STATS 369 Homework 3

Exercise 1.9

(a) We first compute

$$\begin{aligned}
\mathbb{E}\left[H(\boldsymbol{\sigma}_1)H(\boldsymbol{\sigma}_2)\right] &= \mathbb{E}\left[\sum_{k=1}^{M} c_k \langle b\mathbf{W}^{(k)}, \boldsymbol{\sigma}^{\otimes k}\rangle \sum_{j=1}^{M} c_j \langle b\mathbf{W}^{(j)}, \boldsymbol{\sigma}^{\otimes j}\rangle\right] \\
&= \sum_{k=1}^{M} c_k^2 b^2 \mathbb{E}\left[\langle \mathbf{W}^{(k)}, \boldsymbol{\sigma}_1^{\otimes k}\rangle \langle \mathbf{W}^{(k)}, \boldsymbol{\sigma}_2^{\otimes k}\rangle\right] \\
&= \sum_{k=1}^{M} c_k^2 b^2 \frac{1}{k!n} \mathbb{E}\left[\langle \sum_{\pi\in\mathcal{G}_k} \mathbf{G}^{\pi}, \boldsymbol{\sigma}_1^{\otimes k}\rangle \langle \sum_{\pi\in\mathcal{G}_k} \mathbf{G}^{\pi}, \boldsymbol{\sigma}_2^{\otimes k}\rangle\right] \\
&= \sum_{k=1}^{M} c_k^2 b^2 \frac{k!}{n} \mathbb{E}\left[\langle \mathbf{G}, \boldsymbol{\sigma}_1^{\otimes k}\rangle \langle \mathbf{G}, \boldsymbol{\sigma}_2^{\otimes k}\rangle\right] \\
&= \sum_{k=1}^{M} c_k^2 b^2 \frac{k!}{n} \langle \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2\rangle^k.
\end{aligned}$$

So $n\xi(\langle \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2\rangle) = \sum_{k=1}^{M} c_k^2 b^2 \frac{k!}{n} \langle \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2\rangle^k$. Thus, we have

$$\xi(x) = \sum_{k=1}^{M} c_k^2 k! x^k$$

For the model in the chapter with $\lambda = 0$, the Hamiltonian does not depend on $\boldsymbol{\sigma}$. It corresponds to $\xi = 0$ in this exercise.

(b) By part (a), we can write

$$\begin{aligned}
Z_n &= \int_{\mathcal{S}^{n-1}} \exp\left[n \sum_{k=1}^{M} c_k \langle \mathbf{W}^{(k)}, \boldsymbol{\sigma}^{\otimes k}\rangle\right] \nu_0(d\boldsymbol{\sigma}) \\
&= \int_{\mathcal{S}^{n-1}} \exp\left[n \sum_{k=1}^{M} c_k \frac{1}{\sqrt{k!n}} \langle \sum_{\pi\in\mathcal{G}_n} \mathbf{G}^{(k),\pi}, \boldsymbol{\sigma}^{\otimes k}\rangle\right] \nu_0(d\boldsymbol{\sigma})
\end{aligned}$$

Using Theorem 13, we only need to prove that $\frac{1}{\sqrt{n}} \log Z_n$ is a Lipschitz function of $\mathbf{G}^{(k)}$. Taking the gradient of $\log Z_n$ w.r.t. $\mathbf{G}^{(k)}$ gives

$$\nabla_{\mathbf{W}^{(k)}} \frac{1}{\sqrt{n}} \log Z_n = \frac{1}{\sqrt{n}} \frac{\int_{\mathcal{S}^{n-1}} \exp\left[n \sum_{k=1}^{M} c_k \langle \mathbf{W}^{(k)}, \boldsymbol{\sigma}^{\otimes k} \rangle\right] \nu_0(d\boldsymbol{\sigma}) \sqrt{n} c_k \boldsymbol{\sigma}^{\otimes k} \nu_0(d\boldsymbol{\sigma})}{\int_{\mathcal{S}^{n-1}} \exp\left[n \sum_{k=1}^{M} c_k \langle \mathbf{W}^{(k)}, \boldsymbol{\sigma}^{\otimes k} \rangle\right] \nu_0(d\boldsymbol{\sigma})}.$$

Since $\|\boldsymbol{\sigma}^{\otimes k}\|_F^2 \leq 1$ when $\boldsymbol{\sigma} \in \mathcal{S}^{n-1}$, we have

$$\|\nabla_{\mathbf{W}^{(k)}} \frac{1}{\sqrt{n}} \log Z_n\|_F \leq c_k$$

Thus, $n^{-1/2} \log Z_n$ is Lipschitz continuous in all the $\mathbf{G}^{(k)}$ with Lipschitz constant $\sum_{k=1}^{M} c_k$.

(c) We start by computing

$$\mathbb{E}\left[Z_n^r\right] = \mathbb{E} \int_{(\mathcal{S}^{n-1})^r} \exp\left[n \sum_{k=1}^{M} c_k \langle \mathbf{W}^{(k)}, \sum_{a=1}^{r} \boldsymbol{\sigma}_a^{\otimes k} \rangle\right] \nu_{0,r}(d\boldsymbol{\sigma}).$$

Here we denote $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_r)$ and $\nu_{0,1} = \nu_0 \times \cdots \times \nu_0$. By (1.2.5-6), we have

$$\mathbb{E}\left[\exp\left(n c_k \langle \mathbf{W}^{(k)}, \sum_{a=1}^{r} \boldsymbol{\sigma}_a^{\otimes k} \rangle\right)\right] = \exp\left(\frac{n c_k^2 k!}{2} \sum_{a,b=1}^{r} \langle \boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b \rangle^k\right).$$

Hence,

$$\begin{aligned}
\mathbb{E}\left[Z_n^r\right] &= \int_{(\mathcal{S}^{n-1})^r} \exp\left(\sum_{k=1}^{M} \frac{n c_k^2 k!}{2} \sum_{a,b=1}^{r} \langle \boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b \rangle^k\right) \nu_{0,r}(d\boldsymbol{\sigma}) \\
&= \int_{(\mathcal{S}^{n-1})^r} \exp\left(\sum_{k=1}^{M} \frac{n c_k^2 k!}{2} \sum_{a,b=1}^{r} \hat{Q}_{a,b}^k\right) \nu_{0,r}(d\boldsymbol{\sigma}) \\
&= \int \exp\left(\sum_{k=1}^{M} \frac{n c_k^2 k!}{2} \sum_{a,b=1}^{r} Q_{a,b}^k\right) \nu_{0,r}(d\boldsymbol{\sigma}) f_{n,r-1}(\mathbf{Q}) d\mathbf{Q} \\
&\doteq \exp(n \sup_{\mathbf{Q}} S(\mathbf{Q})),
\end{aligned}$$

where

$$S(\mathbf{Q}) = \frac{1}{2} \sum_{k=1}^{M} c_k k! \sum_{a,b=1}^{r} Q_{a,b}^k + \frac{1}{2} \operatorname{tr} \log(\mathbf{Q}).$$

Under the 1RSB ansatz with block size m, we have

$$\sum_{a,b=1}^{r} Q_{a,b}^k = r + q_1^k m(m-1) \frac{r}{m} + q_0^k (r^2 - m^2 \frac{r}{m}),$$

and

$$\mathrm{tr}\log(\mathbf{Q}) = \log(1+(m-1)q_1+(r-m)q_0) + (\frac{r}{m}-1)\log(1-q_1+m(q_1-q_0)) + \frac{r}{m}(m-1)\log(1-q_1).$$

Therefore,

$$\Psi_{\mathrm{1RSB}} = \lim_{r\to 0}\frac{1}{r}S(Q) = \frac{1}{2}\sum_{k=1}^{M} c_k k!(1-(1-m)q_1^k - mq_0^k) + \frac{1}{2}\frac{q_0}{1-(1-m)q_1-mq_0} + \frac{1}{2m}\log(1-(1-m)q_1-mq_0) - \frac{1-m}{2m}\log(1-q_1).$$

The derivatives w.r.t. q_0 is

$$\begin{aligned}\frac{\partial\Psi}{\partial q_0} &= \frac{1}{2}\sum_{k=1}^{M} c_k k!(-mkq_0^{k-1}) + \frac{mq_0}{2(1-(1-m)q_1-mq_0)^2}\\ &= -\frac{m}{2}(\xi'(q_0) - \frac{q_0}{(1-(1-m)q_1-mq_0)^2}).\end{aligned}$$

So q_0 should satisfy

$$\xi'(q_0) = \frac{q_0}{(1-(1-m)q_1-mq_0)^2}.$$

The derivative w.r.t. q_1 is

$$\frac{\partial\Psi}{\partial q_1} = \frac{1}{2}\sum_{k=1}^{M} c_k k!(-(1-m)kq_1^{k-1}) + \frac{q_0(1-m)}{2(1-(1-m)q_1-mq_0)^2} + \frac{1-m}{2}\frac{q_1-q_0}{(1-(1-m)q_1-mq_0)(1-q_1)}.$$

So q_1 should satisfy

$$\xi'(q_1) = \frac{q_0}{(1-(1-m)q_1-mq_0)^2} + \frac{1}{2}\frac{q_1-q_0}{(1-(1-m)q_1-mq_0)(1-q_1)}.$$