \tag{3}$$

This formula is certainly nothing else than Boltzmann probability for canonical system and Z is the normalization constant called partition function. Formula 3 corresponds exactly to the distribution for Ising magnet if we assume that:

$$\begin{aligned} \Phi_i(x_i) &= exp(hs_i) \\ \Psi_{ij}(x_i, x_j) &= exp\Big(\frac{Js_is_j}{T}\Big) \end{aligned} \tag{4}$$

In the basic version of belief propagation algorithm we need indeed two equations to determine the behaviour of sample. The belief that the state of x_i variable will be equal to the one of the possible values may be calculated:

$$b_i(x_i) = C_{norm} \Phi_i(x_i) \prod_{j \in N(i)} m_{ji}(x_i) \tag{5}$$

It is obvious that all probabilities should be normalized to 1. In the situations when the real probabilities are calculated by the final normalization of earlier obtained values we will use the C_{norm} notation. The m_{ij} values are the messages from the neighbouring nodes determined in the way:

$$m_{ij}(x_j) = C_{norm} \sum_{x_i} \Phi_i(x_i) \Psi_{ij}(x_i, x_j) \prod_{k \in N(i) \setminus j} m_{ki}(x_i) \tag{6}$$

Let us now shortly comment the two above equations. First of all, formula 5 makes it possible to define easy the magnetization of sample. The b_i value defined there is indeed the n-dimensional vector, where n is the number of possible states at selected position. In our case $n = 2$ because in the Ising model we have only, mentioned in earlier paragraph, two states up/down. It is important that due to symmetry of model all positions are equivalent and in the steady state the beliefs of reaching the same directions are equal for all positions. Thus, considering beliefs as probabilities, the average magnetization per site can be written:

$$m = p(-1) * (-1) + p(1) * 1 = -b_i(0) + b_i(1) \tag{7}$$

The variable exchange between the sets characteristic for magnetism and this one for Belief Propagation, visible in the above formula, was described earlier.

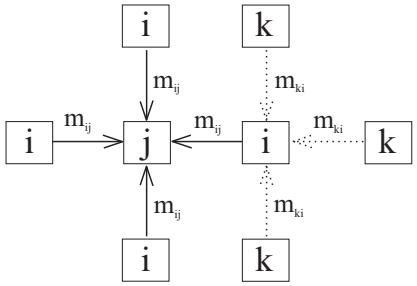

Figure 1. Scheme of message flows between nodes in Belief Propagation method illustrating equations 5 and 6

The schematic view of message and belief propagation between nearest neighbours on the two-dimensional square matrix is presented on figure 1. The node with updated beliefs is denoted by j. Solid lines correspond to the messages m_{ji} from its nearest neighbours (all numbered with i) as presented in formula 5. These messages are hovewer affected by the their nearest neighbours (k) and the messages m_{ki} ploted with dotted lines for only one selected i node. It is important to notice that there is no dotted arrow between nodes i and j, so the message between them doesn't have an influence on itself.

Equation 6 shows that messages have to be determined self-consistently so it defines also clearly the message update procedure.

In order to calculate thermodynamical potentials we have to do successive steps. First of all the two-node beliefs should be calculated:

$$b_{ij}(x_i, x_j) = C_{norm} \Psi_{ij}(x_i, x_j) \Phi_i(x_i) \Phi_j(x_j) \prod_{k \in N(i) \setminus j} m_{ki}(x_i) \prod_{l \in N(j) \setminus i} m_{lj}(x_j) \tag{8}$$

and on this basis the entropy and energy of system in the Bethe approximation can be calculated:

$$S = \sum_{ij} \sum_{x_i,x_j} b_{ij}(x_i,x_j) ln b_{ij}(x_i,x_j) + \sum_i (q_i - 1) \sum_{x_i} b_i(x_i) ln b_i(x_i) \tag{9}$$

where qi is the number of nearest neighbours and energies have their physical meaning:

$$U = \sum_{ij} \sum_{x_i,x_j} b_{ij}(x_i,x_j) E_{ij}(x_i,x_j) + \sum_i (q_i - 1) \sum_{x_i} b_i(x_i) E_i(x_i) \tag{10}$$

$$\begin{aligned} E_i(x_i) &= -ln\Phi_i(x_i) \\ E_i(x_i,x_j) &= -Tln\Psi_i(x_i,x_j) - ln\Phi_i(x_i) - ln\Phi_j(x_j) \end{aligned} \tag{11}$$

The Gibbs free energy can be calculated form the formula identical with the equation (2).

$$G\Big(\{x\}\Big) = U\Big(\{x\}\Big) - TS\Big(\{x\}\Big) \tag{12}$$

The most important feature of described algorithm is that the steady state of belief propagation algorithm corresponds to the minimal Gibbs free energy of the system for which the message model has been created what has been proved by Yedidia et al. [15]. Please, notice also the presence of temperature in formulas (10) and (11). All calculations in the mentioned earlier Yedidia's papers were performed for the temperature equal to 1 (in k_B units) so this factor was omitted.

As it will be shown the standard BP algorithm doesn't lead to the sensible results therefore we enhanced our model with the generalization also proposed by Yedidia and coworkers. Main idea of this generalization is the use of spin clusters instead of single spins and the determination of messages on the basis of exchange of information between whole clusters called also regions. Formally the message between regions r and s is given by:

$$m_{rs} = \frac{[\sum_{x_{r/s}} \Psi_{r/s}(x_{r/s}) \prod_{m_{r''s''} \in M(r) \setminus M(s)} m_{r''s''}]}{\prod_{m_{r's'} \in M(r)} m_{r's'}} \tag{13}$$

2 Results

All the calculations presented in this paragraph have been performed for the two-dimensional Ising sample. Although there exist simulational results for Ising model in higher dimensions, D=2 is the highest one for which the exact analytical result has been found. Its author was Lars Onsager [8] who confirmed the existence of spontaneous non-zero magnetization in such a sample in low temperatures even without external magnetic field (in opposition to the solved by Ising 1D model). The exact dependence of magnetization on the temperature is described by the formula.

$$m = \{1 - [sinh(2\beta J)]^{-4}\}^{\frac{1}{8}} \tag{14}$$

where the usual substitution $\beta = 1/k_B T$ is made. During the further calculations, also as usually, the value of Boltzmann constant is for simplicity assumed to be equal

to 1. Finally this assumption means that we express the temperature in the units of exchange integral J.

The evolution system has been performed in the following way. Initially all message values were sampled from the uniform $[0, 1]$ distribution with respect to the normalization conditions. Then the system undergoes the self-consistent update described by the formula (6) for BP or (12) for GBP. In this paper we will use the clusterization in GBP to the two-spin regions. As a unit of update we chose the smallest possible change, ie. the change of single message between exactly two nodes. Every time when we will refer to the notion of step we will think about such a single update of randomly selected pair. The results of magnetization vs. temperature dependence are presented on figure 2.

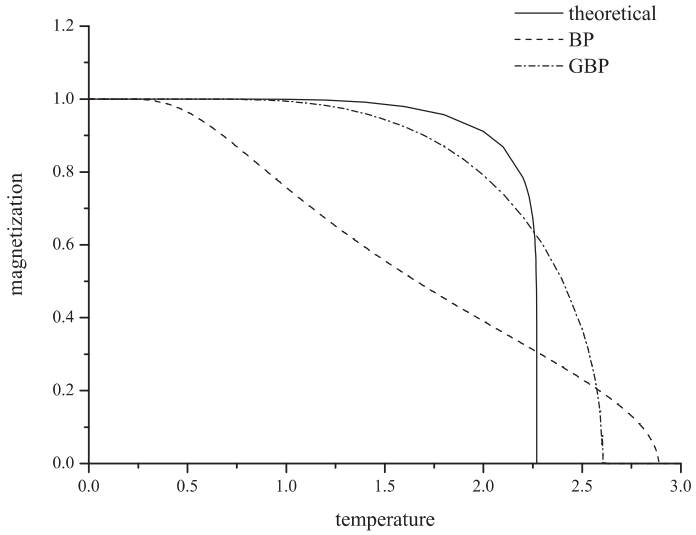

Figure 2. The comparison of magnetization curves for analytical solution, Belief Propagation and Generalized Belief Propagation models.

The solid line is the Onsager solution, the dashed one corresponds to the standard Belief Propagation algorithm, while the dashed-dotted is made for the Generalized BP. The plots confirm the observation made in other calculations using cluster-based approaches. BP as well as GBP overestimates the temperature of phase transition. For Ising model the phase transition temperature (T_c) is this one when the spontaneous magnetization vanishes ($m = 0$). However, besides of T_c comparison we have to look at the shape of $m(T)$ curves. It is obvious that for standard BP it is unacceptable. This simplified model doesn't reflect sufficiently the correlations between different nodes and leads to such an awkward course of magnetization dependence. For the GBP the curve looks closer to the correct one. As well the critical temperature is closer as convexity estimates it better.

The idea to use GBP with pair clusters allowed to compare its results with those obtained from the Evolutionary Algorithm with entropy calculated using pair approximation [6]. The strict theoretical basis is the fact that both approaches are inspired by the seminal paper by Kikuchi [5] in which he proposed such clusterization. The re-

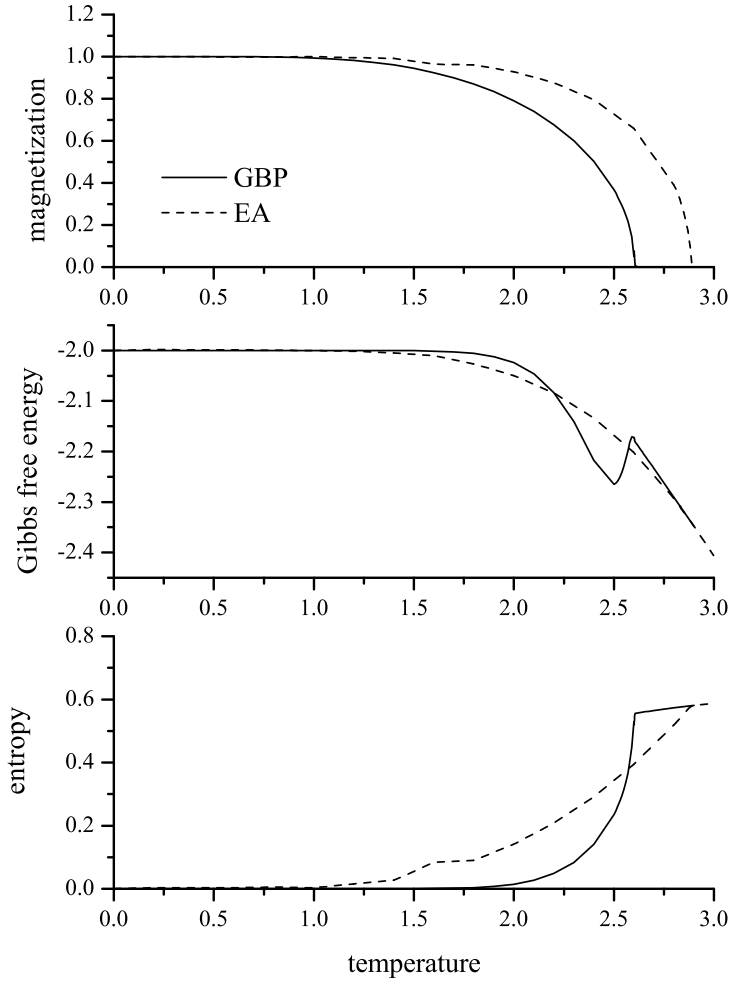

Figure 3. The comparison of magnetization, Gibbs free energy and entropy per one spin for Generalized Belief Propagation models and Evolutionary Algorithm with entropy pair approximation.

sults of comparison for three quantities: magnetization, entropy and Gibbs free energy (all normalized to single spin) are presented on figure 3. The plots can be compared in two ways: qualitatively and quantitatively. From one point of view the results are different, the relative difference between critical temperatures is of order of about 10% with the GBP approximation better than EA. Also the absolute values of entropy differ one from another. Certainly, the course of Gibbs free energy should reflect the entropy differences because it is calculated with formula 12 and the energy values are for both

models strongly connected with total magnetization and similar.

Considering the critical properties of system we cannot limit ourselves to the analysis of T_C. There are important factors which allow to classify the membership of model in the given universality class called critical exponents. Here we will point out only one of them, which is connected to the observed $m(T)$ dependence. This is so called β exponent calculated from the formula

$$m = \Big(\frac{|T - T_C|}{T_C}\Big)^{\beta} \tag{15}$$

The possibility of defining such an exponents come from the fact that in the vicinity of critical temperature many quantities like magnetization, heat capacity , correlation factor behaves according to the power-law dependence.

	T_C	β
analytical	2.269	0.125
EA with pair approximation	2.85	0.63
GBP with two spin clusters	2.607	1.32

Table 1. Critical properties of two-dimensional Ising system using different approaches.

It may be observed that using GBP improves critical temperature when compared with global optimization however makes worse the β coefficient.

Finally let us to pay attention to the efficiency of GBP algorithm. On the figure 4 there are presented the dependences of stabilization times on the temperature and sample size. The stabilization time is expressed in steps and understood as a time when the standard deviation of values of messages leading to the chosen state is smaller than some initially assumed value. In our simulation we accepted the value of 10^{-3}. The interesting observation is that the hardest cases are those close to the critical point. It is in opposition to the classical EA where it was most difficult to obtain correct values for the low temperatures when configuration doesn't differ strongly from the ferromagnetic ordering. The dependence on the size of sample has visible power-law character with the exponent about 3 different for various temperatures.

3 Conclusions

In their paper Yedidia et al. wrote "The success of BP and GBP algorithms is exciting because it means that many different kinds of problems that seemed so difficult to handle can actually be handled using efficient and systematically correctable algorithms.". The results presented in this paper confirm in my opinion this statement. Although Ising model is frequently indicated as an example for GBP algorithm we didn't meet any comprehensive (except of some quite new, initial works [2]) analysis of it as a method of determining the critical properties. Among the advantages one has especially to enlist the time of calculations. From the figure 3 we can easily read that even for computationally hardest region of phase transition the stabilization occurs after on average few millions steps. What is important but not shown here it is the lack of finite size effect. In all simulations or EA optimizations on the lattice with the given edge size the result depends

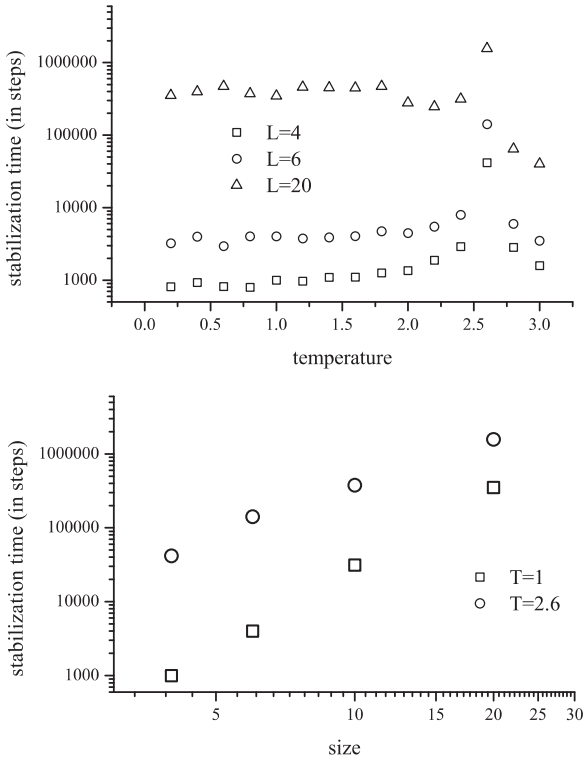

Figure 4. The efficiency of steady state calculation in GBP.

on this size. It can be eg. noticed by comparison of figure 3 and table 2. In the table there is T_C value equal to 2.85 which is the extrapolation to the infinite lattice size while on the plot there is 2.89 obtained directly for $L = 100$.

Certainly GBP has some limitations. The main is that it cannot be used to obtain the ground state of a lot of systems, ie. in zero temperature, The convergence is also the problem of algorithm. The steady state can either be not reached in simulation or the simulation can stopin the state which doesn't correspond to the minimum Gibbs energy and needs the repetition of procedure.

We hope that further improvement of algorithm, especially using larger clusters can significantly change the results and make it possible to use it in real systems analysis.

Bibliography

[1] Dirk Jan Bukman, Guozhong An, and J. M. J. van Leeuwen. Cluster-variation approach to the spin-1/2 xxz model. *Phys. Rev. B*, 43(16):13352–13364, Jun 1991.

[2] L. Casetti and M. Kastner. Partial equivalence of statistical ensembles and kinetic energy. *Physica A*, 384:318, 2007.

[3] T.M. Gwizdałła. Ising model studied using evolutionary approach. *Mod.Phys.Lett.B*, 19:169, 2005.

[4] T.M. Gwizdałła. Magnetic problems as a benchmark for genetic algorithms operators. In *Proceedings of X KAEiOG Conference, Bedlewo*, pages 115–123, 2007.

[5] R. Kikuchi. A theory of cooperative phenomena. *Phys.Rev.*, 6:988–1003, 1981.

[6] H. Meirovitch. Computer simulation study of hysteresis and free energy an the fcc Ising antiferromagnet. *Phys.Rev.B*, 30(5):2866–2874, 1984.

[7] J.M. Mooij and H.J Kappen. On the properties of the Bethe approximation and loopy belief propagation on binary networks. *J.Stat.Mech.*, 5:P11012, 2005.

[8] L. Onsager. Crystal statistics. I. a two-dimensional model with an order-disorder transition. *Phys.Rev.*, 65:117, 1944.

[9] P. Pakzad and V Anantharam. Minimal graphical representation of Kikuchi regions. In *Proceedings of the Allerton Conference on Communication, Control, and Computing. Piscataway*. IEEE, 2002.

[10] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.

[11] A. Szarecka, R.P. White, and H Meirovitch. Absolute entropy and free energy of fluids using the hypothetical scanning method. I. calculation of transition probabilities from local grand canonical partition functions. *J.Chem.Phys.*, 119:12084–12095, 2003.

[12] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *J.Stat.Phys.*, 52:479–487, 1988.

[13] J.S. Yedidia, W.T. Freeman, and Y Weiss. Bethe free energy, Kikuchi approximations and belief propagation algorithms. In *MERL Report*, 2000.

[14] J.S. Yedidia, W.T. Freeman, and Y Weiss. Understanding belief propagation and its generalizations. In G. Lakemeyer and B. Nebel, editors, *Exploring Artificial Intelligence in the New Millenium*, pages 239–269. Morgan Kaufmann, 2003.

[15] J.S. Yedidia, W.T. Freeman, and Y Weiss. Constructing free-energy approximations and generalized belief propagation algorithms. *IEEE Transactions on Information Theory*, 51:2282–2312, 2005.