Belief Propagation and the RFIM

Trying out belief propagation for graph coloring

One cannot fully appreciate the interest of belief propagation without ever coding it and looking at its behavior. In this homework we will hence code the belief propagation as covered in the lecture for Erdős-Rényi graphs coloring for $\beta \to \infty$. And test the following:

(a) Initialize BP close to be uniform fixed point, i.e. $1/q + \epsilon_s^{j\to i}$ and iterate the equations until convergence. Define converge as the time when the

$$\frac{1}{2qM}\sum_{(ij)\in E}\sum_{s}\left|\chi_s^{i\to j}(t+1)-\chi_s^{i\to j}(t)\right| < \tilde{\epsilon}$$

with suitably chosen small $\tilde{\epsilon}$.

(b) Check how the behavior depends on the order of update, i.e. compare what happens if you update all messages at once or randomly one by one.

(c) For parameters where the update converges, plot the convergence time as a function of the average degree c. Do this on as large graphs as is feasible with your code.

(d) Check how the behavior depends on the initialization. What if initial messages are random? What if they are all points towards the first color?

[Hint: It is not necessary to implement your own random graph generation code, for Python users please check `networkx` package, for R users please check `igraph` package.]

Note that the time spend on this assignment will come handy in future homeworks where we will look at simple variants of exactly this assignment in order to study replica symmetry breaking.

Solution.

(a) In the code, the BP messages are initialized as $1/q+\epsilon_s^{j\to i}$ where $\epsilon_s^{j\to i} \sim \mathsf{Uniform}[-\alpha/q, \alpha/q]$, $\alpha = 0.1$, and then normalize. $\tilde{\epsilon}$ is chosen to be 10^{-4}.

(b) Take a specific case, $N = 1000$, $\beta = 2$, $q = 3$ and $c = 5$. If we update BP messages parallelly, BP will not converge in 1000 iterations, but if we update BP messages randomly one by one, BP converge in around 30 iterations.

After each iteration, the corresponding Bethe free energy is computed from the current BP messages.

(c) Here I choose $N = 200, 500, 1000, 2000$, $\beta = 2$, $q = 3$, and $c \in [0.1, 7]$ with step size 0.1, the BP update schedule is randomly one by one. For each c, I drawn 5 random graphs to run BP till converge or reach 1000 iterations, and in the plot the median converge time of these five trials is plotted. We can see there is a jump at $c = 6$.

(d) Here I choose $N = 1000$, $\beta = 2$, $q = 3$ and $c = 5$. And three different initializations: small perturbation from $1/q$, totally random, point mass on first color. One can see for this case they converge to same point where the Bethe free entropy is around 2.19. However, since BP fixed point can be viewed as the stationary point of Bethe free entropy, when Bethe free entropy has multiple stationary points, the BP evolution may converge to different fixed points, which is determined by the initialization.

□

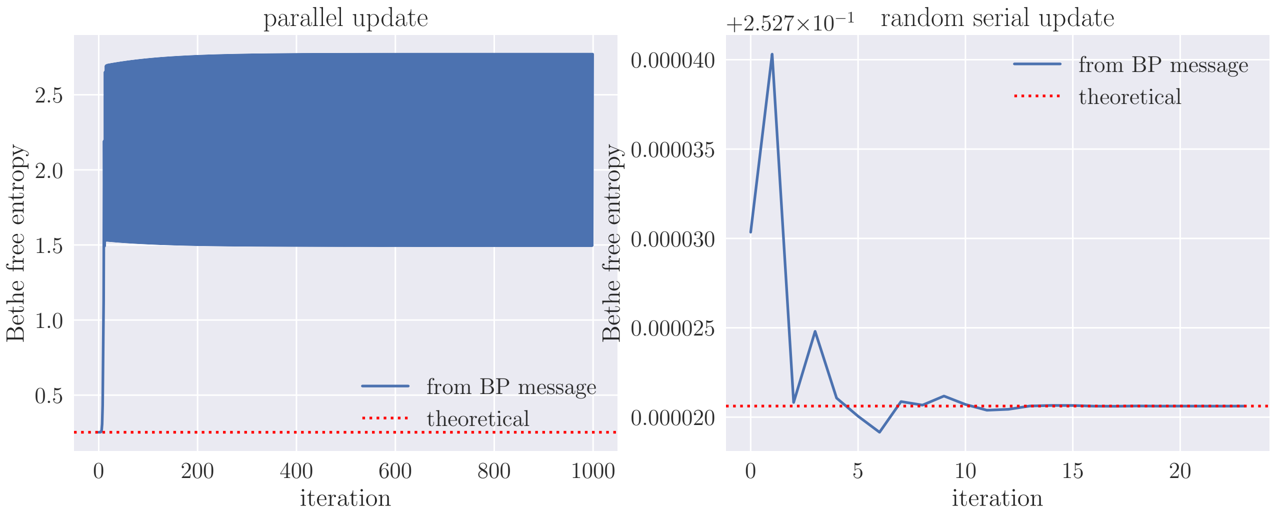

Figure 1: Bethe free entropy trace under different BP update schedule for case $N = 1000$, $\beta = 2$, $q = 3$ and $c = 5$

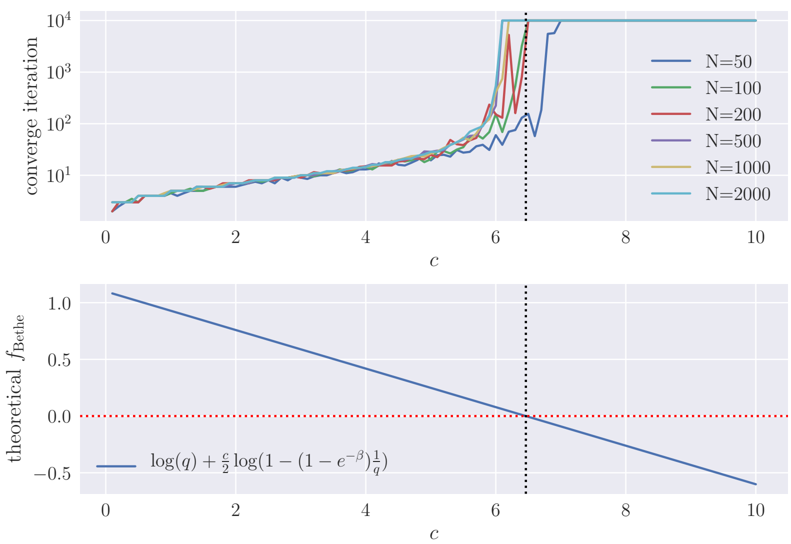

Figure 2: Convergence time vs. c for the case $N = 1000$, $\beta = 2$, $q = 3$.

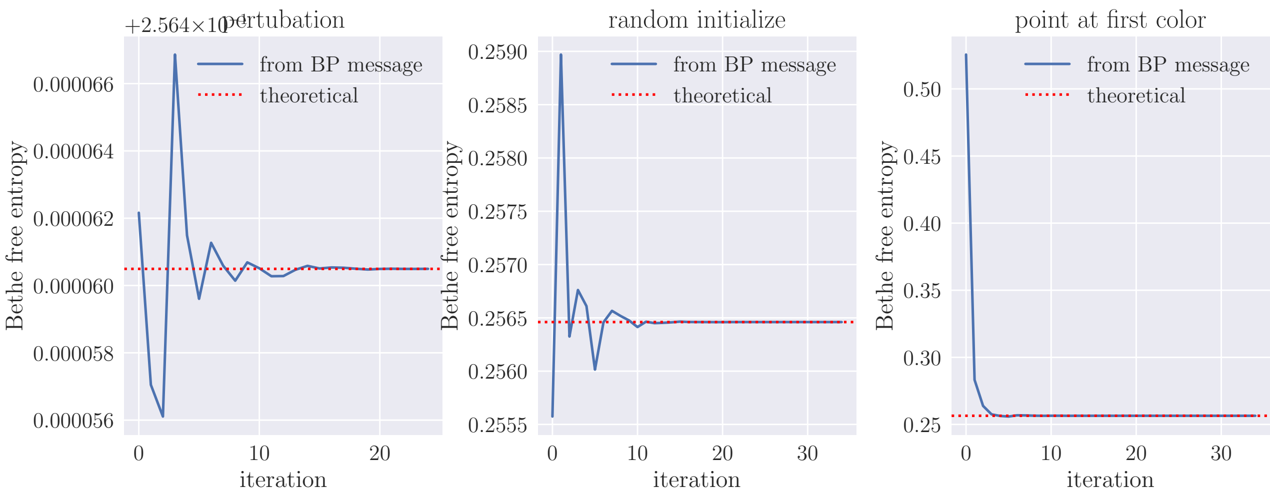

Figure 3: Bethe free entropy trace under different initialization for various N, $\beta = 2$, $q = 3$ and $c = 5$

Concentration of the free energy in the Random Field Ising Model

The Random Field Ising Model (RFIM) is a generalization of the Curie-Weiss model with random fields. Again, given N spins variables $S_i = \pm 1$, the Hamiltonian reads

$$\mathcal{H}\left(\vec{S}, \vec{h}\right) = -N\left(\sum_i \frac{S_i}{N}\right)^2 - \sum_i h_i S_i$$

where the N values of the random fields are chosen randomly from $\mathcal{N}(0, \Delta)$.

For a given realization of the disorder (denoting $\vec{h}$ the ensemble of all values of h), the partition function is therefore an explicit function of $\vec{h}$:

$$F(\beta, \vec{h}) = \log\left(Z(\beta, \vec{h})\right)$$

(a) Show that the function $F(\beta, \vec{h})$ seen as function of h_1 is a Lipschitz function with constant β, i.e.: $|\partial_{h_1} F| \leq \beta$.

(b) Show that the function $F(\beta, \vec{h})$ seen as a function of $\vec{h}$ is Lipschitz with constant $\beta\sqrt{N}$, i.e. that $|f(x) - f(y)| \leq \beta\sqrt{N}\,\|x - y\|_2$, $\forall\, x, y \in \mathbb{R}^n$

(c) Prove that the free entropy per spin (i.e. divided by N) is concentrated with high probability around its mean value using the following theorem: Let $(X_1, \ldots, X_n)$ be a vector of i.i.d. standard Gaussian variables, and let $f\colon \mathbb{R}^n \to \mathbb{R}$ be L-Lipschitz with respect to the Euclidean norm. Then for all $t \geq 0$:

$$\mathbb{P}\left(|f(X) - \mathbb{E}[f(X)]| \geq t\right) \leq 2\mathrm{e}^{-\frac{t^2}{2L^2}}$$

Solution. First let's write $F(\beta, \vec{h})$ explicitly

$$F(\beta, \vec{h}) = \log\left(Z(\beta, \vec{h})\right) = \log\left(\sum_{\vec{S}} \exp\left(-\beta\mathcal{H}(\vec{S}, \vec{h})\right)\right)$$
$$= \log\left(\sum_{\vec{S}} \exp\left(\beta N\left(\sum_i \frac{S_i}{N}\right)^2 + \beta\sum_i h_i S_i\right)\right)$$

(a) Take partial derivative w.r.t. h_1 and note $S_i = \pm 1$, we have

$$\left|\frac{\partial F(\beta, \vec{h})}{\partial h_1}\right| = \left|\frac{\sum_{\vec{S}} \beta S_i \exp\left(\beta N\left(\sum_i \frac{S_i}{N}\right)^2 + \beta\sum_i h_i S_i\right)}{\sum_{\vec{S}} \exp\left(\beta N\left(\sum_i \frac{S_i}{N}\right)^2 + \beta\sum_i h_i S_i\right)}\right|$$
$$\leq \beta\frac{\sum_{\vec{S}} |S_i| \exp\left(\beta N\left(\sum_i \frac{S_i}{N}\right)^2 + \beta\sum_i h_i S_i\right)}{\sum_{\vec{S}} \exp\left(\beta N\left(\sum_i \frac{S_i}{N}\right)^2 + \beta\sum_i h_i S_i\right)} = \beta$$

(b) Given any $\vec{x}, \vec{y} \in \mathbb{R}^N$, we start with defining a auxiliary function $g(t) \equiv F(\beta, t\vec{x} + (1-t)\vec{y})$. Since $F(\beta, \vec{h})$ is a continuous function of $\vec{h}$, by mean-value theorem, there exists $\xi \in [0, 1]$

such that

$$
\begin{aligned}
F(\beta,\vec{x})-F(\beta,\vec{y}) &= g(1)-g(0)=(1-0)\,g'(\xi)\\
&= \left[\nabla_{\vec{h}}F(\beta,\vec{h})\Big|_{\vec{h}=\xi\vec{x}+(1-\xi)\vec{y}}\right]^{\mathrm{T}} \frac{\mathrm{d}\left(t\vec{x}+(1-t)\vec{y}\right)}{\mathrm{d}t}\Bigg|_{t=\xi}\\
&= \begin{bmatrix}\frac{\partial F(\beta,\vec{h})}{\partial h_1} & \dots & \frac{\partial F(\beta,\vec{h})}{\partial h_N}\end{bmatrix}\Bigg|_{\vec{h}=\xi\vec{x}+(1-\xi)\vec{y}} (\vec{x}-\vec{y})\\
&= \sum_{i=1}^{N} \frac{\partial F(\beta,\vec{h})}{\partial h_i}\Bigg|_{\vec{h}=\xi\vec{x}+(1-\xi)\vec{y}} (x_i-y_i)
\end{aligned}
$$

Using Cauchy-Schwartz inequality and the result of part (a), we have

$$
\begin{aligned}
\left|F(\beta,\vec{x})-F(\beta,\vec{y})\right| &= \left|\sum_{i=1}^{N} \frac{\partial F(\beta,\vec{h})}{\partial h_i}\Bigg|_{\vec{h}=\xi\vec{x}+(1-\xi)\vec{y}} (x_i-y_i)\right|\\
&\le \sqrt{\sum_{i=1}^{N}\left|\frac{\partial F(\beta,\vec{h})}{\partial h_i}\Bigg|_{\vec{h}=\xi\vec{x}+(1-\xi)\vec{y}}\right|^2 \times \sum_{i=1}^{N}(x_i-y_i)^2}\\
&\le \sqrt{\sum_{i=1}^{N}\beta^2 \times \sum_{i=1}^{N}(x_i-y_i)^2} = \beta\sqrt{N}\,\|\vec{x}-\vec{y}\|_2
\end{aligned}
$$

Hence, we finished the proof that $F(\beta,\vec{h})$ is a function of $\vec{h}$ with Lipschitz constant $\beta\sqrt{N}$.

(c) Let $f(\beta,\vec{h}) = F(\beta,\vec{h})/N$ to be the free entropy per spin, From the result of part (b) we have

$$
\left|f(\beta,\vec{x})-f(\beta,\vec{y})\right| \le \frac{\beta}{\sqrt{N}}\|x-y\|_2 \quad\Rightarrow\quad f(\beta,\vec{h}) \text{ is } \frac{\beta}{\sqrt{N}}\text{-Lipschitz}
$$

Notice that we cannot apply the theorem above directly, since our $h_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\Delta)$ may not be standard Gaussian.

For convenient, define $\tilde{f}(\beta,\vec{\tilde{h}}) \equiv f(\beta,\sqrt{\Delta}\,\vec{\tilde{h}})$, then we can claim by choosing $\tilde{h}_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, the distribution equivalence $\tilde{f}(\beta,\vec{\tilde{h}}) \overset{\mathrm{d}}{=} f(\beta,\vec{h})$ holds.

$$
\begin{aligned}
\left|\tilde{f}(\beta,\vec{x})-\tilde{f}(\beta,\vec{y})\right| = \left|f(\beta,\sqrt{\Delta}\,\vec{x})-f(\beta,\sqrt{\Delta}\,\vec{y})\right| \le \beta\sqrt{\frac{\Delta}{N}}\,\|\vec{x}-\vec{y}\|_2\\
\Rightarrow\qquad \tilde{f}(\beta,\vec{h}) \text{ is } \beta\sqrt{\frac{\Delta}{N}}\text{-Lipschitz}
\end{aligned}
$$

Now apply the above theorem: $\left(\tilde{h}_1,\dots,\tilde{h}_N\right)$ is a vector of i.i.d. standard Gaussian variables, $\tilde{f}(\beta,\vec{\tilde{h}})$ is a $\beta\sqrt{\frac{\Delta}{N}}$-Lipschitz function of $\vec{\tilde{h}}$ w.r.t. Euclidean norm, then

$$
\mathbb{P}\left(\left|\tilde{f}(\beta,\vec{\tilde{h}})-\mathbb{E}_{\vec{\tilde{h}}}\left[\tilde{f}(\beta,\vec{\tilde{h}})\right]\right|\ge t\right) \le 2\mathrm{e}^{-N\frac{t^2}{2\beta^2\Delta}}
$$

Using the distributional equality before, we finally get

$$
\mathbb{P}\left(\left|f(\beta,\vec{h})-\mathbb{E}_{\vec{h}}\left[f(\beta,\vec{h})\right]\right|\ge t\right) \le 2\mathrm{e}^{-N\frac{t^2}{2\beta^2\Delta}} \xrightarrow{N\to\infty} 0
$$

which means when N is large, the free entropy per spin $f(\beta,\vec{h})$ is concentrated around its expectation w.h.p.

$\square$

Phase diagram

We have seen in our lecture that the free entropy of the random field ising model is given by

$$\Phi(\beta,\Delta)=\lim_{N\to\infty}\frac{\log\left(Z(\beta,\Delta)\right)}{N}=\mathsf{Extr}_m\left\{-\beta\frac{m^2}{2}+\mathbb{E}_h\left[\log\left(2\cosh(\beta(h+m))\right)\right]\right\}$$

(a) Write the self-consistent equation that should satisfies m in order to extremize the expression

(b) In order to solve the self-consistent equation, you need to use a computer for computing the integral, and solving the non-linear equation. Your task it plot the phase diagram as a function of $T=1/\beta$ and Δ

(c) For which value of $T=1/\beta$ and Δ is m^* different than 0? Does the phase transition seen in the Ising model survives in presence of a small random field?

Solution.

(a) Simply take partial derivative of the expression to be extremized w.r.t. m and set it to zero

$$\begin{aligned}0&=\frac{\partial}{\partial m}\left\{-\beta\frac{m^2}{2}+\mathbb{E}_h\left[\log\left(2\cosh(\beta(h+m))\right)\right]\right\}\\&=-\beta m+\mathbb{E}_h\left[\frac{\partial}{\partial m}\log\left(2\cosh(\beta(h+m))\right)\right]\\&=-\beta m+\mathbb{E}_h\left[\beta\tanh(\beta(h+m))\right]\end{aligned}$$

Hence, after arranging terms, the self-consistent equation reads

$$m=\mathbb{E}_h\left[\tanh(\beta(h+m))\right]$$

(b) Notice the self-consistent equation has a trivial solution $m=0$, and if there is another root m^*, $-m^*$ must also be a root since the expectation is taken over $\mathcal{N}(0,\Delta)$ which is symmetric around zero. Therefore, it is sufficient to search for $m^*>0$.

The phase diagram is plotted on $T=\beta^{-1}\in[0,2]$, $\Delta\in[0,1]$, where the heap map is on the value of m^*.

(c) The exterior of left-bottom part describes the phase transition

- When $\Delta=0$, the phase transition happens at $T=1$, which matches the Curie-Weiss model.
- When $T=0$, the phase transition happens around $\Delta=0.63$.

Whether the phase transition survive under "small" random field depends on how do we define "small": for each β, there exists an upper bound $\Delta^{\text{ub}}(\beta)$ such that the phase transition behavior will disappear when the variance of random field exceeds $\Delta^{\text{ub}}(\beta)$.

□

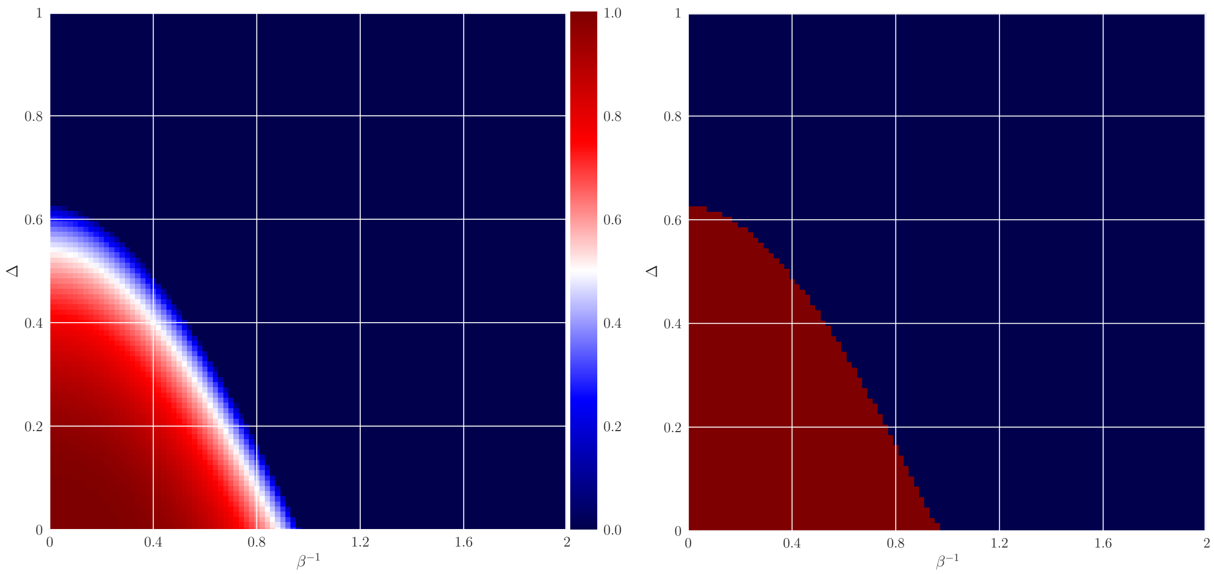

Figure 4: Phase diagram of random field Ising model (left), binary phase digram after thresholding $m > 10^{-12}$ (right)

Belief-Propagation for the RFIM

We now consider the RFIM model on a large regular random graph with connectivity d, using the Hamiltonian:

$$\mathcal{H}\left(\{S_i\}_{i=1}^N, J, \vec{h}\right) = -\sum_{(ij)\in\mathcal{G}} JS_iS_j - \sum_i h_iS_i$$

(a) Show that you can write the problem on a factor graph, and that the BP equations reads

$$\chi_{S_i}^{i\to a} = \frac{e^{\beta h_i S_i}}{Z_\chi^{i\to a}} \prod_{b\in\partial i\backslash a} \psi_{S_i}^{b\to i} \quad \text{and} \quad \psi_{S_i}^{b\to i} = \frac{1}{Z_\psi^{b\to i}} \sum_{\{S_j\}_{j\in\partial b\backslash i}} e^{\beta J \prod_{j\in\partial b} S_j} \prod_{j\in\partial b\backslash i} \chi_{S_j}^{j\to b}$$

(b) We shall use the following transformation:

$$\chi_{S_i}^{i\to a} = \frac{e^{\beta v^{i\to a} S_i}}{2\cosh\left(\beta h^{i\to a}\right)} \quad \text{and} \quad \psi_{S_i}^{b\to i} = \frac{e^{\beta u^{b\to i} S_i}}{2\cosh\left(\beta u^{b\to i}\right)}$$

(c) Show that with these notations, one can rewrite BP as

$$v^{i\to a} = h_i + \sum_{b\in\partial i\backslash a} u^{b\to i} \quad \text{and} \quad \tanh(\beta u^{b\to i}) = \tanh(\beta J) \prod_{j\in\partial b\backslash i} \tanh\left(\beta v^{j\to b}\right)$$

(d) Finally, as in the coloring problem, show that this reduces to a set of equations on $v^{i\to a}$ only. Writting the cavity magnetization as $m^{i\to a} = \tanh\left(\beta v^{i\to a}\right)$, show that in particular

$$m^{i\to a} = \tanh\left[\beta h_i + \operatorname{atanh}\left(\sum_{b\in\partial i\backslash a} \tanh\left(\beta J m^{b\to i}\right)\right)\right]$$

(e) Consider now the large connectivity limit $d = N$, with $J = 1/N$ such that the Hamiltonian reduces to the one of the fully connected model seen in the lecture and show that:

$$m^{i\to a} = \tanh\left(\beta h_i + \mathbb{E}_{\mathcal{G}}\left[\beta m^{b\to i}\right]\right)$$

(f) Show that to first order, one has $m^{i\to a} = m^i$. Denoting the total magnetization as $m = \mathbb{E}_{\mathcal{G}}[m^i]$, show that the solution of the last equation is given by the same self-consistent condition as in the lecture:

$$m = \mathbb{E}_h \left[\tanh(\beta h + \beta m)\right]$$

Solution.

(a) Given the Hamiltonian, we can write down the Boltzmann distribution

$$P_{\beta,J,\vec{h}}\left(\{S_i\}_{i=1}^N\right) = \frac{1}{Z_{J,\vec{h}}(\beta)} \mathrm{e}^{-\beta\mathcal{H}\left(\{S_i\}_{i=1}^N, J, \vec{h}\right)} = \frac{1}{Z_{J,\vec{h}}(\beta)} \prod_i \mathrm{e}^{\beta h_i S_i} \prod_{a\in\mathcal{G}} \mathrm{e}^{\beta J \prod_{j\in\partial a} S_j}$$

Then, directly applying the BP update equation yields

$$\begin{aligned}
\chi_{S_i}^{i\to a} &= \frac{1}{Z_\chi^{i\to a}} g_i(S_i) \prod_{b\in\partial i\backslash a} \psi_{S_i}^{b\to i} = \frac{\mathrm{e}^{\beta h_i S_i}}{Z_\chi^{i\to a}} \prod_{b\in\partial i\backslash a} \psi_{S_i}^{b\to i} \\
\psi_{S_i}^{b\to i} &= \frac{1}{Z_{S_i}^{b\to i}} \sum_{\{S_j\}_{j\in\partial b\backslash i}} f_b\left(\{S_j\}_{j\in\partial b}\right) \prod_{j\in\partial b\backslash i} \chi_{S_j}^{j\to b} = \frac{1}{Z_{S_i}^{b\to i}} \sum_{\{S_j\}_{j\in\partial b\backslash i}} \mathrm{e}^{\beta J \prod_{j\in\partial b} S_j} \prod_{j\in\partial b\backslash i} \chi_{S_j}^{j\to b}
\end{aligned}$$

(b) Notice that $S_i \in \{\pm 1\}$,

$$\begin{aligned}
\chi_{S_i}^{i\to a} &= \frac{\mathrm{e}^{\beta v^{i\to a} S_i}}{2\cosh(\beta v^{i\to a})} = \frac{\mathrm{e}^{\beta v^{i\to a} S_i}}{\mathrm{e}^{\beta v^{i\to a}\cdot(+1)} + \mathrm{e}^{\beta v^{i\to a}\cdot(-1)}} \quad \Rightarrow \quad v^{i\to a} = \frac{1}{2\beta}\log\left(\frac{\chi_{+1}^{i\to a}}{\chi_{-1}^{i\to a}}\right) \\
\psi_{S_i}^{b\to i} &= \frac{\mathrm{e}^{\beta u^{b\to i} S_i}}{2\cosh(\beta u^{b\to i})} = \frac{\mathrm{e}^{\beta u^{b\to i} S_i}}{\mathrm{e}^{\beta u^{b\to i}\cdot(+1)} + \mathrm{e}^{\beta u^{b\to i}\cdot(-1)}} \quad \Rightarrow \quad u^{b\to i} = \frac{1}{2\beta}\log\left(\frac{\psi_{+1}^{b\to i}}{\psi_{-1}^{b\to i}}\right)
\end{aligned}$$

Here u and v are called *log-likelihood-ratio* (LLR), which is a compact parameterization when the spin only takes two values.

(c) Directly apply LLR parameterization into BP equations we have

$$\begin{aligned}
v^{i\to a} &= \frac{1}{2\beta}\log\left(\frac{\chi_{+1}^{i\to a}}{\chi_{-1}^{i\to a}}\right) = \frac{1}{2\beta}\log\left(\frac{\frac{\mathrm{e}^{\beta h_i\cdot(+1)}}{Z_\chi^{i\to a}}\prod_{b\in\partial i\backslash a}\psi_{+1}^{b\to i}}{\frac{\mathrm{e}^{\beta h_i\cdot(-1)}}{Z_\chi^{i\to a}}\prod_{b\in\partial i\backslash a}\psi_{-1}^{b\to i}}\right) \\
&= \frac{1}{2\beta}\log\left(\mathrm{e}^{2\beta h_i}\cdot\prod_{b\in\partial i\backslash a}\mathrm{e}^{2\beta u^{b\to i}}\right) = h_i + \prod_{b\in\partial i\backslash a} u^{b\to i}
\end{aligned}$$

$$\begin{aligned}
\tanh\left(\beta u^{b\to i}\right) &= \frac{\mathrm{e}^{2\beta u^{b\to i}} - 1}{\mathrm{e}^{2\beta u^{b\to i}} + 1} = \frac{\psi_{+1}^{b\to i} - \psi_{-1}^{b\to i}}{\psi_{+1}^{b\to i} + \psi_{-1}^{b\to i}} \\
&= \frac{\sum_{\{S_j\}_{j\in\partial b\backslash i}} \mathrm{e}^{\beta J\prod_{j\in\partial b\backslash i} S_j}\prod_{j\in\partial b\backslash i}\chi_{S_j}^{j\to b} - \sum_{\{S_j\}_{j\in\partial b\backslash i}} \mathrm{e}^{-\beta J\prod_{j\in\partial b\backslash i} S_j}\prod_{j\in\partial b\backslash i}\chi_{S_j}^{j\to b}}{\sum_{\{S_j\}_{j\in\partial b\backslash i}} \mathrm{e}^{\beta J\prod_{j\in\partial b\backslash i} S_j}\prod_{j\in\partial b\backslash i}\chi_{S_j}^{j\to b} + \sum_{\{S_j\}_{j\in\partial b\backslash i}} \mathrm{e}^{-\beta J\prod_{j\in\partial b\backslash i} S_j}\prod_{j\in\partial b\backslash i}\chi_{S_j}^{j\to b}} \\
&= \frac{\sum_{\{S_j\}_{j\in\partial b\backslash i}}\left(\mathrm{e}^{\beta J\prod_{j\in\partial b\backslash i} S_j} - \mathrm{e}^{-\beta J\prod_{j\in\partial b\backslash i} S_j}\right)\prod_{j\in\partial b\backslash i}\mathrm{e}^{2\beta v^{j\to b}\delta_{S_j,+1}}}{\sum_{\{S_j\}_{j\in\partial b\backslash i}}\left(\mathrm{e}^{\beta J\prod_{j\in\partial b\backslash i} S_j} + \mathrm{e}^{-\beta J\prod_{j\in\partial b\backslash i} S_j}\right)\prod_{j\in\partial b\backslash i}\mathrm{e}^{2\beta v^{j\to b}\delta_{S_j,+1}}} \\
&= \frac{\left(\mathrm{e}^{\beta J} - \mathrm{e}^{-\beta J}\right)\left[\sum_{\{S_j\}_{j\in\partial b\backslash i}}^{\text{even}} \mathrm{e}^{2\beta\sum_{j\in\partial b\backslash i} v^{j\to b}\delta_{S_j,+1}} - \sum_{\{S_j\}_{j\in\partial b\backslash i}}^{\text{odd}} \mathrm{e}^{2\beta\sum_{j\in\partial b\backslash i} v^{j\to b}\delta_{S_j,+1}}\right]}{\left(\mathrm{e}^{\beta J} + \mathrm{e}^{-\beta J}\right)\left[\sum_{\{S_j\}_{j\in\partial b\backslash i}}^{\text{even}} \mathrm{e}^{2\beta\sum_{j\in\partial b\backslash i} v^{j\to b}\delta_{S_j,+1}} + \sum_{\{S_j\}_{j\in\partial b\backslash i}}^{\text{odd}} \mathrm{e}^{2\beta\sum_{j\in\partial b\backslash i} v^{j\to b}\delta_{S_j,+1}}\right]}
\end{aligned}$$

$$= \tanh(\beta J) \cdot \frac{\frac{\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}+1\right)+\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}-1\right)}{2} - \frac{\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}+1\right)-\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}-1\right)}{2}}{\frac{\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}+1\right)+\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}-1\right)}{2} + \frac{\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}+1\right)-\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}-1\right)}{2}}$$

$$= \tanh(\beta J) \cdot \frac{\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}-1\right)}{\prod_{j\in\partial b\setminus i}\left(\mathrm{e}^{2\beta v^{j\to b}}+1\right)} = \tanh(\beta J) \cdot \prod_{j\in\partial b\setminus i} \tanh\left(\beta v^{j\to b}\right)$$

where $\sum^{\text{even}}_{\{S_j\}_{j\in\partial b\setminus i}}$, $\sum^{\text{odd}}_{\{S_j\}_{j\in\partial b\setminus i}}$ are summed over $\{S_j\}_{j\in\partial b\setminus i}$ but restricted to terms with $\prod_{j\in\partial b\setminus i} S_i = +1$ and -1, respectively.

(d) Since in this model, all interactions are pairwise, the BP equations can be simplified as

$$v^{i\to(ij)} = h_i + \sum_{k\in\partial^* i\setminus j} u^{(ik)\to i}, \quad \tanh(\beta u^{(ij)\to j}) = \tanh(\beta J)\tanh\left(\beta v^{i\to(ij)}\right), \quad \forall\,(ij)\in\mathcal{G}$$

By letting $m^{i\to(ij)} = \tanh\left(\beta v^{i\to(ij)}\right)$ we have

$$\begin{aligned} m^{i\to(ij)} = \tanh\left(\beta v^{i\to(ij)}\right) &= \tanh\left(\beta h_i + \beta \sum_{k\in\partial^* i\setminus j} u^{(ik)\to i}\right) \\ &= \tanh\left(\beta h_i + \sum_{k\in\partial^* i\setminus j} \operatorname{atanh}\left(\tanh(\beta J)\tanh\left(\beta v^{k\to(ik)}\right)\right)\right) \\ &= \tanh\left(\beta h_i + \sum_{k\in\partial^* i\setminus j} \operatorname{atanh}\left(\tanh(\beta J) m^{k\to(ik)}\right)\right) \end{aligned}$$

(e) In the large connectivity limit $d = N$, the graph grows to fully connected. Besides, as $J = 1/N$, the Hamiltonian reduces to the RFIM in the lecture

$$\mathcal{H}\left(\{S_i\}_{i=1}^N, \vec{h}\right) = -\frac{1}{N}\sum_{i,j} S_i S_j - \sum_i h_i S_i$$

Notice that when $x \to 0$, $\tanh(x) = x + \mathrm{O}(x^3)$, $\operatorname{atanh}(x) = x + \mathrm{O}(x^3)$. Therefore, in the large-N limit, the cavity magnetization equation for the whole random graph ensemble becomes

$$\begin{aligned} m^{i\to(ij)} &= \tanh\left(\beta h_i + \mathbb{E}_{\mathcal{G}}\left[\sum_{k\in\partial^* i\setminus j} \operatorname{atanh}\left(\tanh(\beta J) m^{k\to(ik)}\right)\right]\right) \\ &= \tanh\left(\beta h_i + \mathbb{E}_{\mathcal{G}}\left[\sum_{\substack{k=1\\k\notin\{i,j\}}}^N \operatorname{atanh}\left(\tanh\left(\frac{\beta}{N}\right) m^{k\to(ik)}\right)\right]\right) \\ &= \tanh\left(\beta h_i + \mathbb{E}_{\mathcal{G}}\left[\sum_{\substack{k=1\\k\notin\{i,j\}}}^N \operatorname{atanh}\left(\frac{\beta}{N} m^{k\to(ik)}\right)\right]\right) \\ &= \tanh\left(\beta h_i + \mathbb{E}_{\mathcal{G}}\left[\sum_{\substack{k=1\\k\notin\{i,j\}}}^N \frac{\beta}{N} m^{k\to(ik)}\right]\right) \\ &= \tanh\left(\beta h_i + \mathbb{E}_{\mathcal{G}}\left[\beta m^{k\to(ik)}\right]\right) \end{aligned}$$

(f) Compare equations between m^i and $m^{i\to(ij)}$, and apply Taylor expansion around $m^{i\to(ij)}$ we have

$$\begin{aligned}
m^i &= \tanh\left(\beta h_i + \mathbb{E}_{\mathcal{G}}\left[\sum_{\substack{k=1\\k\neq i}}^{N} \frac{\beta}{N} m^{k\to(ik)}\right]\right)\\
&= \tanh\left(\operatorname{atanh}\left(m^{i\to(ij)}\right) + \frac{\beta}{N}\mathbb{E}_{\mathcal{G}}\left[m^{j\to(ij)}\right]\right)\\
&= m^{i\to(ij)} + \frac{\beta}{N}\left(1 - \left[m^{i\to(ij)}\right]^2\right)\mathbb{E}_{\mathcal{G}}\left[m^{j\to(ij)}\right] + \mathrm{O}(N^{-2})
\end{aligned}$$

which indicates that to the first order $m^{i\to(ij)} = m^i$. Use this conclusion and result of part (e), we can claim that to the first order

$$m^i = \tanh\left(\beta h_i + \beta\mathbb{E}_{\mathcal{G}}[m^k]\right)$$

Note $m = \mathbb{E}_{\mathcal{G}}[m^i]$ does not depend on $\mathcal{G}$, taking expectation over $\mathcal{G}$ and $\vec{h}$ on both sides yields

$$\begin{aligned}
m &= \mathbb{E}_{\vec{h}}\left[\frac{1}{N}\sum_i m\right] = \frac{1}{N}\mathbb{E}_{\vec{h}}\left[\sum_i \mathbb{E}_{\mathcal{G}}[m^i]\right] = \frac{1}{N}\mathbb{E}_{\mathcal{G}}\left[\sum_i \mathbb{E}_{h_i}[m^i]\right]\\
&= \frac{1}{N}\mathbb{E}_{\mathcal{G}}\left[\sum_i \mathbb{E}_{h_i}\left[\tanh\left(\beta h_i + \beta\mathbb{E}_{\mathcal{G}}[m^k]\right)\right]\right]\\
&= \frac{1}{N}\mathbb{E}_{\mathcal{G}}\left[\sum_i \mathbb{E}_{h_i}\left[\tanh\left(\beta h_i + \beta m\right)\right]\right] = \mathbb{E}_h\left[\tanh\left(\beta(h+m)\right)\right]
\end{aligned}$$

which is same as the self-consistent condition in the lecture and finishes the proof.

$\square$