

6.S891 Problem Set 3

Due: November 30, 2023 (11:59 PM EST)

Please turn in a PDF with solutions typed in \LaTeX . You may collaborate on this problem set, but please attempt the problems yourself first. It goes without saying, please do not try to look up solutions online. Feel free to use Wikipedia, etc. for material pertaining to reasonable course prerequisites (e.g. basic probability, linear algebra, etc.) If you find a bug in a problem, please let me know.

Exercise 1 (Hardcore Model on $K_{n,n}$). Consider the hardcore model with fugacity $\lambda \geq 0$ on the complete bipartite graph $K_{n,n}$. Let μ denote the Gibbs distribution, and $Z(\lambda)$ denote the partition function.

- (a) Give a closed form expression for $Z(\lambda)$.
- (b) Prove that there is a constant $C = C(\lambda) > 0$ such that for each $0 \leq k \leq 2n - 2$, there exists $S \subseteq V$ with $|S| = k$ and a pinning $\tau : S \rightarrow \{\text{in}, \text{out}\}$ such that $\lambda_{\max}(\Psi_{\mu^\tau}) \geq C \cdot (2n - k)$.
(If it makes life convenient, you may assume k is even.)
- (c) Show that for every fixed $\lambda \geq 0$ independent of n , there is a constant $C = C(\lambda) > 0$ such that Glauber dynamics has spectral gap at most $\exp(-Cn)$. In particular, its worst-case mixing time is exponentially large.
- (d) Conclude that the “local-to-global theorem”, i.e. that $O(1)$ -spectral independence for all pinnings implies inverse polynomial spectral gap, **cannot** admit an “average-case”¹ analog of the following form:

“Let μ be a probability distribution on $\{\pm 1\}^n$. Suppose for every $\Omega(\log n) \leq k \leq n - 2$, the conditional measure μ^τ is $O(1)$ -spectrally independent with high probability (e.g. $1 - \frac{1}{\text{poly}(n)}$) over a random pinning $\tau \sim \mu_S$ on a uniformly random set S of k coordinates. Then Glauber dynamics for μ has spectral gap at least $1/\text{poly}(n)$, and polynomial mixing time.”

Exercise 2 (Continuous to Discrete). Let μ be a probability measure on $\{\pm 1\}^n$.

- (a) Suppose μ can be decomposed as

$$\mu(\sigma) = \int_{\mathbb{R}^n} (\mathcal{T}_x \nu)(\sigma) d\xi(x), \quad \forall \sigma \in \{\pm 1\}^n,$$

for some distribution ν on $\{\pm 1\}^n$ and some distribution ξ on \mathbb{R}^n satisfying the following properties:

- **Poincaré Inequality for Mixture:** There is a constant $C > 0$ such that for every differentiable test function $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\text{Var}_{x \sim \xi}[\psi(x)] \leq \alpha \cdot \int_{\mathbb{R}^n} \|\nabla \psi(x)\|_2^2 d\xi(x). \tag{1}$$

- **Spectral Independence for Components:** There is a constant $\eta > 0$ such that for every $x \in \mathbb{R}^n$, the exponential tilt $\mathcal{T}_x \nu$ of ν is η -spectrally independent.

Prove that for every $x \in \mathbb{R}^n$, the exponential tilt $\mathcal{T}_x \mu$ of μ is $(\eta + C(\alpha, \eta))$ -spectrally independent for some constant $C(\alpha, \eta) > 0$.² (Note that since this holds for all exponential tilts, we also get entropic independence for all exponential tilts.)

¹There do exist other kinds of “average-case” analogs which are useful [ALV22; Ana+23].

²One can take $C(\alpha, \eta) = \alpha(1 + \eta)^2$.

(b) Let A be a symmetric positive definite matrix. Show that

$$\exp\left(\frac{1}{2}\sigma^\top A\sigma\right) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\|x\|_2^2 + \langle A^{1/2}x, \sigma \rangle\right) dx$$

for every $\sigma \in \mathbb{R}^n$.³

(c) Let $\mu(\sigma) \propto \exp\left(\frac{1}{2}\sigma^\top A\sigma\right)$ be the Gibbs distribution over $\{\pm 1\}^n$ of an Ising model with interaction matrix $A \in \mathbb{R}^{n \times n}$ satisfying $\frac{\epsilon}{2} \cdot \text{Id} \preceq A \preceq (1 - \frac{\epsilon}{2}) \cdot \text{Id}$ for some constant $0 < \epsilon < 1$. Note that we can always shift A by a diagonal matrix without affecting the distribution μ , and so we can assume A is positive (semi)definite without loss of generality; the key here is ensuring a bound on the spectral diameter $\lambda_{\max}(A) - \lambda_{\min}(A) \leq 1 - \epsilon$. Prove that $\mathcal{T}_x \mu$ is $O(1/\epsilon)$ -spectrally independent for all $x \in \mathbb{R}^n$.

Since the emphasis here is on the connection between the continuous and discrete, you may use the following fact from the continuous world without proof:

Theorem 0.1. Let $\xi(x) \propto \exp(-V(x))$ be an α -strongly log-concave probability distribution on \mathbb{R}^n , i.e. for some constant $\alpha > 0$, $\nabla^2 V(x) \succeq \alpha \cdot \text{Id}$ for all $x \in \mathbb{R}^n$. Then ξ satisfies the Poincaré Inequality⁴ Eq. (1) with constant $C = \frac{1}{\alpha}$.

(d) **The Sherrington–Kirkpatrick Spin Glass:** Let $G \sim \text{GOE}(n)$, i.e. $G_{ij} \sim \mathcal{N}(0, 1/n)$ i.i.d. for $i < j$, $G_{ji} = G_{ij}$ for $i > j$, and $G_{ii} = 0$ for $i \in [n]$. Prove that if $\beta < \frac{1-\epsilon}{4}$ for a constant $0 < \epsilon < 1$,⁵ then with high probability over the randomness of G , all exponential tilts of the Ising Gibbs measure with interaction matrix $A = \beta G$ are $O(1/\epsilon)$ -spectrally independent.⁶ Conclude that w.h.p. over G , Glauber dynamics mixes in polynomial-time.⁷ On the other hand, what is the largest value of β for which you can make path coupling contract?

For this problem, you may use the following standard fact from random matrix theory, which can be proved using the trace moment method.

Theorem 0.2. With high probability over $G \sim \text{GOE}(n)$, we have $-2(1 + o(1)) \cdot \text{Id} \preceq G \preceq 2(1 + o(1)) \cdot \text{Id}$.

³The “fancy name” for this is the *Hubbard–Stratonovich transform*. Up to scaling, the right-hand side is just a moment generating function.

⁴Under this strong log-concavity assumption, a log-Sobolev inequality also holds. This is a consequence of the beautiful Bakry–Émery theory.

⁵The phase transition threshold is $\beta = 1$, not $1/4$. Obtaining an FPAS up to this threshold remains an outstanding open problem as of this writing. Weaker samplers are known [AMS22].

⁶It is known that for this distribution, we have $\|\Psi_\mu\|_{\ell_\infty \rightarrow \ell_\infty} \gtrsim \sqrt{n}$, roughly for the same reason that $\|G\|_{\ell_\infty \rightarrow \ell_\infty} \gtrsim \sqrt{n}$.

⁷The naïve bound one would get for the mixing time is $n^{1/\epsilon}$, although using concentration bounds for the norms of random submatrices, one can bring this down to $C^{1/\epsilon} n^c$ for (reasonable) universal constants $C, c > 0$. Using stochastic localization, one can actually establish the optimal $O(n \log n)$ -mixing.

References

- [ALV22] Nima Anari, Yang P. Liu, and Thuy-Duong Vuong. “Optimal Sublinear Sampling of Spanning Trees and Determinantal Point Processes via Average-Case Entropic Independence”. In: *arXiv preprint arXiv:2204.02570* (2022). To Appear in FOCS 2022 (cit. on p. 1).
- [AMS22] A. Alaoui, A. Montanari, and M. Sellke. “Sampling from the Sherrington-Kirkpatrick Gibbs measure via algorithmic stochastic localization”. In: *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*. Los Alamitos, CA, USA: IEEE Computer Society, Nov. 2022, pp. 323–334. DOI: [10.1109/FOCS54457.2022.00038](https://doi.org/10.1109/FOCS54457.2022.00038) (cit. on p. 2).
- [Ana+23] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. “Universality of Spectral Independence with Applications to Fast Mixing in Spin Glasses”. In: *arXiv preprint arXiv:2307.10466* (2023) (cit. on p. 1).