Dynamics of the Belief Propagation for the Ising Model

G. Siudem[a,b] AND G. Świątek[b,*]

[a]Faculty of Physics, Warsaw University of Technology, Koszykowa 75, PL–00662 Warsaw, Poland

[b]Faculty of Mathematics and Information Science, Warsaw University of Technology,
Koszykowa 75, PL–00662 Warsaw, Poland

The Belief Propagation algorithm is a popular technique of solving inference problems for different graph-like structures. We present a discussion of the dynamics of that algorithm for the Ising model on the square lattice. Our main goal was to describe limit fixed points for that algorithm, which are strictly connected with the marginal probabilities and stationary points of the Bethe Free Energy. Analytical considerations provide an exact analysis of a class of symmetrical points while numerical simulations confirm that for small lattices there are no non-symmetrical points. Notwithstanding the prevalent use of the Belief Propagation as an inference tool we present a sociophysical interpretation of its dynamics. In that case our considerations may be viewed as an investigation of the possible fixed points of the social dynamics.

1. Introduction

The Belief Propagation Algorithm was originally proposed by Judea Pearl in 1982 [1]. It was introduced as an inference technique for the Bayesian trees. Pearl prove that for a graph without cycles the algorithm always converges to its fixed point, and at this fixed point beliefs are equal to marginal probabilities. Pearl's paper started a new direction of research in the statistics and computer science — belief propagation was changed, improved and prepared for different applications (e.g. decoding, computer vision problems and Bayesian networks [2]). Some vision problems, as well as some problems of combinatorial optimalisation may be connected to the application of BP algorithm to Markov Random Fields related to the Gibbs distribution for the Ising model on the square lattice [2–4]. The Ising model was very intensively investigated in statistical physics [5]. Although the main goal of our work is the description of the dynamics of BP algorithm, not the inferring of the properties of the Ising model, we use some well known results for the Ising model as a benchmarks of our calculations. The most important results for our considerations are those connected with the so-called Bethe approximation i.e. solution of the Ising model on the Bethe lattice instead of the standard two-dimensional lattice [6]. It can be shown [2] that the BP algorithm minimizes the Bethe Free Energy i.e. Bethe's approximation of the Gibbs Free Energy. However, unfortunately, for the square lattice there is no global description of the dynamics of BP. The simplest tree case was solved by Pearl [1], but, as he write "if we ignore the existence of loops and permit the nodes to continue communicating with each other as if the network were singly connected, messages may circulate

indefinitely around these loops, and the process may not converge to a stable equilibrium" [2, 7].

The Ising model is also an inspiration for the many sociophysical models [8]. In fact, the BP Algorithm may be interpreted as one of many models describing social dynamics (or opinion dynamics), which is a fast growing branch of science straddling physics, mathematics and sociology [9–15].

Of course beliefs in the BP fixed points have well known physical interpretation (i.e. they are marginal distribution for the Gibbs distribution in the Bethe approximation), although there is no exact results about messages in the BP dynamics for the Ising model. One can calculate beliefs with knowing messages, but different sets of messages may cause with the same value of the beliefs. Our work presents such a results and use above interpretation of the beliefs in the fixed points as an illustration of the BP approach.

Moreover, Belief Propagation for the Ising model was also recently intensively investigated from the statistical physics point of view [16–18], but even despite important results about conditions of convergence and its attraction basins [19] the literature lacks a complete analysis of the dynamics of BP for Ising model for the square lattice. Our paper contributes by considering one of class of fixed points — symmetrical ones.

The article is organised as follows — in Sec. 2 we specify what we mean by the Belief Propagation Algorithm. This sections establishes the terminology. Our main results are demonstrated in Sec. 3 in which we present discussion about the dynamics of the messages in the BP Algorithm. It is well known [2], that every fixed point of the BP algorithm is a stationary point of the Bethe Free Energy. This gives a natural, and well known in the literature, interpretation of the beliefs as the probability distribution for the Bethe's approximation of the Ising model. On the other hand, to the best of our knowledge there are no results concerning the number or stability of

_____

*corresponding author; e-mail: swiatek@mini.pw.edu.pl

fixed points of the message passing, and one can imagine that the beliefs obtained from the BP algorithm may depend on initial messages. Our results establish that there are only two cases — first of them (symmetrical matrices of messages) is somehow typical, and we solve this case exactly. This insight into the BP dynamics allows one to better understand what is BP for the Ising model and may help in further improvement of this algorithm for the Ising model or one of many other applications. Furthermore we claim some hypothesis about the non-symmetrical case and confirm our results by numerical simulations — presented in Sec. 4.

2. Belief Propagation Algorithm for the Ising model

Let us consider the standard Ising model on the $N \times N$ square lattice with periodic boundary conditions. For such a model one can write the Hamiltonian as

$$\mathcal{H}(\sigma, J, h) = -\sum_i \sum_{j \in \mathcal{N}(i)} J\sigma_i\sigma_j - \sum_i h\sigma_i, \qquad (1)$$

where every spin can take two values $\sigma_i = \pm 1$, $i = 1, \ldots, N^2$, h is an external magnetic field and J corresponds to the energy of interaction between nearest neighbours in the square lattice. By $\mathcal{N}(i)$ we mean the set of the four nearest neighbours of the node i in the lattice. With the Hamiltonian given by Eq. (1) one can obtain the partition function

$$Z(J, h, \beta) = \sum_{\sigma \in \{-1, 1\}^{N^2}} \exp\left(-\beta\mathcal{H}(\sigma, J, h)\right), \qquad (2)$$

where $\beta = \frac{1}{k_B T}$. In the further discussion, for brevity we will use only β and call this parameter temperature. Partition function defines the Gibbs distribution on spin configurations — it is simply the probability from the canonical distribution

$$\mathbb{P}_\beta(\sigma) = \frac{\exp\left(-\beta\mathcal{H}(\sigma, J, h)\right)}{Z(J, h, \beta)} = $$
$$\frac{1}{Z} \prod_{i, j \in \mathcal{N}_i} \exp\left(\beta J\sigma_i\sigma_j\right) \prod_i \exp\left(\beta h\sigma_i\right), \qquad (3)$$

where the first product takes place over all pairs of nodes which are connected and the second one over all nodes in the lattice.

As it was mentioned in Sect. 1, the Belief Propagation Algorithm may be used for finding marginal distributions for a very large class of mathematical models. Common for every class are graph-like structures. In our case, the proper structure is a Random Markov Field (for more details see [4]) which is related to the Gibbs distribution given with Eq. (3). The BP algorithm allows one to infer the marginal distribution for a given Random Markov Field i.e. find the Gibbs distribution for the given Ising model in the Bethe approximation [2, 18].

The main idea which is behind the BP algorithm is in fact translating the problem of finding marginal probabilities to the sociophysical model in which every node has opinion about some problem e.g. political view. Every node feels the impact of the external world (e.g. newspapers, television, gossip), which is modelled with the external field h. Furthermore, every node interacts with its neighbours via the coupling J, which means that, for the ferromagnetic case the system prefers nodes which are in agreement. This may be explained as conformism of the modelled population. The most important step in the description of the sociophysical face of the BP algorithm is an introduction of the abstract quantities μ_{ij}, $i, j = 1, \ldots N^2$ which are called *messages*. They may be interpreted as information sent from node j to its neighbour i about node j's opinion about the node i's opinion. It is worth noticing that in that opinion dynamics there are two different mechanisms of interactions. The first of them is more static, connected with the natural (non-)conformism of every node and described with J. The other is concerned with how the message passing represents the impact of one agent on its neighbours. A scheme of this message sending process is presented in Fig. 1. Based on above-mentioned phenomena (external impact, interactions between nodes and message exchange) every node establishes its own beliefs. There are also so-called two-beliefs, which describe pairwise correlations and play a crucial role in the calculus of the BP updating rule. One can repeat this procedure — i.e. allow nodes to change their messages and interact and change the values of the beliefs. If this procedure converges, beliefs are equal to the marginal probabilities.
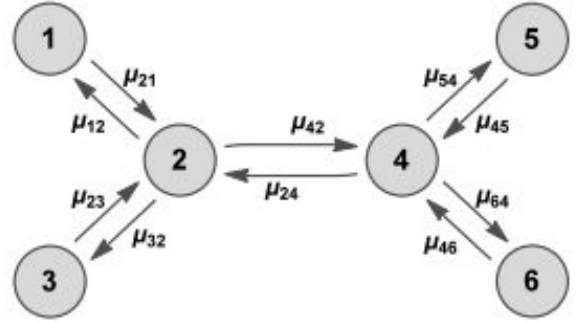

Fig. 1. Scheme of the propagation of messages in the BP algorithm.

Let $\mu = [\mu_{ij}] \in \mathbb{M}^{N^2 \times N^2}(\mathbb{R})$ be the matrix of messages. In general one can not assume that μ is symmetrical, hence $\mu_{ij} \neq \mu_{ji}$. There are also no *a priori* boundary values of the messages. Based on the Hamiltonian given by Eq. (1) we introduce beliefs

$$b_{ij}(\sigma_i, \sigma_j) = k_{ij} \exp\left(\beta\left(h\sigma_i + h\sigma_j + J\sigma_i\sigma_j\right.\right.$$

$$\left.\left. +\sigma_i \sum_{n \in \mathcal{N}_i, n \neq j} \mu_{in} + \sigma_j \sum_{n \in \mathcal{N}_j, n \neq i} \mu_{jn}\right)\right),$$

$$b_i(\sigma_i) = k_i \exp\left(\beta\left(h\sigma_i + \sigma_i \sum_{n \in \mathcal{N}_i} \mu_{in}\right)\right),$$

where k_{ij} and k_i are normalisation constants. If the iterations reach the fixed point and beliefs are equal to the

Gibbs distribution, one can calculate the magnetisation as follows

$$\chi_i = b_i(1) - b_i(-1) = \tanh\left(\beta h + \beta \sum_{n \in \mathcal{N}_i} \mu_{in}\right). \quad (4)$$

We assume that in the limit marginalisation condition [2] will be satisfied

$$b_i(\sigma_i) = \sum_{s_j} b_{ij}(\sigma_i, \sigma_j). \quad (5)$$

This condition is natural — we expect that system is described with the one-beliefs and two-beliefs, so this descriptions must be equivalent. Marginalisation condition in Eq. (5) may be used for the derivation of the updating rule of the Belief Propagation Algorithm. For that purpose let us divide Eq. (5) for $\sigma_i = \pm 1$

$$\frac{b_i(1)}{b_i(-1)} = \frac{\sum_{\sigma_j} b_{ij}(1, \sigma_j)}{\sum_{\sigma_j} b_{ij}(-1, \sigma_j)} \Rightarrow \frac{e^{\beta\mu_{ij}}}{e^{-\beta\mu_{ij}}} =$$

$$\frac{\sum_{\sigma_j} \exp\left(\beta\left(h\sigma_j + J\sigma_j + \sigma_j \sum_{n \in \mathcal{N}_j, n \neq i} \mu_{jn}\right)\right)}{\sum_{\sigma_j} \exp\left(\beta\left(h\sigma_j - J\sigma_j + \sigma_j \sum_{n \in \mathcal{N}_j, n \neq i} \mu_{jn}\right)\right)}, \quad (6)$$

where we use the definitions of b_i and b_{ij}. From the assumption that beliefs in the stationary point of the BP algorithm must satisfy the marginalisation condition one can observe that following function has this property (compare with Eq. (6))

$$\mu'_{ij} = \frac{1}{2\beta} \ln \frac{\cosh\left(\beta\left(h + J + \sum_{n \in \mathcal{N}_j, n \neq i} \mu_{jn}\right)\right)}{\cosh\left(\beta\left(h - J + \sum_{n \in \mathcal{N}_j, n \neq i} \mu_{jn}\right)\right)}. \quad (7)$$

Which means that from the dynamical point of view, the BP algorithm is a problem of iteration of the function (7) parametrised with the three parameters h, J and β over the space of matrices $[\mu_{ij}]$. Because of the construction of the BP algorithm the key features of the dynamics are the existence and uniqueness of the limit points. We present an analytical discussion about limit points of the BP in Sect. 3 and some numerical results in Sect. 4. Our main object of interest are limit points of iterations of BP and such fixed points have to be attracting or at least saddles.

3. General results

Let us consider a fixed point of the BP algorithm μ, i.e. $[\mu_{ij}] = [\mu'_{ij}]$ according to Eq. (7). Furthermore we assume that μ is a symmetric matrix (see Fig. 2). As we will show in this section, the class of symmetrical solutions can be exactly studied and has a natural physical interpretation. In Sect. 4 we present numerical examples which suggest that, for small lattices we tested there were no non-symmetrical limit points. Nonetheless such non-physical solutions exist in higher dimensional problems [20], which is a meaningful drawback of the BP algorithm as a inferring method.
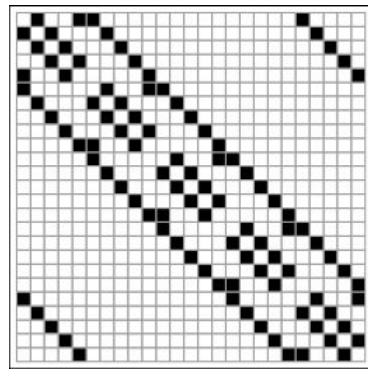

Fig. 2. Visualisation of the one of the fixed points of the BP algorithm for 5×5 lattice. The matrix is symmetrical. Every black square represents the same value of $\mu = 0.588$ ($\beta = 0.4 > \beta_{bif}$, $h = 0$, $J = 1$) and white squares denote zeros. It is worth noticing that because of the fact that $[\mu_{ij}]$ describes messages between every node its size is $N^2 \times N^2$, in this case 25×25.

Let us consider the update formula for the BP algorithm from Eq. (7)

$$\mu'_{ij} = \frac{1}{2\beta} \ln \frac{\cosh\left(\beta\left(h + J + \theta_j - \mu_{ji}\right)\right)}{\cosh\left(\beta\left(h - J + \theta_j - \mu_{ji}\right)\right)} =$$

$$\frac{1}{2\beta} \ln \frac{\cosh\left(\beta\left(\tilde{h} + J - \mu_{ij}\right)\right)}{\cosh\left(\beta\left(\tilde{h} - J - \mu_{ij}\right)\right)},$$

where $\theta_j = \sum_{n \in \mathcal{N}_j} \mu_{jn}$ is the sum of messages coming out from node j. With the substitution $\tilde{h}_j = h + \theta_j$ and the symmetricity of $[\mu_{ij}]$ one can obtain that BP dynamic for μ_{ij} is equivalent to the dynamics of the following function

$$h_{\beta, h, J}(x) = \frac{1}{2\beta} \ln \frac{\cosh\left(\beta\left(h + J - x\right)\right)}{\cosh\left(\beta\left(h - J - x\right)\right)},$$

where $h = \tilde{h}_j$ and $x = \mu_{ij}$. The derivative of $h_{\beta, h, J}(x)$ satisfies the condition $|dh_{\beta, h, J}(x)/dx| < 1$ which means that $h_{\beta, h, J}(x)$ has exactly one fixed point. From the fact that this fixed point depends only on J, β, and $\tilde{h}_j$ one can infer that for every neighbour of node j, limit point's messages are equal $\mu_{kj} = \mu_{k'j}$ for $k, k' \in \mathcal{N}(j)$. This observation with the assumptions of μ being symmetric leads to the conclusion that all non-zero messages are equal and satisfy a fixed point condition for the following f_β function

$$f_\beta(\mu) = \frac{1}{2\beta} \ln \frac{\cosh\left(\beta\left(h + J + 3\mu\right)\right)}{\cosh\left(\beta\left(h - J + 3\mu\right)\right)}, \quad (8)$$

where we ignore h and J in the index for brevity. With the substitutions $A = e^{2\beta h}$, $B = e^{2\beta J}$, $z = e^{2\beta\mu}$ one can simplify the equation of the fixed point of f_β from Eq. (8) to

$$Az^4 - ABz^3 + Bz - 1 = 0, \quad (9)$$

with conditions A, B, $z > 0$. Eq. (9) provides limitations on the number of different fixed point of BP algorithm for the given values of parameters (β, h and J) — there are at most four such solutions. Eq. (9) may be seen as

a solution of the problem — one can solve this algebraic equation (e.g. in Wolfram Mathematica) and find possible values of fixed points of BP algorithm. We skip this calculation, as a technical one and present above discussion about properties of those limit points. We split this into two parts — ferromagnetic ($J > 0$) and antiferromagnetic ($J < 0$).

Firstly, let us assume that $J > 0$ and $h = 0$, which means, that f_β is a monotone increasing function. For sufficiently small values of β one can see that there is only one stable fixed point $x = 0$. This situation changes at the bifurcation point, where $x = 0$ becomes unstable, which means that

$$\frac{df_\beta(x)}{dx} = \frac{3}{2}\left(\tanh\left(\beta(h + J + 3x)\right)\right.$$
$$\left. - \tanh\left(\beta(h - J + 3x)\right)\right),$$

satisfies the condition $df_\beta/dx|_{x=0} = 1$, which corresponds to the following value of bifurcation temperature

$$\beta_{bif} = \frac{\mathrm{artgh}(\frac{1}{3})}{J} = \frac{\ln 2}{2J}. \tag{10}$$

For $\beta < \beta_{bif}$ there is only one stable fixed point at $x = 0$ and for $\beta > \beta_{bif}$ two stable fixed points while $x = 0$ has lost stability, but is still a fixed point. There were no another bifurcations, which one can infer from the plot of function f_β, which is presented in Fig. 3.
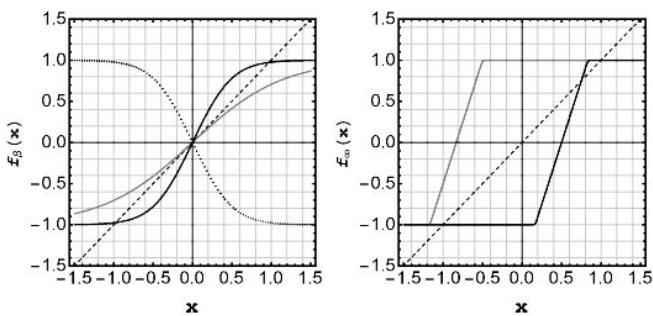

Fig. 3. Plots of functions f_β (left, see Eq. (8)) and f_∞ (right, see Eq. (11)). On the left plot every curve is the plot of function for $h = 0$ and $\beta = 0.3 < \beta_{bif}$, $J = 1$ (black solid), $\beta = 0.8 > \beta_{bif}$, $J = 1$ (gray solid), $\beta = 0.8$, $J = -1$ (black dotted). On the right plot every curve is drawn for $J = 1$, the grey solid line corresponds to $h = 2.5$ and the black solid line to $h = -1.5$.

Introduction of a non-zero external field $h \neq 0$ does not change the dynamic dramatically, but for the brevity we describe temperature of bifurcation at zero external field as β_{bif} and for non-zero field β_b. For sufficiently small values of $|h|$ there is no change in the dynamic — below the bifurcation point (in that case β_b is higher for β_{bif} given by Eq. (10)) there is only one fixed point $x = -h/3$ and for higher temperatures two stable and one unstable points appear. Sufficiently large external field destroys the bifurcation — for that situation there is only one fixed point for every values of β. Those observations may be proven as follows — let us consider f_∞, which is the limit as $\beta \to \infty$ of Eq. (8)

$$f_\infty(x) = \begin{cases} h + 3x & x \in \left(-\frac{1}{3}(J+h), \frac{1}{3}(J-h)\right), \\ J\,\mathrm{sgn}(x) & \text{otherwise.} \end{cases} \tag{11}$$

$f_\beta(x)$ converges monotonically to $f_\infty(x)$ for every x, so if any bifurcation occurs we may observe more than one fixed point of f_∞. The analysis of the graph (Fig. 3) of function f_∞ convinces one that the necessary and sufficient condition for the bifurcation to occur is $|h| < 2J$. For a stronger external field there is no bifurcation and there is only one stable fixed point.

The antiferromagnetic case ($J < 0$) situation is much easier for description. Firstly, let us notice that the change of the sign of J causes changes of sign in $f_\beta(x)$, which can be easy deduce from Eq. (8). It implies that for the antiferromagnetic case the function f_β is monotonically decreasing, which means that there is exactly one stable fixed point below the bifurcation point. For the temperature above the bifurcation the situation strongly depends on the value of external field. What is surprising — the role of external field is opposite than in ferromagnetic case. For higher values of $|h|$ there is exactly one stable point, but for $|h| < 2J$ there is no fixed point — the dynamics leads to stable orbits of order 2, but these are unacceptable for the BP algorithm, as it was mentioned in Sect. 1.

This discussion, which is shortly presented in Table may be summed up as follows: if one considers only symmetrical fixed points of the BP algorithm, then there are exact constrains for the number of stable fixed points, which depends on sign of values of β, h and J. The bifurcation, if it occurs, results in two new stable points in the place of one for smaller values of β ($J > 0$), or the fixed point vanishing ($J < 0$). Of course the vanishing does not mean that there is no fixed point of BP — there are only no symmetrical ones.

TABLE

The number of symmetrical fixed points of BP, β_b denotes bifurcation temperature for the non-zero external field, for $h = 0$, $\beta_b = \beta_{bif}$ (see Eq. (10)).

	$|h| < 2J$	$|h| > 2J$
$\beta < \beta_b$	$1_{ferro}\ 1_{anti}$	$1_{ferro}\ 1_{anti}$
$\beta > \beta_b$	$2_{ferro}\ 0_{anti}$	$1_{ferro}\ 1_{anti}$

4. Numerical simulations

We investigate numerically BP algorithm on the $N \times N$ lattices, where $N \leqslant 10$. For those lattices the only fixed points into which BP algorithm converge were symmetrical matrices, which means that results from Sect. 3 are applicable. In Fig. 2 we present a visualisation of the matrix, which was the fixed point of BP, and as it can be seen — it is a symmetrical matrix. Of course sometimes the BP algorithm gets stuck in a periodic orbit and in that situation we repeat our computations.

The next step in our numerical simulation was checking the correctness of magnetisations calculated from the BP algorithm. BP converges to its fixed point, which

allows us to calculate beliefs from Eq. (5) and, finally, the magnetisations (4). We compare this simulation results with well known calculations of magnetisation for Bethe lattices [6]. As is presented in Fig. 4, there is a perfect correspondence between numerical and analytical results.
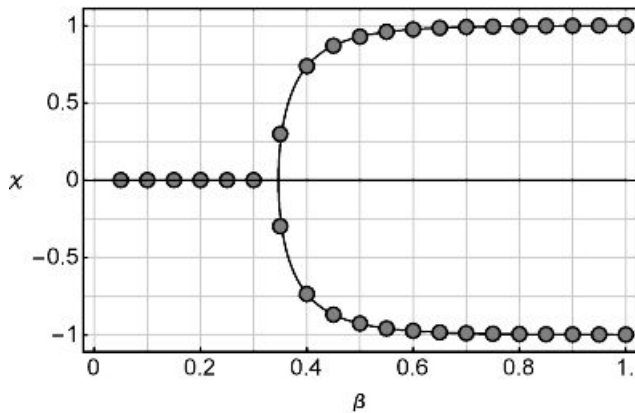

Fig. 4. Comparison between BP results (dots from simulations and Eq. (4)) and Bethe approximation (solid line, for details see [6]) for the magnetisation. We assume that $J = 1$ and $h = 0$. Note that β_{bif} from Eq. (10) is equal to the critical temperature for the Bethe approximation.
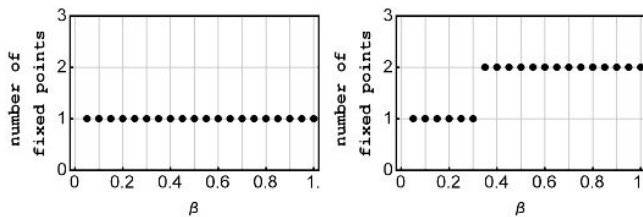

Fig. 5. Obtained simulations scenarios. No (left) or one (right) bifurcation.

The last numerical test was the examination of the phase space of BP. We run the algorithm for the random values of the external field and $J = 1$ and count different fixed points. The plot of the number of such points versus β is presented in Fig. 5. As one can see there are only two scenarios — one includes a single bifurcation in which one fixed point changes into two different fixed points. The second of them shows no bifurcation — for every value of β there is always one stationary stable point.

5. Concluding remarks

We present exact, analytical discussion of the dynamics of the BP algorithm converging to a symmetrical fixed point. Numerical simulations suggest that for lattices we considered such fixed points are the only ones possible. Notably for larger lattices there exist non-symmetrical fixed points [20]. Notwithstanding the application of the BP algorithm for solving graphical models, one can see

that algorithm as a sociophysical model of opinion dynamic. From that point of view we have investigated of properties of a global compromise state.

References

[1] J. Pearl, *Proceedings of the Second National Conference on Artificial Intelligence*, 1982, p. 133.

[2] J.S. Yedidia, W.T. Freeman, Y. Weiss, *Exploring Artificial Intelligence in the New Millennium*, Eds. G. Lakemeyer, B. Nebel, Morgan Kaufmann Publishers, San Francisco 2003, Ch. 8, p. 239.

[3] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods. A Mathematical Introduction*, Springer-Verlag, New York 1991.

[4] P. Brémaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulation, And Queues*, Springer-Verlag, New York 1999.

[5] B.M. McCoy, *Advanced Statistical Mechanics*, Oxford Univ. Press, Oxford 2010.

[6] R.J. Baxter *Exactly Solved Models in Statistical Mechanics*, Academic Press, London 1989.

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

[8] C. Castellano, S. Fortunato, V. Loreto, *Rev. Mod. Phys.* **81**, 591 (2009).

[9] M.J. Krawczyk, K. Malarz, R. Korff, K. Kułakowski, *Lect. Notes Artif. Int.* **6421**, 90 (2010).

[10] J. Sienkiewicz, G. Siudem, J.A. Hołyst, *Phys. Rev. E* **82**, 057101, (2010).

[11] A. Czaplicka, A. Chmiel, J.A. Hołyst, *Acta Phys. Pol. A* **117**, 688 (2010).

[12] K. Malarz, K. Kułakowski, *Acta Phys. Pol. A* **121**, B86 (2012).

[13] P. Nyczka, K. Sznajd-Weron, J. Cislo, *Phys. Rev. E* **86**, 011105 (2012).

[14] P. Nyczka, K. Sznajd-Weron, *J. Stat. Phys.,* **151**, 174 (2013).

[15] P. Sobkowicz, *J. Artif. Soc. Soc. Simulat.* **12**, 11 (2009).

[16] S. Yoon, A.V. Goltsev, S.N. Dorogovtsev, J.F.F. Mendes, *Phys. Rev. E* **84**, 041144, (2011).

[17] J. Ohkubo, M. Yasuda, K. Tanaka, *Phys. Rev. E* **72**, 046135 (2005).

[18] S. Dorogovtsev, A. Goltsev, J. Mendes, *Rev. Mod. Phys.* **80**, 1275 (2008).

[19] J.M. Mooij, H.J. Kappen, *J. Stat. Mech. Theor. Exp.* **11**, P11012 (2005).

[20] G. Siudem, G. Świątek, *Diagonal Stationary Points of Bethe's Free Energy*, in preparation.