Statistics 291: Lecture 22 (April 11, 2024)

Proof of Spherical Parisi Formula

Instructor: Mark Sellke

Scribe: Kenny Gu

1 Introduction

Last class, we used Guerra's interpolation method to prove an upper bound for the free energy of spherical spin glasses. That is, if $\beta = 1$ and ξ satisfies $\xi''(t) \geq 0$ for all $t \in [-1,1]$, we have

$$\begin{aligned}\limsup_{N\to\infty} \mathbf{E} F_N &\leq \inf_{\substack{b>1 \\ m_0<\cdots<m_{r-1} \\ 0\leq q_1\leq\cdots\leq q_{r-1}}} \frac{1}{2}\left(\frac{\xi'(q_0)}{D_r} + \sum_k \frac{\log(D_{k+1}/D_k)}{m_k} - \sum_k m_k(\theta(q_k)-\theta(q_{k-1})) + b - 1 - \log b\right) \\ &\equiv \mathrm{Par}(\xi, \beta=1).\end{aligned}$$

where $\theta(t) = t\xi'(t) - \xi(t)$ and $D_k = b - \sum_{d=1}^k m_d(\xi'(q_d) - \xi'(q_{d-1}))$.

2 Crisanti-Sommers formula

We can make the formula for $\mathrm{Par}(\xi, \beta = 1)$ slightly nicer by replacing $(\vec{m}, \vec{q})$ with the step function f where $f(q) = m_{d-1}$ for $q \in [q_{d-1}, q_d)$. Then, we can rewrite

$$\mathrm{Par}(\xi, \beta=1) = \frac{1}{2}\left(\frac{\xi'(0)}{D(0)} + \int_0^1 \frac{\xi''(q)}{D(q)} - f(q)q\xi''(q)dq + b - 1 - \log b\right)$$

where $D(q) = b - \int_0^q f(s)\xi''(s)ds$ is the continuous analog of the D_k terms.

We can actually make this formula even nicer: the Crisanti-Sommers formula states that

$$\mathrm{Par}(\xi, \beta=1) = \inf_{\substack{f \text{ increasing} \\ f:[0,1]\to[0,1]}} \frac{1}{2}\left(\int_0^1 \xi'(q)f(q)dq + \int_0^{q^*} \frac{dq}{\hat{f}(q)} + \log(1-q^*)\right)$$

where $q^* < 1$ is a point such that $f(q^*) = 1$ and we define $\hat{f}(q) = \int_q^1 f(u)du$. In this form, we can see that this bound essentially amounts to minimizing a functional over CDFs on $[0,1]$, which, as discussed in previous lectures, we can interpret as an overlap distribution. This view is further confirmed by the following theorem:

Theorem 2.1 (Jagannath-Tobasco)**.** *There exists a piecewise analytic function f^* that minimizes the Crisanti-Sommers functional.*

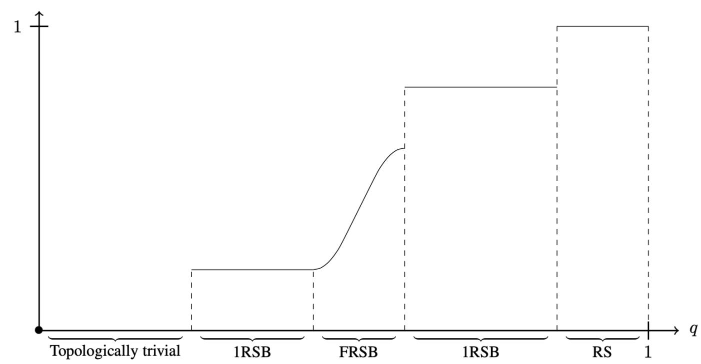

Figure 1: Example of a minimizing function $f^* : [0,1] \to [0,1]$

This theorem says that there exists a CDF on $[0,1]$ that achieves this infimum and that can be split up into finitely many intervals such that the CDF is analytic on each of these intervals. That is, the overlap distribution decomposes into finitely many “nice" components, as illustrated in Figure 1.

Note that each of these components corresponds to either a topologically trivial phase (when $f^*(q) = 0$), a 1RSB phase (when $f^*(q)$ is constant and nonzero over some interval), a full RSB phase (when $f^*(q)$ is strictly monotonically increasing), or a replica symmetric phase (when $f^*(q) = 1$).

To complete the proof of the spherical Parisi formula, we want a lower bound on $\liminf_{N\to\infty} \mathbf{E}F_N$ involving $\mathrm{Par}(\xi, \beta = 1)$. We do so by taking the minimizing function f^* in the Crisanti-Sommers lower bound, connecting it to the geometry of the Gibbs measure, and bounding the individual components.

We predict that, for a Gibbs measure, we should be able to construct a tree (like we did with Subag's algorithm) with nodes on the subspheres of radii Q_d, where Q_d correspond to the breaking points between different components of f^*.

To show a lower bound, we can cut f^* into simple parts, prove sharp lower bounds for each part, and then combine these bounds/parts geometrically to conclude the proof.

Between the spheres of radii Q_d and Q_{d-1}, we can consider the band with smaller covariance ξ_d such that $f_*^{\xi_d}$ is approximately the same as f_*^{ξ} restricted to the interval $[Q_d, Q_{d+1}]$. For $\|\sigma\| = \sqrt{Q_d N}$, consider $\rho = \sigma + \sqrt{Q_{d+1} - Q_d}\tau$ and $\tilde{\rho} = \sqrt{Q_{d+1} - Q_d}\tilde{\tau}$ where $\tau, \tilde{\tau} \perp \sigma$ and $\|\tau\| = \|\tilde{\tau}\| = \sqrt{N}$. Then, the overlap is $R(\rho, \tilde{\rho}) = Q_d + (Q_{d+1} - Q_d)R(\tau, \tilde{\tau})$. Similar to our discussion of spin glasses at high temperature with external fields (Lecture 11), we can then write the effective covariance $\xi_d(R) = \xi(Q_d + (Q_{d+1} - Q_d)R) - (Q_{d+1} - Q_d)\xi'(Q_d)R$.

We've already discussed/established the ground state energy for the topologically trivial part (using the Kac-Rice in Lecture 8) and the ground state energy for the full RSB part (using our analysis of Subag's algorithm from Lecture 15). It therefore suffices to solve the 1RSB and RS cases.

3 RS case

Suppose that ξ is replica symmetric. Then, because $f_*(q) = 1$ for all $q \in [0,1]$, we have $\mathrm{Par}(\xi, \beta = 1) = \xi(1)/2 = \frac{1}{N}\log \mathbf{E}Z_N$, where the last equality follows from following the computations from Lecture 2.

We could then try to follow the second moment method computations from Lecture 3, which yields

$$\mathbf{E}Z_N^2 \approx \exp\left(N\left(\xi(1) + \max_q\left(\xi(q) + \frac{\log(1-q^2)}{2}\right)\right) + o(N)\right)$$

It follow that $\mathbf{E}[Z_n^2] \le \exp(\xi(1)N + o(N))$ if and only if $\xi(q) + \frac{\log(1-q^2)}{2} \le 0$ for all $q \in [0,1]$. Therefore, the second moment method yields the desired lower bound for RS ξ satisfying this condition.

To show that this lower bound holds for all RS models, we need to improve upon this second moment method. We'll do so by *truncation*. For a fixed (ϵ, δ), call $\sigma \in S_N$ typical (which we denote $\sigma \in T$) if

$$\frac{1}{N}\log \int_{\substack{\tilde{\sigma}\in S_N \\ |R(\sigma,\tilde{\sigma})|\geq \delta}} e^{H_N(\tilde{\sigma})}\, d\tilde{\sigma} \leq \frac{\xi(1)}{2} + \epsilon.$$

That is, $\sigma \in T$ if the points that are nearly orthgonal to σ account for an exponentially small amount of the free energy.

Now, if we define $\tilde{Z}_N = \int_{\sigma \in T} e^{H_N(\sigma)}\, d\sigma$ such that $\tilde{Z}_N \leq Z_N$, a nice bound on the second moment follows directly from the definition of T:

Proposition 3.1. $\mathbf{E}[\tilde{Z}_N^2] \leq \exp(N\xi(1) + O(\epsilon+\delta)N)$.

Proof. We can write

$$\begin{aligned}
\mathbf{E}[\tilde{Z}_N^2] &= \mathbf{E}\int_{T\times T} e^{H_N(\sigma)+H_N(\tilde{\sigma})}\, d\sigma\, d\tilde{\sigma} \\
&= \mathbf{E}\int_{\substack{\sigma,\tilde{\sigma}\in T \\ |R(\sigma,\tilde{\sigma})|\leq \delta}} e^{H_N(\sigma)+H_N(\tilde{\sigma})}\, d\sigma\, d\tilde{\sigma} + \mathbf{E}\int_{\substack{\sigma,\tilde{\sigma}\in T \\ |R(\sigma,\tilde{\sigma})|\geq \delta}} e^{H_N(\sigma)+H_N(\tilde{\sigma})}\, d\sigma\, d\tilde{\sigma}
\end{aligned}$$

We pick up the $O(\delta)$ term from the first integral and the $O(\epsilon)$ term from the second integral. Combining these two terms yields $\exp(N\xi(1) + O(\epsilon+\delta)N)$ as desired. □

Even though this truncation yields a nice second moment bound, we might be concerned that it kills the first moment. If, however, we can show that $\mathbf{E}[\tilde{Z}_N] \geq \mathbf{E}[Z_N]/2$, then by Paley-Zygmund, $\mathbb{P}[\tilde{Z}_N \geq e^{N\xi(1)/2 - o(N)}] \geq e^{-o(N)}$ and concentration of F_N allows us to establish the desired lower bound.

For a fixed $\sigma_0 \in S_N$, we can write

$$\frac{\mathbf{E}[\tilde{Z}_N]}{\mathbf{E}[Z_N]} = \frac{\mathbf{E}[e^{H_N(\sigma_0)} \cdot \mathbb{1}[\sigma_0 \in T]]}{\mathbf{E}[e^{H_N(\sigma_0)}]} = \tilde{\mathbb{P}}[\sigma_0 \in T]$$

where $\tilde{\mathbb{P}}$ is the tilted measure with $\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = e^{H_N(\sigma_0)}$ and $\mathbb{P} = \mathrm{Law}(H_N)$ Therefore, we can write

$$\mathrm{Law}_{\tilde{P}P}(H_N) \stackrel{d}{=} \mathrm{Law}_{\mathbb{P}}(H_N) + \sum_{p\geq 1} \gamma_p^2 \langle \sigma_0, \cdot\rangle^p$$

which implies that under $\mathbb{P}$, $H_N = \tilde{H}_N = \sum_{p\geq 1}\gamma_p^2 \sigma_0^{\otimes p}$ and $H_N(\sigma) = \tilde{H}_N(\sigma) + \xi(R(\sigma,\sigma_0))$. To show that $\tilde{\mathbb{P}}[\sigma_0 \in T] \geq 1/2$, it suffices to show that the restricted free energy of H_N is at most $\frac{\xi(1)}{2} + o(1)$.

Let $\rho = q\sigma_0 + \sqrt{1-q^2}\tau$ and $\tilde{\rho} = q\sigma_0 + \sqrt{1-q^2}\tilde{\tau}$ for $\tau, \tilde{\tau} \in S_N$ and $\tau, \tilde{\tau} \perp \sigma_0$. Then, we have

$$\mathbf{E}[\tilde{H}_N(\rho)\tilde{H}_N(\tilde{\rho})] = N\xi(q^2 + (1-q^2)R(\tau,\tilde{\tau})) \equiv N\xi_q(R(\tau,\tilde{\tau}))$$

from which we can express $\mathbf{E}F(H_N;\mathrm{Band}_q(\sigma_0)) = \xi(q) + \frac{\log(1-q^2)}{2} + \mathbf{E}F_N(\xi_q)$.

If we let $f(q) = \mathbb{1}_{q\geq t}$, then the interpolation upper bound yields

$$\limsup_{N\to\infty} \mathbf{E}F_N \leq \frac{1}{2}\left(\xi(1) - \xi(t) + \frac{t}{1-t} + \log(1-t)\right).$$

Consider taking $t = \frac{q}{q+1}$. We do this partially because the formulas are nice, but there's also a nice geometric interpretation: if $R(\tau,\tilde{\tau}) = \frac{q}{q+1}$, then $R(\rho,\tilde{\rho}) = q$. Using this in the upper bound, we get

$$\mathbf{E}F_N(\xi_q) \leq \frac{1}{2}(\xi_q(1) - \xi_q(t) + q - \log(1-q))$$

where $\xi_q(1) = \xi(1)$ and $\xi_q(t) = \xi(q)$. Therefore,

$$\mathbf{E}F(H_N;\mathrm{Band}_q(\sigma_0)) = \frac{1}{2}(\xi(1) + \xi(q) + q + \log(1-q))$$

Replica symmetric implies that $\xi(q) + q + \log(1-q) \leq 0$, so $\mathbf{E}F(H_N;\mathrm{Band}_q(\sigma_0)) \leq \frac{\xi(1)}{2}$.

4 1RSB case

Proving a lower bound in the 1RSB follows by a similar process. The following table summarizes the differences between the lower bounds in the RS and the 1RSB cases.

	RS	1RSB
bound needed	free energy F_N	ground state energy GS_N
2nd moment method on	Z_N	# of critical points w/ $\approx$ extremal energy
truncation	typical points	ground state typical points

Here, ground state typical points are points σ_0 such that $H_N(\sigma_0) \approx E_0 N + \epsilon n$, $\nabla_{\mathrm{sph}} H_N(\sigma_0) = 0$, and no other $\tilde{\sigma}$ has $H_N(\tilde{\sigma}) \geq (E_0 + \epsilon)N$.

5 Combining

Finally, to justify combining our lower bounds for each of the individual components of f^*, we need a form of uniform concentration, similar to the one from the analysis of Subag's algorithm. For $\|\sigma\| = \sqrt{Q_d N}$, if we let

$$f_{Q_d, Q_{d+1}, k}(\sigma) = \frac{1}{kN} \max_{\substack{\sigma^1, \dots, \sigma^k \\ \sigma \perp \sigma^1 - \sigma \cdots \perp \sigma^k - \sigma}} \sum_{i=1}^{k} H_N(\sigma^i) - H_N(\sigma)$$

then f concentrates uniformly for large k, which justifies the idea that the free energy of a pure state can be bounded by bounding the free energy along these paths.