8 Spin glass theory and beyond

Spin glasses phenomenology. We have seen that the behavior of *glasses* is qualitatively reproduced by the p-spin spherical models: a dynamic transition without a thermodynamic one, aging at low temperature, a low-temperature thermodynamic transition as the result of an entropy crisis. As a result, these models and others with a 1RSB low-temperature phase are thought of as mean-field models of glasses. However, spin glasses have some behaviors which are qualitatively different from these, namely memory and rejuvenation in dynamics. Mean-field models of spin glasses are usually FRSB, where the number of RSB breakings is taken to infinity and the function $P(q)$ has a continuous section.

1. What does the landscape of free energy pure states look like in a system with many levels of RSB? Discuss how this landscape is explored by aging at a specific temperature.

2. In FRSB models, the aging solution to dynamics has an effective temperature that varies *continuously* with C for $C < q$ for some q. What are the implications for the pure states being explored by aging dynamics in such systems? What are the implications regarding the emergent timescales in such systems?

3. We saw in pure models that each pure state of the free energy is connected to a minimum of the zero-temperature energy landscape, and that this correspondence survives at all temperatures until the state is destabilized. General models do not have this property, and in fact many exhibit *temperature chaos*, a phenomenon where there is little overlap between pure states at nearby temperatures. Discuss how aging at two successive temperatures would differ in models with and without temperature chaos.

4. Rejuvenation consists of the following behavior: a spin glass aged at a certain temperature exhibits a smaller integrated response after aging than before aging. When cooled to slightly lower temperature, the effect of the aging at the previous temperature goes away, and the integrated response is just as large as it would have been if the system was not aged at the previous temperature. Discuss how the considerations above might explain this phenomenon.

5. Memory consists of the following behavior: a spin glass that has undergone the procedure described in the previous question is reheated back to the temperature at which it was aged. There, its integrated response is found to be similar to that at the end of aging, despite the fact it has not undergone aging at that temperature a second time, and that at lower temperatures the effects of the earlier aging appeared to have vanished. Discuss how the considerations above might explain this phenomenon.

Statistical inference. The tensor denoising problem is a common problem in inference,[1] and consists of the following problem: given a low-rank tensor constructed that has been corrupted by Gaussian noise, can the original low-rank tensor be recovered? To be more concrete, given a signal $x \in \mathbb{R}^N$, the rank-one version of this problem is the recovery of x from data of the form

$$Y_{i_1,\ldots,i_p} = \lambda x_{i_1} \cdots x_{i_p} + J_{i_1 \cdots i_p} \tag{1}$$

where the Js are independent random Gaussian numbers with zero mean and unit variance and λ sets the signal-to-noise ratio.

1. Show that the probability of a specific entry of Y given a signal x is

$$p(Y_{i_1,\ldots,i_p} \mid x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(Y_{i_1,\ldots,i_p} - \lambda x_{i_1} \cdots x_{i_p})^2} \tag{2}$$

2. Given only Y, how can we find x in a principled way? We often want to maximize the probability that our guess for x is correct given that we have the specific data tensor Y, which is called the *likelihood*. Using Bayes' theorem

$$p(Y)p(x \mid Y) = p(Y \mid x)p(x) \tag{3}$$

and supposing we know that x belongs to the sphere of radius $\sqrt{N}$, show that

$$p(x \mid Y) = \exp\left[\lambda \sum_{i_1,\ldots,i_p}^{N} Y_{i_1,\ldots,i_p} x_{i_1} \cdots x_{i_p}\right] \delta(\|x\|^2 - N) \tag{4}$$
$$\times \text{ constants independent of } x$$

Does this look familiar? What does maximizing the likelihood consist of?

3. Unlike the spherical p-spin model, the data tensor Y is not IID random: we suppose that we know it was generated using the procedure above for some ground truth signal x^*. Show that under this assumption, the likelihood is proportional to

$$\exp\left[\lambda \sum_{i_1,\ldots,i_p}^{N} J_{i_1,\ldots,i_p} x_{i_1} \cdots x_{i_p} + \lambda^2 (x \cdot x^*)^p\right] \delta(\|x\|^2 - N) \tag{5}$$

Argue that the scaling in N which separates recovery of the signal from no recovery is $\lambda \propto N^{-\frac{p}{2}}$.

[1] For a long time it was the principle way that recommendation algorithms worked.

4. A key object of study in statistical inference is the *mutual information*, which is the difference between the Shannon entropy

$$H(X) = -\int dx\, p(x) \log p(x) \tag{6}$$

of Y and that of Y given x. For us, this is

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y \mid X) = H(Y) - H(J) \\ &= -\int dY\, p(Y) \log p(Y) + \int dJ\, p(J) \log p(J) \\ &= -\int dY\, dx^*\, p(x^*) p(Y \mid x^*) \log \int dx\, p(x) p(Y \mid x) - N^p \frac{1}{2}(1 + \log 2\pi) \end{aligned} \tag{7}$$

where we have used $p(Y) = \int dx\, p(x) p(Y \mid x)$. The mutual information is nonnegative; if it is positive some information about the signal exists in the data, while if it zero there is no information about the signal in the data. Show that the mutual information consists in part of the free energy of the likelihood (5) averaged over all possible ground truth signals.

5. The transition between zero and nonzero mutual information is a transition in whether the signal can *in principle* be recovered. Argue that this corresponds to a phase transition in the model defined by (5).

   Hint: In a paramagnetic phase, annealed and quenched averages are equal!

6. In practice, the signal must be recovered through some algorithm, which typically consist of an effective quench on the negative logarithm of the likelihood (5) (log-likelihood) from a random initial condition. Argue that whether such a procedure will succeed or fail depends on the low-temperature phase of the model.

7. Depending on the low-temperature phase of the model, the mutual information may be positive but signal cannot be recovered. Under what circumstances does this happen? This is known as a *statistical to computational gap*.