Solutions for Week 2

Bruno and Luca

March 18, 2021

1 Exercise 1.4

Consider again the Hamiltonian of the Curie-Weiss model. A very practical way to sample configurations of N spins from the Gibbs probability distribution

$$P\left(\boldsymbol{S}=\boldsymbol{s};\beta,h\right)=\frac{\exp\left(-\beta\mathcal{H}\left(\boldsymbol{s};h\right)\right)}{Z_N(\beta,h)} \tag{1.1}$$

is the Monte-Carlo-Markov-Chain (MCMC) method, and in particular the Metropolis-Hastings algorithm. It works as follows:

1. Choose a starting configuration for the N spins values $s_i=\pm 1$ for $i=1,\ldots,N$.
2. Choose a spin i at random. Compute the current value of the energy E_{now} and the value of the energy E_{flip} if the spins i is flipped (that is if $S_i^{\text{new}}=-S_i^{\text{old}}$).
3. Sample a number r uniformly in $[0,1]$ and, if $r<e^{\beta(E_{\text{now}}-E_{\text{flip}})}$ perform the flip (i.e. $S_i^{\text{new}}=-S_i^{\text{old}}$) otherwise leave it as it is.
4. Go back to step 2.

If one is performing this program long enough, it is guarantied that the final configuration will have been chosen with the correct probability.

(a) Write a code to perform the MCMC dynamics, and start by a configuration where all spins are equal to $S_i=1$. Take $h=0$, $\beta=1.2$ and try your dynamics for a long enough time (say, with $t_{\max}=100N$ attempts to flips spins) and monitor the value of the magnetization per spin $m=\sum_i S_i/N$ as a function of time. Make a plot for $N=10,50,100,200,1000$ spins. Compare with the exact solution at $N=\infty$. Remarks? Conclusions?

Solution: See Jupyter notebook for solution and discussion.

(b) Start by a configuration where all spins are equal to 1 and take $h=-0.1$, $\beta=1.2$. Monitor again the value of the magnetization per spin $m=\sum_i s_i/N$ as a function of time. Make a plot for $N=10,50,100,200,1000$ spins. Compare with the exact solution at $N=\infty$. Remarks? Conclusions?

Solution: See Jupyter notebook for solution and discussion.

2 Exercise 1.5

An alternative local algorithm to sample from the measure eq. 1 is known as the Glauber or Heat bath algorithm. Instead of flipping a spin at random, the idea is to thermalise this spin with its local environment.

Part I: The algorithm

(a) Let $\bar{S} = \frac{1}{N}\sum_{i=1}^{N} s_i$ be the total magnetisation of a system of N spins. Show that for all $i = 1, \cdots, N$, the probability of having a spin at $S_i = \pm 1$ given that all other spins are fixed is given by:

$$\mathbb{P}\left(S_i = \pm 1 | \{S_j\}_{j\neq i}\right) \equiv P_{\pm} = \frac{1 \pm \tanh\big(\beta(\bar{S}+h)\big)}{2}$$

Solution: As in Metropolis-Hastings, Glauber dynamics is defined in terms of random local updates of the spins. But different from the the former, the latter seeks a transition rate $P_{\pm}$ which preserves equilibrium at every step. To find this transition rate, we therefore need to look at the probability of finding a spin at a state $S_i = \pm$ according to the Boltzmann distribution given that all the other spins are fixed:

$$P_{\pm} = \mathbb{P}\left(S_i = \pm 1 | \{S_j\}_{j\neq i}\right) \tag{2.1}$$

This conditional probability can be obtained by splitting the Boltzmann probability of the N spins system:

$$\mathbb{P}\left(\{S_j\}_{j=1}^{N}\right) = \mathbb{P}\left(S_i, \{S_j\}_{j\neq i}\right) = \mathbb{P}\left(S_i | \{S_j\}_{j\neq i}\right)\mathbb{P}\left(\{S_j\}_{j\neq i}\right) \tag{2.2}$$

and therefore the probability we are after is given by:

$$\mathbb{P}\left(S_i = \pm 1 | \{S_j\}_{j\neq i}\right) = \frac{\mathbb{P}\left(\{S_j\}_{j\neq i}\right)}{\mathbb{P}\left(\{S_j\}_{j=1}^{N}\right)} = \frac{1}{\mathcal{Z}_N} e^{\beta \mathcal{H}_N(\boldsymbol{s})}\mathbb{P}\left(\{S_j\}_{j\neq i}\right) \tag{2.3}$$

Note that we can simply split the Hamiltonian of the N spin system in a part that depends on S_i and the complement:

$$\begin{aligned}
\mathcal{H}_N(\boldsymbol{s}) &= -\frac{1}{2N}\sum_{j,k=1}^{N} s_j s_k - h\sum_{j=1}^{N} s_j \\
&= -\frac{(s_i)^2}{2N} - \frac{s_i}{N}\sum_{j\neq i} s_j - \frac{1}{2N}\sum_{j,k\neq i} s_j s_k - h s_i - h\sum_{j\neq i} s_j \\
&= -\frac{1}{2N} - s_i\left(\frac{1}{N}\sum_{j\neq i} s_j + h\right) + \mathcal{H}_N(\{s_j\}_{j\neq i})
\end{aligned} \tag{2.4}$$

Inserting this in the expression above:

$$\mathbb{P}\left(S_i = s | \{S_j\}_{j\neq i}\right) = \underbrace{\frac{e^{\beta\left[-\frac{1}{N} + \mathcal{H}_N\left(\{S_j\}_{j\neq i}\right)\right]}}{\mathcal{Z}_N}\mathbb{P}\left(\{S_j\}_{j\neq i}\right)}_{\tilde{\mathcal{Z}}(\{S_j\}_{j\neq i})} e^{\beta s\left(\frac{1}{N}\sum_{j\neq i} s_j + h\right)} \tag{2.5}$$

Note that the function of the $j \neq i$ spins $\tilde{\mathcal{Z}}$ can be computed explicitly by requiring that the conditional probability is properly normalised:

$$\tilde{\mathcal{Z}} = \sum_{s\in\{-1,+1\}} \mathbb{P}\left(S_i = s|\{S_j\}_{j\neq i}\right) = e^{\beta\left(\frac{1}{N}\sum\limits_{j\neq i} s_j + h\right)} + e^{-\beta\left(\frac{1}{N}\sum\limits_{j\neq i} s_j + h\right)} \tag{2.6}$$

Putting together:

$$\mathbb{P}\left(S_i = \pm|\{S_j\}_{j\neq i}\right) = \frac{e^{\pm\beta\left(\frac{1}{N}\sum\limits_{j\neq i} s_j + h\right)}}{e^{\beta\left(\frac{1}{N}\sum\limits_{j\neq i} s_j + h\right)} + e^{-\beta\left(\frac{1}{N}\sum\limits_{j\neq i} s_j + h\right)}} = \frac{1}{2}\left[1 \pm \tanh\left(\beta\left(\frac{1}{N}\sum_{j\neq i} s_j + h\right)\right)\right] \tag{2.7}$$

That's the most generic transition probability for the Glauber dynamics, and it doesn't assume anything on the top of the fact that the equilibrium distribution is the Boltzmann distribution. To get the expression of the exercise, we simply note that:

$$\frac{1}{N}\sum_{j\neq i} S_j = \frac{1}{N}\sum_{j=1}^{N} S_j - \frac{S_i}{N} \tag{2.8}$$

and therefore when $N \gg 1$, we can neglect the last term and write:

$$P_\pm = \mathbb{P}\left(S_i = \pm|\{S_j\}_{j\neq i}\right) = \frac{1 \pm \tanh(\beta(m+h))}{2} \tag{2.9}$$

for $m = \frac{1}{N}\sum\limits_{i=1} S_i$.

(b) The Glauber algorithm is defined as follows:

1. Choose a starting configuration for the N spins. Compute the magnetisation m_t and the energy E_t corresponding to the configuration.
2. Choose a spin S_i at random. Sample a random number uniformly $r \in [0,1]$. If $r < P_+$, set $S_i = +1$, otherwise set $S_i = -1$. Update the energy and magnetisation.
3. Repeat step 2 until convergence.

Write a code implementing the Glauber dynamics. Repeat items (a) and (b) of exercise 1 using the same parameters. Compare the dynamics. Comment on the observed differences.

Solution: See Jupyter notebook for solution and discussion.

Part II: Mean-field equations from Glauber

Let's now derive the mean-field equations for the Curie-Weiss model from the Glauber algorithm.

(a) Let m_t denote the total magnetisation at time t, and define $P_{t,m} = \mathbb{P}(m_t = m)$. For simplicity, consider $\beta = 1$ and $h = 0$. Show that for $\delta \ll 1$ we can write:

$$\begin{aligned} P_{t+\delta t,m} = {} & P_{t,m+\frac{2}{N}} \times \left\{\frac{1}{2}\left(1 + m + \frac{2}{N}\right)\right\} \times \frac{1 - \tanh(m + 2/N)}{2} \\ & + P_{t,m-\frac{2}{N}} \times \left\{\frac{1}{2}\left(1 - m + \frac{2}{N}\right)\right\} \frac{1}{2}\left(1 + \tanh(m - 2/N)\right) \\ & + P_{t,m}\left\{\frac{1}{2}(1+m)\frac{1+\tanh(m)}{2} + \frac{1}{2}(1-m)\frac{1-\tanh(m)}{2}\right\}. \end{aligned}$$

This is known as the **master equation**.

Solution: Before going into the solution, note the following useful facts:

- The total number of spins N can be decomposed as $N = N_+ + N_-$, where $N_\pm$ is the total number of spins with values $S_i = \pm 1$. Similarly, the magnetisation can be written:

$$m = \frac{N_+ - N_-}{N} = \frac{N_+ - (N - N_+)}{N} = \frac{2N_+}{N} - 1 = 1 - \frac{2N_-}{N} \tag{2.10}$$

- Therefore, the probability of selecting a spin ± 1 among N spins with magnetisation m, given by $N_\pm/N$, can be simply written in terms of the magnetisation as

$$\mathbb{P}(\text{select } S_i = \pm 1 | \bar{S} = m) = \frac{1 \pm m}{2}. \tag{2.11}$$

- The effect of flipping a single spin from ± 1 to ∓ 1 makes $m \to m \mp \frac{2}{N}$

Now let's go back to the question. To get a magnetisation $m = \frac{1}{N} \sum_{i=1}^{N} s_i$ at time $t + \delta$, there are three possibilities:

i) We had already magnetisation m at time t. We choose a spin $S_i = \pm$ at step 2. but we didn't change it. The probability for this event is:

$$\begin{aligned} &\mathbb{P}\left(\text{select } S_i = -1 \middle| \bar{S} = m\right) P_-(m) + \mathbb{P}\left(\text{select } S_i = +1 \middle| \bar{S} = m\right) P_+(m) \\ &= \frac{1}{2}\left[(1+m)P_+(m) + (1-m)P_-(m)\right] \end{aligned} \tag{2.12}$$

ii) We had magnetisation $m + \frac{2}{N}$, and we flipped a spin from $+1$ to -1. As before, we can compute the probability of this event as:

$$\mathbb{P}\left(\text{select } S_i = -1 \middle| \bar{S} = m + \frac{2}{N}\right) P_-\left(m + \frac{2}{N}\right) = \frac{1}{2}\left(1 + m + \frac{2}{N}\right) P_-\left(m + \frac{2}{N}\right) \tag{2.13}$$

iii) Similarly, if we had a magnetisation $m - \frac{2}{N}$ and we flipped a spin from -1 to $+1$:

$$\mathbb{P}\left(\text{select } S_i = +1 \middle| \bar{S} = m - \frac{2}{N}\right) P_+\left(m - \frac{2}{N}\right) = \frac{1}{2}\left(1 - m + \frac{2}{N}\right) P_+\left(m - \frac{2}{N}\right) \tag{2.14}$$

Putting together:

$$\begin{aligned} P_{t+\delta t, m} = &\frac{1}{2}\left[(1+m)P_+(m) + (1-m)P_-(m)\right] P_{t,m} \\ &+ \frac{1}{2}\left(1 + m + \frac{2}{N}\right) P_-\left(m + \frac{2}{N}\right) P_{t, m+\frac{2}{N}} \\ &+ \frac{1}{2}\left(1 - m + \frac{2}{N}\right) P_+\left(m - \frac{2}{N}\right) P_{t, m-\frac{2}{N}}. \end{aligned} \tag{2.15}$$

Inserting the result from Part I yield the desired expression.

(b) Defining the mean magnetisation with respect to $P_{t,m}$

$$\langle m(t) \rangle = \int \mathrm{d}m \; m \; P_{t,m}$$

and using the master equation above, show we can get an equations for the expected magnetisation:

$$\langle m(t+\delta t)\rangle = \int P_{t,m+2/N} \times \left\{\frac{1}{2}\left(1+m+2/N\right)\right\} \times \frac{1-\tanh(m+2/N)}{2} \times mdm$$
$$+ \int P_{t,m-2/N} \left\{\frac{1}{2}(1-m+2/N)\right\} \times \frac{1+\tanh(m-2/N)}{2} \times mdm$$
$$+ \int P_{t,m} \times \left\{\frac{1+m}{2} \times \frac{1+\tanh(m)}{2} + \frac{1-m}{2} \times \frac{1-\tanh(m)}{2}\right\} \times mdm$$

Solution: Multiplying both sides of eq. (2.13) by m and integrating with respect to m gives the desired result:

$$\begin{aligned}\int_{\mathbb{R}} \mathrm{d}m\ P_{t+\delta t,m} m \equiv \langle m(t+\delta t)\rangle = \frac{1}{4}\int_{\mathbb{R}} \mathrm{d}m\ m \bigg[& P_{t,m+2/N}\left(1+m+\frac{2}{N}\right)\left(1-\tanh\left(m+\frac{2}{N}\right)\right) \\ & + P_{t,m-2/N}\left(1-m+\frac{2}{N}\right)\left(1+\tanh\left(m-\frac{2}{N}\right)\right) \\ & + P_{t,m}(1+m)\left(1+\tanh(m)\right) + P_{t,m}(1-m)\times(1-\tanh(m))\bigg]\end{aligned} \tag{2.16}$$

(c) Making the change of variables $m \to m + 2/N$ in the first integral and $m \to m - 2/N$ in the second and choosing $\delta = \frac{1}{N}$, conclude that for $N \to \infty$ we can write the following continuous dynamics for the mean magnetisation:

$$\frac{d}{dt}\langle m(t)\rangle = -\langle m(t)\rangle + \tanh\langle m(t)\rangle$$

Solution: Making the suggested change of variables in the first and second integrals give:

$$\begin{aligned}&\int_{\mathbb{R}} \mathrm{d}m\ m\ P_{t,m\pm\frac{2}{N}}\left(1\pm m+\frac{2}{N}\right)\left(1\mp\tanh\left(m\pm\frac{2}{N}\right)\right) \overset{m\leftarrow m\pm 2/N}{=} \\ &= \int_{\mathbb{R}} \mathrm{d}m\ \left(m\mp\frac{2}{N}\right) P_{t,m}\left(1\pm m\right)\left(1\mp\tanh(m)\right)\end{aligned} \tag{2.17}$$

Inserting this in eq. (2.16) and simplifying gives:

$$\begin{aligned}\langle m(t+\delta t)\rangle &= \int \mathrm{d}m\ P_{t,m}\left[m + \frac{1}{N}\left(-m+\tanh(m)\right)\right] \\ &= \langle m(t)\rangle + \frac{1}{N}\left[-\langle m(t)\rangle + \langle\tanh m(t)\rangle\right]\end{aligned} \tag{2.18}$$

Taking $\delta = 1/N \ll 1$ as $N \to \infty$, we recognise this as the Euler discretisation of the following ODE:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle m(t)\rangle = -\langle m(t)\rangle + \langle\tanh m(t)\rangle \tag{2.19}$$

Note that as $N \to \infty$, we can further simplify the above by noting the $\langle\tanh m\rangle = \tanh\langle m\rangle + O(1/N)$, yielding the desired result:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle m(t)\rangle = -\langle m(t)\rangle + \tanh\langle m(t)\rangle \tag{2.20}$$

(d) Conclude that the stationary expected magnetisation satisfies the Curie-Weiss mean-field equation.

Solution: By definition, the stationary points are given by:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle m(t)\rangle = 0 \qquad \Leftrightarrow \qquad \langle m(t)\rangle = \tanh\langle m(t)\rangle \tag{2.21}$$

which is exactly the Curie-Weiss mean-field equations. Therefore, for $N \to \infty$ and long enough time $t \gg 1$, indeed our algorithm converge to equilibrium solutions of the Boltzmann distribution P_β.

(e) We can now repeat the experiment of the previous exercise, but using the theoretical ordinary differential equation: start by a configuration where all spins are equal to 1 and take different values of h and β. For which values will the Monte-Carlo chain reach the equilibrium value? When will it be trapped in a spurious maximum of the free entropy $\phi(m)$? Compare your theoretical prediction with numerical simulations.

Solution: See Jupyter notebook for solution and discussion.

3 Exercise 2.1

In order to prove the Gaussian Poincaré inequality, we first need to prove the very generic Efron-Stein inequality, which is at the roots of many important result in probability theory:

Theorem 1 (Efron-Stein). *Suppose that $X_1, \dots, X_n$ and $X'_1, \dots, X'_n$ are independent random variable, with X_i and X'_i having the same law for all i. Let $X = (X_1, \dots, X_i, \dots X_n)$ and $X^{(i)} = (X_1, \dots, X_{i-1}, X'_i, X_{i+1}, \dots X_n)$. Then for any function $f : \mathbb{R}^n \to \mathbb{R}$ we have:*

$$\mathrm{var}(f(X)) \leq \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}[(f(X) - f(X^{(i)}))^2]. \tag{3.1}$$

We are going to prove Efron-Stein using the so-called Lindeberg trick, by considering averages over mixed ensembles of the X_i and X'_i. First we define the set $X_{(i)}$ as the set of random variable that are the prime one up to i, and the original one for all larger indices, i.e. $X_{(i)} = (X'_1, \dots, X'_{i-1}, X'_i, X'i+1, \dots X_n)$. In particular $X_{(0)} = X$ and $X_{(n)} = X'$.

a) Show that (this is called the Lindeberg replacement trick):

$$\mathrm{Var}\left[f(X)\right] = \mathbb{E}[f(X)(f(X) - f(X'))] = \sum_{i=1}^{n} \mathbb{E}[f(X)(f(X_{(i-1)}) - f(X_{(i)}))] \tag{3.2}$$

Solution: By definition, we have that:

$$\mathrm{Var}\left[Y\right] = \mathbb{E}\left[Y^2\right] - (\mathbb{E}\left[Y\right])^2 = \mathbb{E}\left[Y^2\right] - \mathbb{E}\left[Y\right]\mathbb{E}\left[Y\right]. \tag{3.3}$$

Replacing $Y = f(X)$, since X and X' are independent but both follow the same distribution we have:

$$\mathrm{Var}\left[f(X)\right] = \mathbb{E}\left[f(X)^2\right] - \mathbb{E}\left[f(X)\right]\mathbb{E}\left[f(X')\right] = \mathbb{E}[f(X)(f(X) - f(X'))]. \tag{3.4}$$

Now, inside the expectation, we can add and subtract each $f(X_{(i)})$ with $i = 1, ..., n-1$ and get:

$$\begin{aligned}
\mathrm{Var}\left[f(X)\right] &= \mathbb{E}[f(X)(f(X) + (f(X_{(1)}) - f(X_{(1)})) + ... - f(X'))] && (3.5)\\
&= \mathbb{E}[f(X)((f(X_{(0)}) - f(X_{(1)})) + ... + (f(X_{(n-1)}) - f(X_{(n)})))] && (3.6)\\
&= \sum_{i=1}^{n} \mathbb{E}[f(X)((f(X_{(i-1)}) - f(X_{(i)}))] && (3.7)
\end{aligned}$$

b) Show that for all i:

$$\mathbb{E}[f(X)(f(X_{(i-1)}) - f(X_{(i)}))] = \mathbb{E}[f(X^{(i)}(f(X_{(i)}) - f(X_{(i-1)}))] \tag{3.8}$$

$$= \frac{1}{2}\mathbb{E}\left[(f(X) - f(X^{(i)})(f(X_{(i-1)}) - f(X_{(i)})\right] \tag{3.9}$$

Solution: We explicit the dependence on the components of the random variables:

$$\mathbb{E}[f(X)((f(X_{(i-1)}) - f(X_{(i)}))] = \tag{3.10}$$

$$= \mathbb{E}[f(\{X_1, ..., X_i, ..., X_n\})(f(\{X'_1, ..., X'_{i-1}, X_i, ..., X_n\}) - f(\{X'_1, ..., X'_i, X_{i+1}, ..., X_n\}))] \tag{3.11}$$

Now, since we are taking the expectation over each independent component of the random variables, we can exchange name between $X_i \leftrightarrow X'_i$ and get an equivalent result:

$$\mathbb{E}[f(X)((f(X_{(i-1)}) - f(X_{(i)}))] = \tag{3.12}$$

$$= \mathbb{E}[f(\{X_1, ..., X'_i, ..., X_n\})(f(\{X'_1, ..., X'_i, X'_{i+1}, ..., X_n\}) - f(\{X'_1, ..., X'_{i-1}, X_i, ..., X_n\}))] \tag{3.13}$$

$$= \mathbb{E}[f(X^{(i)})(f(X_{(i)}) - f(X_{(i-1)}))] = \frac{1}{2}\mathbb{E}[(f(X) - f(X^{(i)}))(f(X_{(i-1)}) - f(X_{(i)}))] \tag{3.14}$$

c) Show that by Cauchy-Schwartz:

$$|\mathbb{E}[f(X)(f(X_{(i-1)}) - f(X_{(i)}))]| \leq \frac{1}{2}\mathbb{E}\left[(f(X) - f(X^{(i)})^2\right] \tag{3.15}$$

and proove the Efron-Stein theorem.

Solution: Usually, we write down Cauchy-Schwartz as:

$$\langle X, Y\rangle^2 \leq \|X\|^2\|Y\|^2, \tag{3.16}$$

but in the case of random variables we can define the inner product through the expectation $\langle X, Y\rangle = \mathbb{E}[XY]$, so we have:

$$(\mathbb{E}[XY])^2 \leq (\mathbb{E}X^2)(\mathbb{E}Y^2). \tag{3.17}$$

Therefore, we have:

$$(\mathbb{E}[(f(X) - f(X^{(i)}))(f(X_{(i-1)}) - f(X_{(i)}))])^2 \leq \mathbb{E}\left(f(X) - f(X^{(i)})\right)^2 \mathbb{E}\left(f(X_{(i-1)}) - f(X_{(i)})\right)^2. \tag{3.18}$$

Note that, in both expectations on the RHS, the random variables differ only in the i-th component, therefore giving the same result. Thus, after taking the square root we substitute back in eq. 3.7 and get:

$$\mathrm{Var}\,[f(X)] = \tag{3.19}$$

$$= \frac{1}{2}\sum_{i=1}^{n}\mathbb{E}[(f(X) - f(X^{(i)})((f(X_{(i-1)}) - f(X_{(i)}))] \tag{3.20}$$

$$\leq \frac{1}{2}\sum_{i=1}^{n}(\mathbb{E}[(f(X) - f(X^{(i)}))^2]) \tag{3.21}$$

Now that we have Efron-Stein, we can proove Poincare's inequality for Gaussian random variables. We shall do it for a single variable, and let the reader generalize the proof to the multi-value case.

With X_i a ± 1 random variable that takes each value with probability $1/2$ (this is called a Rademacher variable), define:

$$S_n = X_1 + X_2 + \ldots + X_n\,. \tag{3.22}$$

d) Using Efron-Stein, show that

$$\mathrm{Var}[f(\frac{S_n}{\sqrt{n}})] \leq \frac{n}{4}\mathbb{E}\left[\left(f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{1}{\sqrt{n}}\right)-f\left(\frac{S_{n-1}}{\sqrt{n}}-\frac{1}{\sqrt{n}}\right)\right)^2\right] \tag{3.23}$$

Solution: We directly apply Efron-Stein, and after isolating the expectation over the i-th component we get:

$$\mathrm{Var}[f(\frac{S_n}{\sqrt{n}})] \leq \frac{1}{2}\sum_{i=1}^{n}\mathbb{E}_{S_{n-1}}\left[\mathbb{E}_i\left(f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{X_i}{\sqrt{n}}\right)-f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{X_i'}{\sqrt{n}}\right)\right)^2\right]. \tag{3.24}$$

Now, the internal expectation can be expanded to a sum of 4 terms, each weighted by a probability 1/4:

$$\begin{aligned}
&\mathbb{E}_i\left(f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{X_i}{\sqrt{n}}\right)-f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{X_i'}{\sqrt{n}}\right)\right)^2 = \\
=\;&\frac{1}{4}\left(f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{1}{\sqrt{n}}\right)-f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{1}{\sqrt{n}}\right)\right)^2+\frac{1}{4}\left(f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{-1}{\sqrt{n}}\right)-f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{-1}{\sqrt{n}}\right)\right)^2 \\
&+\frac{1}{4}\left(f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{1}{\sqrt{n}}\right)-f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{-1}{\sqrt{n}}\right)\right)^2+\frac{1}{4}\left(f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{-1}{\sqrt{n}}\right)-f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{1}{\sqrt{n}}\right)\right)^2 \\
=\;&\frac{1}{2}\left(f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{1}{\sqrt{n}}\right)-f\left(\frac{S_{n-1}}{\sqrt{n}}-\frac{1}{\sqrt{n}}\right)\right)^2.
\end{aligned} \tag{3.25}$$

Therefore, since every term in the sum in eq. 3.24 gives the same contribution, we get:

$$\mathrm{Var}[f(\frac{S_n}{\sqrt{n}})] \leq \frac{n}{4}\mathbb{E}\left[\left(f\left(\frac{S_{n-1}}{\sqrt{n}}+\frac{1}{\sqrt{n}}\right)-f\left(\frac{S_{n-1}}{\sqrt{n}}-\frac{1}{\sqrt{n}}\right)\right)^2\right]. \tag{3.26}$$

e) Using the central limit theorem, show that this leads, as $n \to \infty$ to the following theorem:

Theorem 2 (Gaussian-Poincaré)**.** *Suppose f: $\mathbb{R} \mapsto \mathbb{R}$ is a smooth function and X is Gaussian $X \sim \mathcal{N}(0,1)$, then*

$$Var[f(X)] \leq \mathbb{E}\left[(f'(X))^2\right]. \tag{3.27}$$

Solution: The CLT tells us that the normalized variables $S_n/\sqrt{n}$ and $S_{n-1}/\sqrt{n}$ will be normally distributed $\mathcal{N}(0,1)$. Then, we can bring the factor $n/4$ on the RHS inside the expectation and get:

$$\mathrm{Var}[f(X)] \leq \mathbb{E}\left[\left(\frac{f\left(X+\frac{1}{\sqrt{n}}\right)-f\left(X-\frac{1}{\sqrt{n}}\right)}{2\frac{1}{\sqrt{n}}}\right)^2\right]. \tag{3.28}$$

Finally, in the limit of $n \to \infty$ the central difference gives us the derivative of f evaluated in X, giving:

$$\mathrm{Var}[f(X)] \leq \mathbb{E}\left[\left(f'(X)\right)^2\right]. \tag{3.29}$$

4 Exercise 2.2

The goal of this exercise is to provide an alternative proof of the the free entropy of the random field Ising model, using a technique close to the cavity method.

a) Show that, by adding one spin to a spin of N spins, on has:

$$A_n(\beta, h) = \mathbb{E} \log \frac{Z_{N+1}}{Z_N} = \mathbb{E}_h \log\langle e^{-\frac{\beta}{2}\bar{S}^2} 2\cosh\left(\beta(\bar{S} + h)\right)\rangle_{N,\beta,\boldsymbol{h}} + o(1) \tag{4.1}$$

Solution: We start by writing explicitly the Hamiltonian of a system of $N+1$ spins, assigning index 0 to the newly added spin, and then collect the terms in order to reconstruct $\mathcal{H}_N$:

$$\begin{aligned}
\mathcal{H}_{N+1} &= -\frac{(N+1)}{2}\left(\sum_{i=0}^{N} \frac{S_i}{N+1}\right)^2 - \sum_{i=0}^{N} h_i S_i = && (4.2)\\
&= \left(-\frac{N}{2}\overline{S}^2_{(N)} - \sum_{i=1}^{N} h_i S_i\right) + \frac{N}{2(N+1)}\overline{S}^2_{(N)} - \left(\frac{N}{N+1}\overline{S}_{(N)} + h_0\right) S_0 && (4.3)\\
&= \mathcal{H}_N + \frac{1}{2}\overline{S}^2_{(N)} - \left(\overline{S}_{(N)} + h_0\right) S_0 + o(1) && (4.4)
\end{aligned}$$

Thus, we have:

$$\begin{aligned}
A_N(\beta, \boldsymbol{h}) &= \mathbb{E} \log \frac{\sum_{\{S_i\}_{i=1}^N} e^{-\beta \mathcal{H}_N(\boldsymbol{h})}\left(e^{-\frac{\beta}{2}\overline{S}^2_{(N)}} \sum_{S_0} e^{\beta(\overline{S}_{(N)}+h_0)S_0}\right)}{\sum_{\{S_i\}_{i=1}^N} e^{-\beta \mathcal{H}_N(\boldsymbol{h})}} = && (4.5)\\
&= \mathbb{E} \log\langle e^{-\frac{\beta}{2}\overline{S}^2_{(N)}} 2\cosh\left(\beta(\overline{S}_{(N)} + h_0)\right)\rangle_{N,\beta,\boldsymbol{h}} && (4.6)
\end{aligned}$$

b) Show that, by adding an external magnetic field B to the Hamiltonian (i.e. a term $B\sum_i S_i$, one can concentration of the magnetisation for almost all B so that:

$$\mathbb{E} \int_{B1}^{B2} \langle \bar{S}^2\rangle_{N,\beta,\boldsymbol{h}} - \langle \bar{S}\rangle^2_{N,\beta,\boldsymbol{h}} dB \leq 2/\beta N \tag{4.7}$$

Solution: Since at any fixed value of $\boldsymbol{h}$ we have that $\beta^{-1}\frac{d}{dB}\frac{\log Z_N(\boldsymbol{h})}{N} = \langle \overline{S}\rangle_{N,\beta,\boldsymbol{h}}$, and $\beta^{-2}\frac{d^2}{dB^2}\frac{\log Z_N(\boldsymbol{h})}{N} = N\mathrm{Var}[\overline{S}]_{N,\beta,\boldsymbol{h}}$, we can write:

$$\begin{aligned}
&\mathbb{E} \int_{B1}^{B2} \langle \bar{S}^2(B)\rangle_{N,\beta,\boldsymbol{h}} - \langle \bar{S}(B)\rangle^2_{N,\beta,\boldsymbol{h}} dB = && (4.8)\\
&= \frac{1}{\beta N}\mathbb{E}\int_{B1}^{B2} \frac{d}{dB}(\langle \bar{S}(B)\rangle_{N,\beta,\boldsymbol{h}}) = && (4.9)\\
&= \frac{1}{\beta N}\mathbb{E}(\langle \bar{S}\rangle_{N,\beta,\boldsymbol{h},B2} - \langle \bar{S}\rangle_{N,\beta,\boldsymbol{h},B1}) \leq \frac{1}{\beta N}\sup_{\boldsymbol{h}}\{\langle \bar{S}\rangle_{N,\beta,\boldsymbol{h},B2} - \langle \bar{S}\rangle_{N,\beta,\boldsymbol{h},B1}\} && (4.10)\\
&\leq \frac{2}{\beta N}. && (4.11)
\end{aligned}$$

Therefore, we just found that, in high dimensions and for any fixed realization of the disorder, the value of $\overline{S}$ rapidly concentrates around its mean value. However, we still haven't connected the various realizations of the disorder: in the following we will prove that the exact realization of $\boldsymbol{h}$ is in fact irrelevant.

c) Explain why this implies, almost everywhere in B, that at large N:

$$A_n(\beta, h, B) = \mathbb{E}_h\left[-\frac{\beta}{2}\langle\bar{S}\rangle^2_{N,\beta,\boldsymbol{h}} + \log 2\cosh\left(\beta(\bar{\langle}S\rangle^2_{N,\beta,\boldsymbol{h}} + h + B)\right)\right] + o(1) \tag{4.12}$$

Solution: Since we have concentration of the magnetization $m(\beta, \boldsymbol{h}) = \langle\overline{S}\rangle_{\beta,\boldsymbol{h}}$, we can obtain the leading as $N \to \infty$ by exchanging the log and the expectation:

$$\begin{aligned}
A_n(\beta, h, B) &= \mathbb{E}\log\langle e^{-\frac{\beta}{2}\overline{S}^2} 2\cosh\big(\beta(\overline{S}_{(N)} + h_0)\big)\rangle_{N,\beta,\boldsymbol{h}} \\
&= \mathbb{E}_{h_0}\left[-\frac{\beta}{2}\langle\overline{S}^2\rangle_{N,\beta,\boldsymbol{h}} + \log 2\cosh\left(\beta(\langle\overline{S}\rangle_{N,\beta,\boldsymbol{h}} + h_0 + B)\right)\right] + o(1) \\
&= \mathbb{E}_{h_0}\left[-\frac{\beta}{2}\langle\overline{S}\rangle^2_{N,\beta,\boldsymbol{h}} + \log 2\cosh\left(\beta(\langle\overline{S}\rangle_{N,\beta,\boldsymbol{h}} + h_0 + B)\right)\right] + o(1)
\end{aligned} \tag{4.13}$$

where in the last equality we used $\langle\overline{S}^2\rangle = \langle\overline{S}\rangle^2 + \mathrm{Var}\overline{\mathrm{S}}$.

d) Show that this implies, as $N \to \infty$, the bound:

$$\Phi(\beta, \Delta, B) \leq \sup_m\left[-\frac{\beta}{2}m^2 + \mathbb{E}_h \log 2\cosh\left(\beta(m + h + B)\right)\right] \tag{4.14}$$

Solution: The average free entropy is exactly the $N \to \infty$ limit of A_N. Here, we can again use the fact that the expectation over the possible values of $\langle\overline{S}\rangle$ can be upper bounded by the supremum over all possible values in its support:

$$\Phi(\beta, \Delta, B) \leq \sup_m\left[-\frac{\beta}{2}m^2 + \mathbb{E}_h \log 2\cosh\left(\beta(m + h + B)\right)\right]. \tag{4.15}$$

e) Use the variational approach of lecture 1 to obtain the converse bound and finally show:

$$\Phi(\beta, \Delta) = \sup_m - \frac{\beta}{2}m^2 + \mathbb{E}_h \log 2\cosh\left(\beta(m + h)\right) \tag{4.16}$$

Solution: As in our approach to the Curie-Weiss model, we observe that our upper bound only depends on a one-dimensional magnetization parameter. So we might attempt looking for a lower bound using a single parameter variational family:

$$\mathcal{Q}(\{S_i\}) = \prod_i\left(\frac{(1+m)}{2}\delta(S_i - 1) + \frac{(1-m)}{2}\delta(S_i + 1)\right), \tag{4.17}$$

which gives:

$$\Phi(\mathcal{Q}) = \frac{1}{N}(\mathcal{S}(\mathcal{Q}) - \beta\langle\mathcal{H}\rangle_{\mathcal{Q}}) = h(m) + \beta(\frac{1}{2}m^2 + \frac{m}{N}\sum_i \mathbb{E}h_i) \leq \Phi(\beta, \Delta) \tag{4.18}$$

and recalling that the Shannon entropy can be rewritten as:

$$h(m) = \log(2\cosh(\mathrm{atanh}(m))) - m\,\mathrm{atanh}(m), \tag{4.19}$$

that the extremization of eq. 4.15 gives:

$$m^\star = \mathbb{E}_h \tanh(\beta(m^\star + h)), \tag{4.20}$$

and using Jansen's inequality $\mathbb{E}f(X) \leq f(\mathbb{E}X)$ we find:

$$\begin{aligned} \Phi(\beta,\Delta) \geq \Phi(\mathcal{Q})(m^\star) \quad &\geq \quad \mathbb{E}_h \log(2\cosh(\beta(m^\star + h))) - \mathbb{E}_h m^\star(\beta(m^\star + h)) + \beta(\frac{1}{2}(m^\star)^2 + m\mathbb{E}_h h) \\ &\geq \quad \mathbb{E}_h \log(2\cosh(\beta(m^\star + h))) - \frac{\beta}{2}(m^\star)^2. \end{aligned} \tag{4.21}$$

So, our lower bound matches the upper bound we previously found, which implies that equality holds.