Optimality of Glauber dynamics for general-purpose Ising model sampling and free energy approximation

Dmitriy (Tim) Kunisky

Department of Computer Science, Yale University

SODA 2024
January 10, 2024

Problems from Statistical Physics

Ising-type models: "Energy" / "Hamiltonian" / "objective" on $\{\pm 1\}^n$, $\boldsymbol{x} \mapsto \boldsymbol{x}^\top \boldsymbol{J} \boldsymbol{x}$.

Problems from Statistical Physics

Ising-type models: "Energy" / "Hamiltonian" / "objective" on $\{\pm 1\}^n$, $\boldsymbol{x} \mapsto \boldsymbol{x}^\top \boldsymbol{J} \boldsymbol{x}$.

Objects of interest: **Gibbs measure** and **partition function**,

$$\mu_{\boldsymbol{J}}(\boldsymbol{x}) := \frac{1}{Z(\boldsymbol{J})} \exp(\boldsymbol{x}^\top \boldsymbol{J} \boldsymbol{x}),$$

$$Z(\boldsymbol{J}) := \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\boldsymbol{x}^\top \boldsymbol{J} \boldsymbol{x}).$$

Problems from Statistical Physics

Ising-type models: "Energy" / "Hamiltonian" / "objective" on $\{\pm 1\}^n$, $\boldsymbol{x} \mapsto \boldsymbol{x}^\top \boldsymbol{J} \boldsymbol{x}$.

Objects of interest: **Gibbs measure** and **partition function**,

$$\mu_{\boldsymbol{J}}(\boldsymbol{x}) := \frac{1}{Z(\boldsymbol{J})} \exp(\boldsymbol{x}^\top \boldsymbol{J} \boldsymbol{x}),$$
$$Z(\boldsymbol{J}) := \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\boldsymbol{x}^\top \boldsymbol{J} \boldsymbol{x}).$$

Combinatorial heuristic: $\boldsymbol{J} = \eta \cdot$ (Laplacian of graph), then

$$\boldsymbol{x}^\top \boldsymbol{J} \boldsymbol{x} = 4\eta \cdot \#\{\text{ edges cut by } \boldsymbol{x} \},$$
$$\mu_{\eta \boldsymbol{J}} \approx \text{uniform measure on cuts of some size},$$
$$Z(\eta \boldsymbol{J}) \approx C \cdot \#\{\text{ cuts of some size }\}.$$

Glauber Dynamics

Simple Markov Chain Monte Carlo algorithm for sampling from $\mu_{\boldsymbol{J}}$:

0. Initialize $\boldsymbol{x} = \boldsymbol{x}^{(0)}$.
1. Choose $i \sim \mathsf{Unif}([n])$.
2. Resample $x_i \in \{\pm 1\}$ according to conditional distribution.
3. Go to Step 1.

Glauber Dynamics

Simple Markov Chain Monte Carlo algorithm for sampling from $\mu_{\boldsymbol{J}}$:

0. Initialize $\boldsymbol{x} = \boldsymbol{x}^{(0)}$.
1. Choose $i \sim \mathsf{Unif}([n])$.
2. Resample $x_i \in \{\pm 1\}$ according to conditional distribution.
3. Go to Step 1.

$$\boldsymbol{x}^{(k)} = \begin{bmatrix} +1 & -1 & +1 & +1 & -1 & -1 & -1 & +1 \end{bmatrix}$$

Glauber Dynamics

Simple Markov Chain Monte Carlo algorithm for sampling from $\mu_{\boldsymbol{J}}$:

0. Initialize $\boldsymbol{x} = \boldsymbol{x}^{(0)}$.
1. Choose $i \sim \mathsf{Unif}([n])$.
2. Resample $x_i \in \{\pm 1\}$ according to conditional distribution.
3. Go to Step 1.

$$\boldsymbol{x}^{(k+1)} \leftarrow \begin{bmatrix} +1 & -1 & +1 & +1 & \pm\mathbf{1} & -1 & -1 & +1 \end{bmatrix}$$

Glauber Dynamics

Simple Markov Chain Monte Carlo algorithm for sampling from μ_J:

0. Initialize $\boldsymbol{x} = \boldsymbol{x}^{(0)}$.
1. Choose $i \sim \mathsf{Unif}([n])$.
2. Resample $x_i \in \{\pm 1\}$ according to conditional distribution.
3. Go to Step 1.

$$\boldsymbol{x}^{(k+1)} \leftarrow \begin{bmatrix} +1 & -1 & +1 & +1 & \pm 1 & -1 & -1 & +1 \end{bmatrix}$$

Stationary distribution = μ_J $\rightsquigarrow$ always works eventually.

Question: For what $\boldsymbol{J}$ can we guarantee it mixes rapidly?

Conditions for Fast Mixing

Classical **Dobrushin uniqueness condition**: enough for $\sum_{j=1}^{n} |J_{ij}| < 1$ for all $i \in [n]$.

Works for, e.g., $\boldsymbol{J} = \eta \cdot$ (Laplacian of sparse graph).

Conditions for Fast Mixing

Classical **Dobrushin uniqueness condition**: enough for $\sum_{j=1}^{n} |J_{ij}| < 1$ for all $i \in [n]$.

Works for, e.g., $\boldsymbol{J} = \eta \cdot$ (Laplacian of sparse graph).

Does not work for dense $\boldsymbol{J}$, e.g., i.i.d. Gaussian entries $J_{ij} \sim \mathcal{N}(0, \frac{\sigma^2}{n})$. ("Sherrington-Kirkpatrick spin glass model")

Conditions for Fast Mixing

Classical **Dobrushin uniqueness condition**: enough for $\sum_{j=1}^{n} |J_{ij}| < 1$ for all $i \in [n]$.

Works for, e.g., $\boldsymbol{J} = \eta \cdot$ (Laplacian of sparse graph).

Does not work for dense $\boldsymbol{J}$, e.g., i.i.d. Gaussian entries $J_{ij} \sim \mathcal{N}(0, \frac{\sigma^2}{n})$. ("Sherrington-Kirkpatrick spin glass model")

Several recent works making progress on this case. The result we will focus on:

Theorem: [Eldan, Koehler, Zeitouni '20] Output of Glauber dynamics is close to μ after $\mathcal{O}(n^2/\delta)$ steps so long as

$$\mathsf{width}(\boldsymbol{J}) := \lambda_{\mathsf{max}}(\boldsymbol{J}) - \lambda_{\mathsf{min}}(\boldsymbol{J}) \leq \frac{1}{2} - \delta.$$

Is "$\mathrm{width}(\boldsymbol{J}) < \frac{1}{2}$" Optimal?

Theorem: [Griffiths, Weng, Langer '66] Glauber dynamics does not mix rapidly when $\boldsymbol{J} = (\frac{1}{2} + \varepsilon)\frac{1}{n}\mathbf{1}\mathbf{1}^\top$. ("Curie-Weiss model")

$\rightsquigarrow$ width condition cannot be improved to $\frac{1}{2} + \varepsilon$.

Is "width(J) $< \frac{1}{2}$" Optimal?

Theorem: [Griffiths, Weng, Langer '66] Glauber dynamics does not mix rapidly when $J = (\frac{1}{2} + \varepsilon)\frac{1}{n}\mathbf{1}\mathbf{1}^\top$. ("Curie-Weiss model")

$\rightsquigarrow$ width condition cannot be improved to $\frac{1}{2} + \varepsilon$.

What about other efficient samplers, more tailored to J?

Theorem: [El Alaoui, Montanari, Sellke '22; Celentano '22] For the Gaussian Sherrington-Kirkpatrick model $J_{ij} \sim \mathcal{N}(0, \frac{\sigma^2}{n})$, a different efficient sampler ("algorithmic stochastic localization") succeeds whenever width(J) $\leq 2 - \delta$.

Is "width$(\boldsymbol{J}) < \frac{1}{2}$" Optimal?

Theorem: [Griffiths, Weng, Langer '66] Glauber dynamics does not mix rapidly when $\boldsymbol{J} = (\frac{1}{2} + \varepsilon)\frac{1}{n}\mathbf{1}\mathbf{1}^\top$. ("Curie-Weiss model")

⇝ width condition cannot be improved to $\frac{1}{2} + \varepsilon$.

What about other efficient samplers, more tailored to $\boldsymbol{J}$?

Theorem: [El Alaoui, Montanari, Sellke '22; Celentano '22] For the Gaussian Sherrington-Kirkpatrick model $J_{ij} \sim \mathcal{N}(0, \frac{\sigma^2}{n})$, a different efficient sampler ("algorithmic stochastic localization") succeeds whenever width$(\boldsymbol{J}) \leq 2 - \delta$.

Question: Can we remove the model dependence from this result? That is, does there exist a **general-purpose sampler** that improves on the width condition?

Lower Bounds

Theorem: [Sly, Sun '12; Galanis, Štefankovič, Vigoda '16] If the width condition can be improved to $1 + \varepsilon$, then $\mathsf{NP} = \mathsf{RP}$ (implicitly, via case of Laplacians of regular graphs).

Lower Bounds

Theorem: [Sly, Sun '12; Galanis, Štefankovič, Vigoda '16] If the width condition can be improved to $1 + \varepsilon$, then $\mathsf{NP} = \mathsf{RP}$ (implicitly, via case of Laplacians of regular graphs).

Conjecture: [Bandeira, K., Wein '20] It is hard to distinguish:

$\mathbb{Q}$: A uniformly random γn-dimensional subspace $\subset \mathbb{R}^n$.

$\mathbb{P}$: A random subspace biased towards $\boldsymbol{x}^\star \sim \mathsf{Unif}(\{\pm 1\}^n)$.

Theorem: (this paper) If the width condition can be improved to $\frac{1}{2} + \varepsilon$, then this Conjecture is false.

Glauber dynamics is an optimal (in width condition) general-purpose sampler.

Evidence for Conjecture: Gaussian Model

$$\boldsymbol{x}^\star \text{ close to } V \subset \mathbb{R}^n \quad \leftrightarrow \quad \boldsymbol{x}^\star \text{ almost orthogonal to } V^\perp.$$

Evidence for Conjecture: Gaussian Model

$\boldsymbol{x}^\star$ close to $V \subset \mathbb{R}^n \quad \leftrightarrow \quad \boldsymbol{x}^\star$ almost orthogonal to $V^\perp$.

We will model $V^\perp := \mathsf{span}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{\alpha n})$, $\alpha = 1 - \gamma$:

$\mathbb{Q}$: $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{\alpha n} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_n)$.

$\mathbb{P}$: $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{\alpha n} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_n - \frac{\beta}{n}\boldsymbol{x}^\star \boldsymbol{x}^{\star\top})$ for $\boldsymbol{x}^\star \sim \mathsf{Unif}(\{\pm 1\}^n)$.

Evidence for Conjecture: Gaussian Model

$\boldsymbol{x}^\star$ close to $V \subset \mathbb{R}^n \quad \leftrightarrow \quad \boldsymbol{x}^\star$ almost orthogonal to $V^\perp$.

We will model $V^\perp := \mathsf{span}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{\alpha n})$, $\alpha = 1 - \gamma$:

$\mathbb{Q}$: $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{\alpha n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$.

$\mathbb{P}$: $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{\alpha n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n - \frac{\beta}{n}\boldsymbol{x}^\star \boldsymbol{x}^{\star\top})$ for $\boldsymbol{x}^\star \sim \mathsf{Unif}(\{\pm 1\}^n)$.

Variant of the classical **spiked matrix model:** [Johnstone '01]

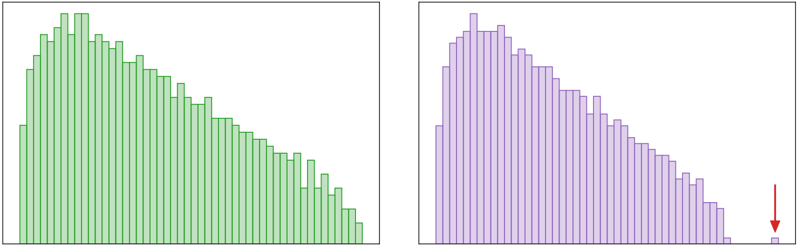
$\mathbb{Q}$: $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{\alpha n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$.

$\mathbb{P}'$: $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{\alpha n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n + \frac{\beta}{n}\boldsymbol{x}^\star \boldsymbol{x}^{\star\top})$ for $\boldsymbol{x}^\star \sim \mathsf{Unif}(\{\pm 1\}^n)$.

Conjecture: It is hard to distinguish in either case, for any choice of $\alpha, \beta \in (0, 1)$.

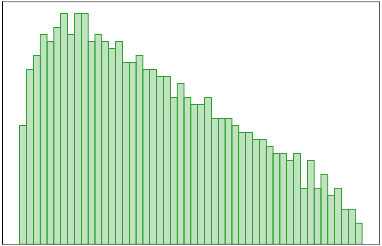
Evidence for Conjecture: BBP Transition

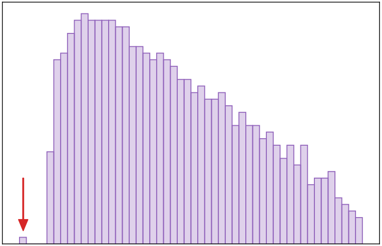
Natural algorithm: threshold extreme eigenvalue of sample covariance, $\lambda_{\mathsf{min}|\mathsf{max}}(\frac{1}{\alpha n}\sum_{i=1}^{\alpha n} \boldsymbol{y}_i \boldsymbol{y}_i^\top)$.



Evidence for Conjecture: BBP Transition

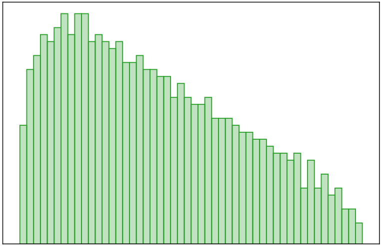
Natural algorithm: threshold extreme eigenvalue of sample covariance, $\lambda_{\mathsf{min}|\mathsf{max}}(\frac{1}{\alpha n}\sum_{i=1}^{\alpha n} \boldsymbol{y}_i \boldsymbol{y}_i^\top)$.

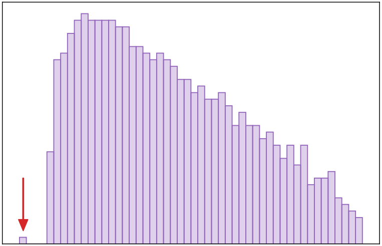




Evidence for Conjecture: BBP Transition

Natural algorithm: threshold extreme eigenvalue of sample covariance, $\lambda_{\mathsf{min|max}}(\frac{1}{\alpha n}\sum_{i=1}^{\alpha n} \boldsymbol{y}_i \boldsymbol{y}_i^\top)$.

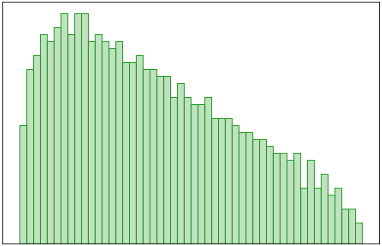


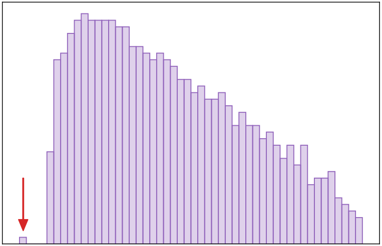


Theorem: [Baik, Ben Arous, Péché '05], [Baik, Silverstein '06] In either case, this works $\iff$ $\alpha\beta^2 > 1$.

Evidence for Conjecture: BBP Transition

Natural algorithm: threshold extreme eigenvalue of sample covariance, $\lambda_{\min|\max}(\frac{1}{\alpha n}\sum_{i=1}^{\alpha n} \boldsymbol{y}_i \boldsymbol{y}_i^\top)$.





Theorem: [Baik, Ben Arous, Péché '05], [Baik, Silverstein '06] In either case, this works $\Leftrightarrow$ $\alpha\beta^2 > 1$.

Theorem: [Bandeira, K., Wein '20] More general algorithms from **low-degree polynomials** work $\Leftrightarrow$ $\alpha\beta^2 > 1$.

Proof Sketch I: Relate to Partition Function

Suppose for the sake of contradiction an efficient sampler works for all $\mathsf{width}(\boldsymbol{J}) < \frac{1}{2} + \varepsilon$.

Proof Sketch I: Relate to Partition Function

Suppose for the sake of contradiction an efficient sampler works for all $\mathrm{width}(\boldsymbol{J}) < \frac{1}{2} + \varepsilon$.

Reduce "counting to sampling," approximating $Z(\boldsymbol{J})$ by fixing $0 < \eta_1 < \cdots < \eta_k < 1$ and **annealing**:

$$Z(\boldsymbol{J}) = Z(\boldsymbol{0}) \cdot \frac{Z(\eta_1 \boldsymbol{J})}{Z(\boldsymbol{0})} \cdot \frac{Z(\eta_2 \boldsymbol{J})}{Z(\eta_1 \boldsymbol{J})} \cdots \frac{Z(\boldsymbol{J})}{Z(\eta_k \boldsymbol{J})},$$

$$Z(\boldsymbol{0}) = 2^n,$$

$$\frac{Z(\eta_{i+1}\boldsymbol{J})}{Z(\eta_i \boldsymbol{J})} = \frac{1}{Z(\eta_i \boldsymbol{J})} \sum_{\boldsymbol{x}} \exp(\boldsymbol{x}^\top \eta_i \boldsymbol{J} \boldsymbol{x}) \exp(\boldsymbol{x}^\top (\eta_{i+1} - \eta_i) \boldsymbol{J} \boldsymbol{x})$$

$$= \mathop{\mathbb{E}}_{\boldsymbol{x} \sim \mu_{\eta_i \boldsymbol{J}}} \exp(\boldsymbol{x}^\top (\eta_{i+1} - \eta_i) \boldsymbol{J} \boldsymbol{x})$$

Proof Sketch I: Relate to Partition Function

Suppose for the sake of contradiction an efficient sampler works for all $\mathrm{width}(\boldsymbol{J}) < \frac{1}{2} + \varepsilon$.

Reduce "counting to sampling," approximating $Z(\boldsymbol{J})$ by fixing $0 < \eta_1 < \cdots < \eta_k < 1$ and **annealing**:

$$Z(\boldsymbol{J}) = Z(\boldsymbol{0}) \cdot \frac{Z(\eta_1 \boldsymbol{J})}{Z(\boldsymbol{0})} \cdot \frac{Z(\eta_2 \boldsymbol{J})}{Z(\eta_1 \boldsymbol{J})} \cdots \frac{Z(\boldsymbol{J})}{Z(\eta_k \boldsymbol{J})},$$

$$Z(\boldsymbol{0}) = 2^n,$$

$$\frac{Z(\eta_{i+1}\boldsymbol{J})}{Z(\eta_i \boldsymbol{J})} = \frac{1}{Z(\eta_i \boldsymbol{J})} \sum_{\boldsymbol{x}} \exp(\boldsymbol{x}^\top \eta_i \boldsymbol{J} \boldsymbol{x}) \exp(\boldsymbol{x}^\top (\eta_{i+1} - \eta_i) \boldsymbol{J} \boldsymbol{x})$$

$$= \mathop{\mathbb{E}}_{\boldsymbol{x} \sim \mu_{\eta_i \boldsymbol{J}}} \exp(\boldsymbol{x}^\top (\eta_{i+1} - \eta_i) \boldsymbol{J} \boldsymbol{x})$$

⇝ can estimate $\exp(-\delta n) Z(\boldsymbol{J}) \le \hat{Z}(\boldsymbol{J}) \le \exp(\delta n) Z(\boldsymbol{J})$.

Proof Sketch II: Design Hypothesis Test

To solve hypothesis testing problem for $\boldsymbol{P}_V$ projection to subspace, try thresholding for large η

$$Z(\eta \boldsymbol{P}_V) = \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\boldsymbol{x}^\top \eta \boldsymbol{P}_V \boldsymbol{x}) = \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \|\boldsymbol{P}_V \boldsymbol{x}\|^2).$$

Proof Sketch II: Design Hypothesis Test

To solve hypothesis testing problem for $\boldsymbol{P}_V$ projection to subspace, try thresholding for large η

$$Z(\eta \boldsymbol{P}_V) = \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\boldsymbol{x}^\top \eta \boldsymbol{P}_V \boldsymbol{x}) = \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \|\boldsymbol{P}_V \boldsymbol{x}\|^2).$$

Idea: For large enough η,

$$\frac{1}{\eta} \log Z(\eta \boldsymbol{P}_V) \approx \max_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{P}_V \boldsymbol{x}\|^2,$$

which distinguishes $V \sim \mathbb{Q}$ from $V \sim \mathbb{P}$.

Question: How large does η need to be to separate these?

Proof Sketch III: Planted Calculations

$V \sim \mathbb{P}:$ with $\boldsymbol{x}^\star$ close to V, lower bound

$$\|\boldsymbol{P}_V \boldsymbol{x}\|^2 \geq \frac{\langle \boldsymbol{x}, \boldsymbol{x}^\star \rangle^2}{n}$$

Proof Sketch III: Planted Calculations

$V \sim \mathbb{P}$: with $\boldsymbol{x}^\star$ close to V, lower bound

$$\|\boldsymbol{P}_V \boldsymbol{x}\|^2 \geq \frac{\langle \boldsymbol{x}, \boldsymbol{x}^\star \rangle^2}{n}$$

with which we can reduce to combinatorics,

$$\begin{aligned} Z(\eta \boldsymbol{P}_V) &\gtrsim \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp\left(\eta \frac{\langle \boldsymbol{x}, \boldsymbol{x}^\star \rangle^2}{n}\right) \\ &= \sum_{r=0}^{n} \binom{n}{r} \exp\left(\eta \frac{(2r-n)^2}{n}\right) \\ &\approx \exp\left(n \cdot \max_{\rho = \frac{r}{n} \in [0,1]} \left\{ H(\rho) + \eta(2\rho - 1)^2 \right\}\right) \end{aligned}$$

$$V \sim \mathbb{P} \quad \Rightarrow \quad Z(\eta \boldsymbol{P}_V) \gtrsim \exp(n \cdot f_{\mathbb{P}}(\eta))$$

Proof Sketch IV: Null Calculations

$V \sim \mathbb{Q}$: Use rotational invariance and Jensen's inequality,

$$\begin{aligned}\mathop{\mathbb{E}}_{V} \log Z(\eta \boldsymbol{P}_V) &= \mathop{\mathbb{E}}_{V} \log \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \boldsymbol{x}^\top \boldsymbol{P}_V \boldsymbol{x}) \\ &= K + \mathop{\mathbb{E}}_{V} \mathop{\mathbb{E}}_{\boldsymbol{Q} \text{ orth.}} \log \mathop{\mathbb{E}}_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \| \boldsymbol{x}^\top \boldsymbol{Q}^\top \boldsymbol{P}_V \boldsymbol{Q} \boldsymbol{x} \|^2)\end{aligned}$$

Proof Sketch IV: Null Calculations

$V \sim \mathbb{Q}$: Use rotational invariance and Jensen's inequality,

$$\begin{aligned}
\mathop{\mathbb{E}}_{V} \log Z(\eta \boldsymbol{P}_V) &= \mathop{\mathbb{E}}_{V} \log \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \boldsymbol{x}^\top \boldsymbol{P}_V \boldsymbol{x}) \\
&= K + \mathop{\mathbb{E}}_{V} \mathop{\mathbb{E}}_{\boldsymbol{Q} \text{ orth.}} \log \mathop{\mathbb{E}}_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \| \boldsymbol{x}^\top \boldsymbol{Q}^\top \boldsymbol{P}_V \boldsymbol{Q} \boldsymbol{x} \|^2) \\
&\leq K + \mathop{\mathbb{E}}_{V} \log \mathop{\mathbb{E}}_{\boldsymbol{x} \in \mathbb{S}^{n-1}(\sqrt{n})} \exp(\eta \boldsymbol{x}^\top \boldsymbol{P}_V \boldsymbol{x})
\end{aligned}$$

Proof Sketch IV: Null Calculations

$V \sim \mathbb{Q}$: Use rotational invariance and Jensen's inequality,

$$
\begin{aligned}
\mathop{\mathbb{E}}_{V} \log Z(\eta \boldsymbol{P}_V) &= \mathop{\mathbb{E}}_{V} \log \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \boldsymbol{x}^\top \boldsymbol{P}_V \boldsymbol{x}) \\
&= K + \mathop{\mathbb{E}}_{V} \mathop{\mathbb{E}}_{\boldsymbol{Q} \text{ orth.}} \log \mathop{\mathbb{E}}_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \|\boldsymbol{x}^\top \boldsymbol{Q}^\top \boldsymbol{P}_V \boldsymbol{Q} \boldsymbol{x}\|^2) \\
&\leq K + \mathop{\mathbb{E}}_{V} \log \mathop{\mathbb{E}}_{\boldsymbol{x} \in \mathbb{S}^{n-1}(\sqrt{n})} \exp(\eta \boldsymbol{x}^\top \boldsymbol{P}_V \boldsymbol{x}) \\
&= K + \log \mathop{\mathbb{E}}_{\boldsymbol{x} \in \mathbb{S}^{n-1}(\sqrt{n})} \exp(\eta (x_1^2 + \cdots + x_{\gamma n}^2)) \\
&= \text{“spherical integral” with clean asymptotics}
\end{aligned}
$$

Proof Sketch IV: Null Calculations

$V \sim \mathbb{Q}$: Use rotational invariance and Jensen's inequality,

$$\begin{aligned}
\mathop{\mathbb{E}}_{V} \log Z(\eta \boldsymbol{P}_V) &= \mathop{\mathbb{E}}_{V} \log \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \boldsymbol{x}^\top \boldsymbol{P}_V \boldsymbol{x}) \\
&= K + \mathop{\mathbb{E}}_{V} \mathop{\mathbb{E}}_{\boldsymbol{Q} \text{ orth.}} \log \mathop{\mathbb{E}}_{\boldsymbol{x} \in \{\pm 1\}^n} \exp(\eta \|\boldsymbol{x}^\top \boldsymbol{Q}^\top \boldsymbol{P}_V \boldsymbol{Q} \boldsymbol{x}\|^2) \\
&\leq K + \mathop{\mathbb{E}}_{V} \log \mathop{\mathbb{E}}_{\boldsymbol{x} \in \mathbb{S}^{n-1}(\sqrt{n})} \exp(\eta \boldsymbol{x}^\top \boldsymbol{P}_V \boldsymbol{x}) \\
&= K + \log \mathop{\mathbb{E}}_{\boldsymbol{x} \in \mathbb{S}^{n-1}(\sqrt{n})} \exp(\eta (x_1^2 + \cdots + x_{\gamma n}^2)) \\
&= \text{"spherical integral" with clean asymptotics}
\end{aligned}$$

$$V \sim \mathbb{Q} \quad \Rightarrow \quad Z(\eta \boldsymbol{P}_V) \lesssim \exp(n \cdot f_{\mathbb{Q}}(\eta))$$

Proof Sketch V: Finding a Threshold

$$V \sim \mathbb{P} \quad \Rightarrow \quad Z(\eta \boldsymbol{P}_V) \gtrsim \exp(n \cdot f_{\mathbb{P}}(\eta)) \tag{1}$$

$$V \sim \mathbb{Q} \quad \Rightarrow \quad Z(\eta \boldsymbol{P}_V) \lesssim \exp(n \cdot f_{\mathbb{Q}}(\eta)) \tag{2}$$

$$\text{calculus} \quad \rightsquigarrow \quad f_{\mathbb{P}}(\eta) > f_{\mathbb{Q}}(\eta) \text{ for all } \eta > \frac{1}{2} \tag{3}$$

Proof Sketch V: Finding a Threshold

$$V \sim \mathbb{P} \quad \Rightarrow \quad Z(\eta \boldsymbol{P}_V) \gtrsim \exp(n \cdot f_{\mathbb{P}}(\eta)) \qquad (1)$$

$$V \sim \mathbb{Q} \quad \Rightarrow \quad Z(\eta \boldsymbol{P}_V) \lesssim \exp(n \cdot f_{\mathbb{Q}}(\eta)) \qquad (2)$$

$$\text{calculus} \quad \rightsquigarrow \quad f_{\mathbb{P}}(\eta) > f_{\mathbb{Q}}(\eta) \text{ for all } \eta > \frac{1}{2} \qquad (3)$$

$$\text{sampler for } \mathsf{width}(\boldsymbol{J}) < \frac{1}{2} + \varepsilon$$

$$\downarrow$$

$$\text{partition function approximator } \widehat{Z} \text{ for } \mathsf{width}(\boldsymbol{J}) < \frac{1}{2} + \varepsilon$$

$$\downarrow$$

$$\text{hypothesis test by } V \mapsto \widehat{Z}\left(\left(\frac{1}{2} + \frac{\varepsilon}{2}\right)\boldsymbol{P}_V\right) \approx Z\left(\left(\frac{1}{2} + \frac{\varepsilon}{2}\right)\boldsymbol{P}_V\right)$$

Next Steps

- Stronger evidence (e.g., concluding $\mathsf{NP} = \mathsf{RP}$ instead of refuting Conjecture)
- Potts models (cut $\rightsquigarrow$ coloring, vector $\rightsquigarrow$ subspace)
- Other hypothesis testing problems for which a partition function is a good (and convenient to analyze) statistic

Thank you!