Chapter 4

The Random Field Ising Model on Random Graphs

I think that I shall never see
A poem lovely as a tree.

Joyce Kilmer - 1913

We shall continue our exploration of Boltzmann-Gibbs distribution

$$\mathbb{P}_{N,\beta}\left(\boldsymbol{S}=\boldsymbol{s}\right)=\frac{e^{-\beta\mathcal{H}(\boldsymbol{s})}}{Z_N(\beta)}. \tag{4.1}$$

For the random field Ising model, but we shall now look at a more complex, and more interesting, topology. We assume the existence of a graph $\mathcal{G}_{ij}$ that connects some of the nodes, such that $\mathcal{G}_{ij}=1$ if i and j are connected, and 0 otherwise. So far, the last chapters dealt only with the case of fully connected graphs $\mathcal{G}_{ij}=1\forall\, i,j$. Now our Hamiltonian can be written as:

$$\mathcal{H}_{N,J,\{h\},\mathcal{G}}(\boldsymbol{s})=-J\sum_{i,j\in\mathcal{G}}s_is_j-\sum_{i=1}^{N}h_iS_i \tag{4.2}$$

where J is a scalar coupling constant.

4.1 The roots of all Cavity arguments

Let us assume we have N spins $i=1,\ldots,N$, that are all isolated. In this case, their probability distribution is simple enough, we have $m_i=\tanh\beta h_i$. Imagine now we are connecting these N spins to a new spin S_0, in the spirit of the cavity method. Of course, the m_i are now changed! Let us thus refer to the old values of m_i as the "cavity ones" and write:

$$m_i^c\equiv\tanh\beta h_i\,. \tag{4.3}$$

With this definition, note that $(1+s_im_i^c)/2=e^{\beta h_is_i}/cosh(\beta h_i)$. Our main question of interest is: can we write the magnetization of the new spin m_0 as a function of the $\{m_i^c\}$?

Let us try! Clearly

$$m_0 = \frac{\sum_{s_0,\{s\}} S_0 e^{\beta s_0 h_0 + \sum_i \beta J s_i s_0 + \beta s_i h_i}}{\sum_{s_0,\{s\}} S_0 e^{\beta s_0 h_0 + \sum_i \beta J s_i s_0 + \beta s_i h_i}} = \frac{\sum_{s_0} S_0 e^{\beta s_0 h_0} \prod_i \sum_{s_i} e^{\beta J s_i s_0} \frac{1+s_i m_i^c}{2}}{\sum_{s_0} e^{\beta s_0 h_0} \prod_i \sum_{s_i} e^{\beta J s_i s_0} \frac{1+s_i m_i^c}{2}} \tag{4.4}$$

$$= \frac{X^+ - X^-}{X^+ X^-} \text{ with } X_s = e^{\beta s h_0} \prod_i \sum_{s_i} e^{\beta J s_i s} \frac{1 + s_i m_i^c}{2} \tag{4.5}$$

Using now the relation $\text{atanh} x = \frac{1}{2} \log \frac{1+x}{1-x}$, and applying the atanh on both side, we reach,

$$\text{atanh}(m_0) = \frac{1}{2} \log \frac{1 + \frac{X^+ - X^-}{X^+ + X^-}}{1 - \frac{X^+ - X^-}{X^+ + X^-}} = \frac{1}{2} \log \frac{X_+}{X_-} \tag{4.6}$$

$$= \beta h_0 + \sum_i \frac{1}{2} \log \frac{e^{\beta J}(1 + m_i^c) + e^{-\beta J}(1 - m_i^c)}{e^{-\beta J}(1 + m_i^c) + e^{\beta J}(1 - m_i^c)} \tag{4.7}$$

$$= \beta h_0 + \sum_i \frac{1}{2} \log \frac{\cosh(\beta J) + m_i^c \sinh(\beta J)}{\cosh(\beta J) - m_i^c \sinh(\beta J)} \tag{4.8}$$

$$= \beta h_0 + \sum_i \frac{1}{2} \log \frac{1 + m_i^c \tanh(\beta J)}{1 - m_i^c \tanh(\beta J)} \tag{4.9}$$

$$= \beta h_0 + \sum_i \text{atanh} \left(m_i^c \tanh(\beta J) \right) \tag{4.10}$$

We can thus finally write the magnetization of the new spins as the function of the spin in the "old" system as

$$m_0 = \tanh \left(\beta h_0 + \sum_i \text{atanh} \left(m_i^c \tanh(\beta J) \right) \right) \tag{4.11}$$

4.2 Exact recursion on a tree

It is quite simple to repeat the same argument iteratively! Consider a tree, with initial conditions then we can write, at any point of the graph

$$m_{i \to j} = \tanh \left(\beta h_i + \sum_{k \in \partial_i \neq j} \text{atanh} \left(m_{k \to i} \tanh(\beta J) \right) \right) \tag{4.12}$$

What is we want to know the true marginal? This is easy, we just write

$$m_i = \tanh \left(\beta h_i + \sum_{k \in \partial_i} \text{atanh} \left(m_{k \to i} \tanh(\beta J) \right) \right) \tag{4.13}$$

This is the root of the so-called Belief propagation approach. We solve the problem for the cavity marginals $m_{i \to j}$ that has a convenient interpretation as a message passing problem. Once we know them all, we can compute the true marginal!

This is a method that first appears in statistical physics with Bethe and Peierls as an approximation of regular lattice. Indeed, we could iterate this method on a large infinite tree of

connectivity, say $c = 2d$ to approximate a lattice of dimension d. Consider for instance the case in zero field, in this case, the magnetization at distance ℓ from the leaves follow

$$m^{\ell+1} = \tanh\Big((c-1)\text{atanh}\Big(m^{\ell}\tanh(\beta J)\Big)\Big) \tag{4.14}$$

We can look for a fixed point of this equation, and check for which value of the (inverse) temperature a non-zero value for the magnetization is possible at the fixed point. This is the same phenomenon as in the Ising model, with just a slightly more complicated fixed point. By ploting this equation, one realizes that, assuming $d = 2c$, this happens at $\beta^{\rm BP} = \infty$ for $d = 1$, $\beta^{\rm BP} = 0.346$ for $d = 2$, $\beta^{\rm BP} = 0.203$ for $d = 3$, $\beta^{\rm BP} = 0.144$ for $d = 4$, and $\beta^{\rm BP} = 0.112$ for $d = 5$. If we compare these numbers to the actual transition on a real hyper-cubic lattice, we find $\beta^{\rm lattice} = \infty$ for $d = 1$, $\beta^{\rm lattice} = 0.44$ for $d = 2$, $\beta^{\rm lattice} = 0.221$ for $d = 3$, $\beta^{\rm lattice} = 0.149$ for $d = 4$ and $\beta^{\rm lattice} = 0.114$ for $d = 5$. Not so bad, and in fact we see that the predictions become exact as d grows! This approach is quite a good one to estimate a critical temperature (in fact, one can show that it gives a rigorous upper bound on the ferromagnetic transition, in any topology!).

4.2.1 Belief propagation on trees for pairwise models

The iterative approach we just discussed can be made completely generic on a tree graph! The example that we have been considering so far reads, in full generality

$$\mathcal{H} = -\sum_{(ij)\in E} J_{ij}S_iS_j - \sum_i h_iS_i\,. \tag{4.15}$$

This is an instance of a very generic type of model, those with pairwise interaction where the probability of each configuration is given by

$$P\left((S)\right) = \frac{1}{Z}\left[\prod_i \psi_i(S_i)\prod_{(ij)\in E}\psi_{ij}(S_i,S_j)\right] \tag{4.16}$$

$$Z = \sum_{\{S_{i=1,\ldots,N}\}}\left[\prod_i \psi_i(S_i)\prod_{(ij)\in E}\psi_{ij}(S_i,S_j)\right] \tag{4.17}$$

and the connection is clear once we define $\psi_{ij}(S_i,S_j) = \exp(\beta J_{ij}S_iS_j)$ and $\psi_i(S_i) = \exp(\beta h_iS_i)$.

Similarly, as what we did for the RFIM, to compute Z on a tree, the trick is to consider instead the variable $Z_{i\to j}(S_i)$, for each two adjacent sites i and j, defined as the *partial* partition function for the sub-tree rooted at i, when excluding the branch directed towards j, with a fixed value S_i of the spin variable on the site i. We also need to introduce $Z_i(S_i)$, the partition function of the entire complete tree when, again, the variable i is fixed to a value S_i. On a tree, these intermediate variables can be exactly computed according to the following recursion

$$Z_{i\to j}(S_i) = \psi_i(S_i)\prod_{k\in\partial i\setminus j}\left(\sum_{S_k} Z_{k\to i}(S_k)\psi_{ik}(S_i,S_k)\right)\,, \tag{4.18}$$

$$Z_i(S_i) = \psi_i(S_i)\prod_{j\in\partial i}\left(\sum_{S_j} Z_{j\to i}(S_j)\psi_{ij}(S_i,S_j)\right)\,, \tag{4.19}$$

where $\partial i \neq j$ denotes the set of all the neighbors of i, except spin j. In order to write these equations, the only assumption that has been made was that, for all $k \neq k' \in \partial i \setminus j$, $Z_{k\to i}(S_k)$ and $Z_{k'\to i}(S_{k'})$ are independent. On a tree, this is obviously true: since there are no loops, the sites k and k' are connected only through i and we have "cut" this interaction when considering the partial quantities. This recursion is very similar, in spirit, to the standard transfer matrix method for a one-dimensional chain.

In practice, however, it turns out that working with partition functions (that is, numbers that can be exponentially large in the system size) is somehow impractical, and we can thus normalize eq. (4.18) and rewrite these recursions in terms of probabilities. Denoting $\eta_{i\to j}(S_i)$ as the marginal probability distribution of the variable S_i when the edge (ij) has been removed, we have

$$\eta_{i\to j}(S_i) = \frac{Z_{i\to j}(S_i)}{\sum_{S'} Z_{i\to j}(S')}\,, \qquad \eta_i(S_i) = \frac{Z_i(S_i)}{\sum_{S'} Z_i(S')}\,. \tag{4.20}$$

So that the recursions (4.18-4.19) now read

$$\eta_{i\to j}(S_i) = \frac{\psi_i(S_i)}{z_{i\to j}} \prod_{k\in\partial i\setminus j} \left(\sum_{S_k} \eta_{k\to i}(S_k)\psi_{ik}(S_i,S_k) \right)\,, \tag{4.21}$$

$$\eta_i(S_i) = \frac{\psi_i(S_i)}{z_i} \prod_{j\in\partial i} \left(\sum_{S_j} \eta_{j\to i}(S_j)\psi_{ij}(S_i,S_j) \right)\,, \tag{4.22}$$

where the $z_{i\to j}$ and z_i are normalization constants defined by:

$$z_{i\to j} = \sum_{S_i} \psi_i(S_i) \prod_{k\in\partial i\setminus j} \left(\sum_{S_k} \eta_{k\to i}(S_k)\psi_{ik}(S_i,S_k) \right)\,, \tag{4.23}$$

$$z_i = \sum_{S_i} \psi_i(S_i) \prod_{j\in\partial i} \left(\sum_{S_j} \eta_{j\to i}(S_j)\psi_{ij}(S_i,S_j) \right)\,. \tag{4.24}$$

The iterative equations (4.21,4.22), and their normalization (4.23,4.24), are called the belief propagation equations. Indeed, since $\eta_{i\to j}(S_i)$ is the distribution of the variable S_i when the edge to variable j is absent, it is convenient to interpret it as the "belief" of the probability of S_i in absence of j. It is also called a "cavity" probability since it is derived by removing one node from the graph. The belief propagation equations are used to define the belief propagation algorithm

- Initialize the cavity messages (or "beliefs") $\eta_{i\to j}(S_i)$ randomly or following a prior information $\psi_i(S_i)$ if we have one.
- Update the messages in a random order following the belief propagation recursion eq. (4.21,4.22) until their convergence to their fixed point.
- After convergence, use the beliefs to compute the complete marginal probability distribution $\eta_i(S_i)$ for each variable. This is the belief propagation estimate on the marginal probability distribution for variable i.

Using the resulting marginal distributions, one can compute, for instance, the equilibrium local magnetization via $m_i = \langle S_i \rangle = \sum_{S_i} \eta_i(S_i) S_i$, or basically any other local quantity of interest.

At this point, since we have switched from partial partition sums to partial marginals, the astute reader could complain that it seems that we have lost out prime objective: the computation of the partition function. Fortunately, one can compute it from the knowledge of the marginal distributions. To do so, it is first useful to define the following quantity for every edge (ij):

$$z_{ij} = \sum_{S_i, S_j} \eta_{j\to i}(S_j)\eta_{i\to j}(S_i)\psi_{ij}(S_i, S_j) = \frac{z_j}{z_{j\to i}} = \frac{z_i}{z_{i\to j}} \,, \tag{4.25}$$

where the last two equalities are obtained by plugging (4.21) into the first equality and realizing that it almost gives eq. (4.24). Using again eqs. (4.21-4.24), we obtain

$$\begin{aligned} z_i &= \sum_{S_i} \psi_i(S_i) \prod_{j\in\partial i} \left(\sum_{S_j} \eta_{j\to i}(S_j)\psi_{ij}(S_i, S_j) \right) \\ &= \sum_{S_i} \psi_i(S_i) \prod_{j\in\partial i} \left(\sum_{S_j} \frac{Z_{j\to i}(S_j)}{\sum_{S'} Z_{j\to i}(S')} \psi_{ij}(S_i, S_j) \right) = \frac{\sum_{S_i} Z_i(S_i)}{\prod_{j\in\partial i} \sum_{S_j} Z_{j\to i}(S_j)} \,, \end{aligned} \tag{4.26}$$

and along the same steps

$$z_{j\to i} = \frac{\sum_{S_j} Z_{j\to i}(S_j)}{\prod_{k\in\partial j\setminus i} \sum_{S_k} Z_{k\to j}(S_k)} \,. \tag{4.27}$$

For any spin S_i, the total partition function can be obtained using $Z = \sum_{S_i} Z_i(S_i)$. We can thus start from an arbitrary spin i

$$Z = \sum_{S_i} Z_i(S_i) = z_i \prod_{j\in\partial i} \left(\sum_{S_j} Z_{j\to i}(S_j) \right) = z_i \prod_{j\in\partial i} \left(z_{j\to i} \prod_{k\in\partial j\setminus i} \sum_{S_k} Z_{k\to j}(S_k) \right) \,, \tag{4.28}$$

and we continue to iterate this relation until we reach the leaves of the tree. Using Eq. (4.25), we obtain

$$Z = z_i \prod_{j\in\partial i} \left(z_{j\to i} \prod_{k\in\partial j\setminus i} z_{k\to j} \cdots \right) = z_i \prod_{j\in\partial i} \left(\frac{z_j}{z_{ij}} \prod_{k\in\partial j\setminus i} \frac{z_k}{z_{jk}} \cdots \right) = \frac{\prod_i z_i}{\prod_{(ij)} z_{ij}} \,. \tag{4.29}$$

We thus obtain the expression of the free energy in a convenient form, that can be computed directly from the knowledge of the cavity messages, often called the *Bethe free energy*

$$\begin{aligned} &fN = -T \log Z = \sum_i f_i - \sum_{(ij)} f_{ij} \,, \\ &f_i = -T \log z_i \,, \quad f_{ij} = -T \log z_{ij} \,, \end{aligned} \tag{4.30}$$

where f_i is a "site term" coming from the normalization of the marginal distribution of site i, and is related to the change in Z when the site i (and the corresponding edges) is added to the system. f_{ij} is an "edge" term that can be interpreted as the change in Z when the edge

(ij) is added. This provides a convenient interpretation of the Bethe free energy eq. (4.30): it is the sum of the free energy f_i for all sites but, since we have counted each edge twice we correct this by subtracting f_{ij}.

We have now entirely solved the problem on a tree. There is, however, nothing that prevents us from applying the same strategy on any graph. Indeed the algorithm we have described is well defined on any graph, but we are not assured that it gives exact results nor that it will converge. Using these equations on graphs with loops is sometimes referred to as *loopy belief propagation* in Bayesian inference literature.

One may wonder if there is a connection between the BP approach and the variational one. We may even wonder if this could be simply the same as using our bound with a better approach than the naive mean field one! Sadly, the answer is no! We cannot prove in general that the BP free entropy is a lower bound on any graph: indeed there are examples where it is larger than $logZ$, and some where it is lower.

Two important remarks:

- There IS however a connection between the variational approach and the BP one. If one writes the variational approach and uses the following parametrization for the guess:

$$\mathcal{Q}(S) = \frac{\prod_{ij} b_{ij}(S_i, S_j)}{\prod_i b_i(S)^{c_i - 1}} \tag{4.31}$$

  then it is possible to show that optimizing on the function b_{ij} and b_i one finds the BP free entropy. The sad news, however, is that $\mathcal{Q}(S)$ does not really correspond to a true probability density, it is not always normalizable, so one cannot apply the variational bound.

- In the case of ferromagnetic models, or more exactly, on attractive potential $\Psi_i j$, and only in this case then it can be shown rigorously, that BP *DOES* give a lower bound on the free entropy, on any graph. For such models (thus including the RFIM!) it is thus effectively equivalent to a variational approach. In fact, it can be further shown that in the limit of zero temperature, BP finds the ground state of the RFIM on any graph (through a mapping to linear programming).

4.3 Cavity on random graphs

4.3.1 Random graphs

We shall discuss now the basic properties of sparse *Erdős-Rényi* (ER) random graphs.

An ER random graph is taken uniformly at random from the ensemble, denoted $\mathcal{G}(N, M)$, of graphs that have N vertices and M edges. To create such a graph, one has simply to add random M edges to an empty graph. Alternatively, one can also define the so called $\mathcal{G}(N, p)$ ensemble where an edge exists independently for each pair of nodes with a given probability $0 < c/N < 1$. The two ensembles are asymptotically equivalent in the large N limit, when $M = c(N-1)/2$. The constant c is called the average degree. We denote by c_i the degree

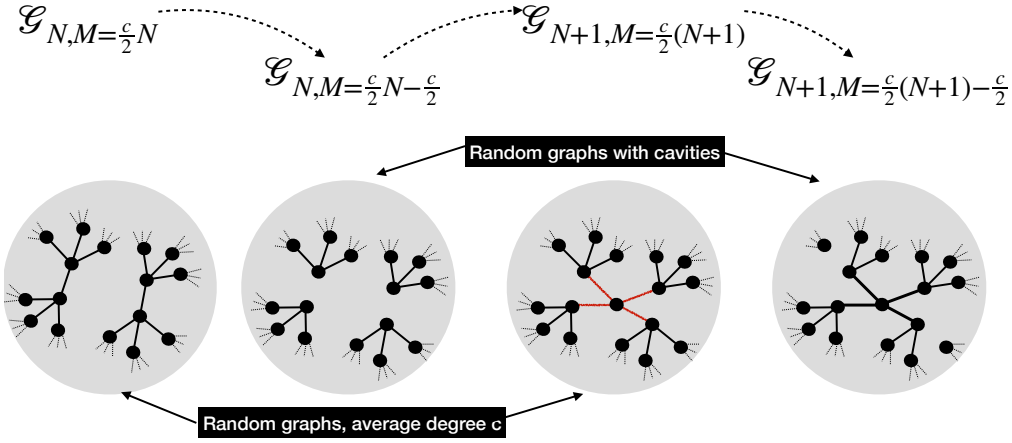

Figure 4.3.1: Iterative construction of a random graph for the Cavity method. The average degree of the graph is c.

of a node i, i.e. the number of nodes to which i is connected. The degrees are distributed according to Poisson distribution, with average c.

Alternatively, one can also construct the so-called *regular* random graphs from the ensemble $\mathcal{R}(N, c)$ with N vertices but where the degree of each vertex is fixed to be exactly c. This means that the number of edges is also fixed to $M = cN/2$.

At the core of the cavity method is the fact that such random graphs locally look like trees, i.e. there are no short cycles going trough a typical node. The key point is thus that, in this limit, such random graphs can be considered locally as trees. The intuitive argument for this result is the following one: starting from a random site, and moving following the edges, in ℓ steps c^ℓ sites will be reached. In order to have a loop, we thus need $c^\ell \sim N$ to be able to come back on the initial site, and this gives $\ell \sim \log(N)$.

4.3.2 Cavity method

Let us now apply our beloved cavity method on a random graph with N links and on average $M = cN/2$ links, this is called an Erdos-Renyi random graph. Let us try to see how this method leads to free entropies using the same telescopic-sum trick as before. However, this should be done with caution. If we apply our usual Cesaro-trick naively and write

$$Z_{N,M} = \frac{Z_{N,M}}{Z_{N-1,M-m_1}} \frac{Z_{N-1,M-m1}}{Z_{N-2,M-m_1-m_2}} \dots Z_1$$

and ask how we choose the m at each time, we encounter a problem. We want the new spin to have on average c neighbors, so we must add a Poissonian variable m with mean c at each time so that the new guy has the correct numbers of neighbors. This however, add a $+1$ contribution to N spins in the previous graphs, so that, on average, we went from $M = cN$ to $M' = c/2N + c$ while we actually wanted to have $M' = c(N + 1)/2$. The difference is $\Delta M = c/2$ so we need to construct our sequence of graph such that we remove $m/2$ link

on average every time we add one spin connected to m previous spins. Therefore, we write instead

$$Z_{N,M} = \frac{Z_{N,M}}{Z_{N-1,M-m_1}} \frac{Z_{N-1,M-m_1}}{Z_{N-1,M-m_1+m_2/2}} \frac{Z_{N-1,M-m_1+m_2/2}}{Z_{N-2,M-m_1+m_1'/2-m_2}} \ldots\ldots Z_1 \tag{4.32}$$

Concretely, this means that we must apply the Cesaro theorem to

$$A_N = \log \frac{Z_{N,M=N\frac{c}{2}}}{Z_{N-1,M-c}} \frac{Z_{N-1,M-c}}{Z_{N-1,M-c+\frac{c}{2}}} = \log \frac{Z_{N,M=N\frac{c}{2}}}{Z_{N-1,M-c}} - \log \frac{Z_{N,M=N\frac{c}{2}}}{Z_{N-1,M-c+\frac{c}{2}}} . \tag{4.33}$$

And we thus obtain

$$\lim_{N,M\to\infty} \mathbb{E}\frac{1}{N} \log Z_N = \lim_{N\to\infty} \mathbb{E} A_N = \lim_{N\to\infty} \mathbb{E}\Phi_N^{(\text{site})} - \lim_{N\to\infty} \mathbb{E}\Phi_N^{(\text{link})} \tag{4.34}$$

so that we have two terms to compute: the shift in free entropy when one add one spin to a graph connecting it to c spins with c cavities, and one when we add $c/2$ links to a graph connecting it to 2 spins with cavities. This leads to

$$\begin{aligned} \Phi_N^{(\text{site})} &= \mathbb{E}\log\langle \sum_{s_0} e^{\beta h_0 s_0 + \beta J \sum_{i\in\partial_0} s_0 s_i} \rangle_{N-1,M-c} = \mathbb{E}\log\langle 2\cosh\left(\beta h_0 + \beta J \sum_{i\in\partial_0} s_i\right)\rangle_{N-1,M-c} \\ \Phi_N^{(\text{link})} &= \mathbb{E}\log\langle \prod_{ij)} e^{\beta J J s_i s_j} \rangle_{N,M} = \frac{c}{2}\mathbb{E}\log\langle e^{\beta J J s_i s_j}\rangle_{N-1,M-c} \end{aligned} \tag{4.35}$$

Interestingly, both these equations depend on the distribution of the cavity magnetizations in the graph so we can write

$$\Phi^{(\text{site})} = \mathbb{E}\int dd P_{\text{e}}(d) \int \prod_{i=1}^{d} dm_i Q^c(\{m_i\}) \log\langle 2\cosh\left(\beta h_0 + \beta J \sum_{i\in\partial_0} s_i\right)\rangle_{\{m_i\}} \tag{4.36}$$

$$\Phi^{(\text{link})} = \mathbb{E}\frac{c}{2}\int dm_1 Q(m_1) dm_2 Q(m_2) \log\langle e^{\beta J J s_i s_j}\rangle_{\{m_1,m_2\}} \tag{4.37}$$

At the level of rigor used in physics, these formulas can be simplified! First, we can first make the assumption that the distribution of cavity field converges to a limit distribution, independent of the disorder. Secondly, and more importantly, a crucial point is now to *assume the distribution of these cavity marginals factorizes*! This makes sense: the different $\{m_i\}$ are all far from each other in the graph; If we are not exactly at a critical point (a phase transition) then the correlations are not infinite. This is case, we this depends on the single point distribution $Q^{\text{c}}(m)$ and can write:

$$\Phi^{(\text{site})} = \int dd P_{\text{e}}(d) \int \prod_{i=1}^{d} \int dm_i Q^c(m_i) \log\langle 2\cosh\left(\beta h_0 + \beta J \sum_{i\in\partial_0} s_i\right)\rangle_{\{m_i\}} \tag{4.38}$$

$$\Phi^{(\text{link})} = \frac{c}{2}\int dm_1 Q(m_1) dm_2 Q(m_2) \log\langle e^{\beta J J s_i s_j}\rangle_{\{m_1,m_2\}} \tag{4.39}$$

For a regular graph with fixed connectivity c, $P_{\text{e}}(d) = \delta(d-(c-1))$ while for a Erdos-renyi random graph, interestingly $P_{\text{e}}(d) = P(d)$!

where $P_{\text{e}}(d)$ is the excess degree probability .

Our task is thus to find the asymptotic distribution $Q(m)$. At the level of rigor used in physics, this is easily done. We obviously assume that $Q(m)$ is unique, and does not depend on the

realization of the disorder (this is not so trivial). Then we realize that the distribution of cavity fields must satisfy a recursion such as

$$Q^{\rm c}_{N+1}(m) = \int dh_0 P(h_0) \int ddP_{\rm e}(d) \prod_{i=1}^{d} \int dm_i Q^{\rm c}_{N+1}(m_i) \delta\left(m - f_{\rm BP}(\{m_i\}, h_0)\right) \tag{4.40}$$

with

$$f_{\rm BP}(\{m_i\}, h_0) = \tanh\left(\beta h_0 + \sum_i \operatorname{atanh}\left(m_i \tanh(\beta J)\right)\right) \tag{4.41}$$

Obviously, once we find the fixed point, we can compute the distribution to total magnetization, which reads almost exactly the same, except now we have to use the actual distribution of neighbors:

$$Q(m) = \int dh_0 P(h_0) \int ddP(d) \prod_{i=1}^{d} \int dm_i Q^{\rm c}(m_i) \delta\left(m - f_{\rm BP}(\{m_i\}, h_0)\right) \tag{4.42}$$

1. Explain the population dynamic
2. Note that the free energy is the same as the one on graphs!!! Dictionary GRAPH to POPULATION
3. Add discussion regular vs random, and excess degree

4.3.3 The relation between Loopy Belief Propagation and the Cavity method

tree-like etc....

Single graph vs POPULATION !

4.3.4 Can we prove it?

Well, we can try !!! First $Q^c(m)$ is not clearly self-averaging, but for sure

$$\begin{aligned} \Phi \leq \max_{Q^c(m)} \quad & \int \quad ddP(d) \prod_{i=1}^{d} \int dm_i Q^c(\{m_i\}) \log\langle 2\cosh\left(\beta h_0 + \beta J \sum_{i\in\partial_0} s_i\right)\rangle_{\{m\}} \\ & - \quad \frac{c}{2} \int dm_1 Q^c(\{m_1\}) dm_2 Q^c(\{m_2\}) \log\langle e^{\beta J J s_i s_j}\rangle_{m_1,m_2} \end{aligned} \tag{4.43}$$

It is possible to show that the extremization leads to $Q(m)$ being a solution of the cavity recursion. This means that we obtain a bound! If we found a distribution $Q(m)$ (or actually, all of them) that satisfies eq., then we have a bound.

Can we get the converse bound easily? Sadly, no. The point is that BP is *NOT* a variational method on a given instance, so we cannot use the mean-field technics! Fortunately, it can be shown rigorously, that, for any ferromagnetic model, BP *IS* gives a lower bound on the free entropy.

It is also instructive to compare to what we had in the fully connected model. Indeed, if $Q(m)$ become a delta (which we expect as c grows) we obtain, using $J = 1/N$

$$\Phi \leq \max_m \mathbb{E}_h \log 2\cosh(\beta h + \beta m) - \frac{N}{2}\log e^{\beta m^2/N} = \mathbb{E}_h \log 2\cosh(\beta h + \beta m) - \beta\frac{m^2}{2} \quad (4.44)$$

which is indeed the result we add in the fully connected limit!

4.3.5 When do we expect this to work?

When do we expect belief propagation to be correct? As we have have discussed, random graphs are locally tree-like: they are trees up to any finite distance. Further assuming that we are in a pure thermodynamic state, we expect that we have short range correlations, so that the large $O(\log N)$ loops should not matter in a large enough system, and these equations should provide a correct description of the model.

Clearly, this must be a good approach for describing a system in a paramagnetic phase, or even a system with a ferromagnetic transition (where we should expect to have two different fixed points of the iterations). It could be, however, that there exists a huge number of fixed points for these equations: how to deal with this situation? Should they all correspond to a given pure state? Fortunately, we do not have such worries, as the situation we just described is the one arising when there is a glass transition. In this case, one needs to use the cavity method in conjunction with the so-called "replica symmetry breaking" approach as was done by Mézard, Parisi, and Virasoro.

Bibliography

The cavity method is motivated by the original ideas from Bethe Bethe (1935) and Peierls Peierls (1936) and used in detail to study ferromagnetism Weiss (1948). Belief propagation first appeared in computer science in the context of Shanon error correction Gallager (1962) and was rediscovered in many different contexts. The name "Belief Propagation" comes in particular from Pearl (1982). The deep relation between loopy belief propagation and the Bethe "cavity" approach was discussed in the early 2000s, for instance in Opper and Saad (2001) and Wainwright and Jordan (2008). The work by Yedidia et al. (2003) was particularly influential. The construction of the cavity method on random graphs presented in this chapter follows the classical papers Mézard and Parisi (2001, 2003). That loopy belief propagation gives a lower bound on the true partition *on any graph* in the case of ferromagnetic (and in general attractive models) is a deep non-trivial result proven (partially) by Willsky et al. (2007) using the loop calculus of Chertkov and Chernyak (2006), and (fully) by Ruozzi (2012). Chertkov (2008) showed how beleif propagation finds the ground state of the RFIM at zero temperature. Finally, that the cavity method gives rigorous upper bounds on the critical temperature for was shown in Saade et al. (2017).

4.4 Exercises

Exercise 4.1: Excess degree in Erdos-Renyi graphs

bla bla bla

Exercise 4.2: The random field ising model on a regular random graph

We have seen that the BP update equation for the RFIM is given by

$$f_{\mathrm{BP}}(\{m_i\}, h_0) = \tanh\left(\beta h_0 + \sum_i \operatorname{atanh}\left(m_i \tanh(\beta J)\right)\right) \tag{4.45}$$

and that the distribution of cavity fields follows (for a random graph with fixed connectivity $c-1$):

$$Q^{\mathrm{cav}\cdot}(m) = \int dh_0 \mathcal{N}(h_0; 0, \Delta) \prod_{i=1}^{c-1} \int dm_i Q^{\mathrm{cav}}(m_i) \delta\left(m - f_{\mathrm{BP}}(\{m_i\}, h_0)\right) \tag{4.46}$$

This can be solved in practice using the population dynamics approach where we represent $Q^{\mathrm{cav}\cdot}(m)$ by a population of $\mathcal{N}$ elements. In this case, formally we iterate a collection, or pool, or elements. Starting from $Q^{\mathrm{cav}\cdot}_{t=0}(m) = [m_1, m_2, m_3, \ldots, m_{\mathcal{N}}]$ with, for instance all $m = 1$ or random initial conditions, we iterate as follows:

- For T steps:
- For all $i = 1 \to \mathcal{N}$:
- Draw a random $h_0 \sim \mathcal{N}(\imath, \cdot)$, and $c-1$ random $\{m_i\}$ from $Q^{\mathrm{cav}\cdot}_{t=0}$.
- Compute $m^{\mathrm{new}} = f_{\mathrm{BP}}(\{m_i\}, h_0)$.
- Assign m to the i elements of the new population $Q_{t+1i} = m^{\mathrm{new}}$.

If $\mathcal{N}$ is large enough (say 10^5) then this is a good approximation of the true population density, and if T is large enough, then we should have converged to the fixed point. Once this is done, we can compute the average magnetization by computing the true marginal as follows:

- Set $\overline{m} = 0$
- For $\mathcal{N}$ steps:
- Draw c random $\{m_i\}$ from $Q^{\mathrm{cav}\cdot}_{t=0}$.
- Compute $m^{\mathrm{new}} = f_{\mathrm{BP}}(\{m_i\}, h_0)$ using this time the c values.
- $\overline{m} += m^{\mathrm{new}}/\mathcal{N}$

1. Consider the RFIM on regular random graphs with connectivity $c = 4$. Compute analytically the phase transition point in $beta$, denoted β_c at zero disordered field ($\Delta = 0$).

2. Implement the population dynamics and find numerically the phase transition point when $\overline{m}(\beta, \Delta)$ become non zero. Draw the phase diagram in (β, Δ) separating the phase where $\overline{m} = 0$ with the one where $\overline{m} \neq 0$.

3. Now, let us specialize to the zero temperature limit. Argue that the iteration has the following limit

$$f_{\rm BP}^{\beta=\infty}(\{m_i\}, h_0) = \text{sign}\left(h_0 + \sum_i \text{sign}\{m_i\}\right) \tag{4.47}$$

and that $Q^c(m)$ admits a simple parametric form that depends on a single parameter $\overline{m}_c$ as

$$Q^c(m) = \frac{1+\overline{m}_c}{2}\delta(m-1) + \frac{1-\overline{m}_c}{2}\delta(m+1) \tag{4.48}$$

4. Compute the critical value of Δ where a non zero magnetization appears.