Trickle-down and new results for Glauber on SK-Ising and related models

Thuy-Duong "June" Vuong

UC Berkeley

Based on joint work with Nima Anari, Frederic Koehler

Ising model

Probability distribution $\mu_{J,h}\colon \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$

where $\log \mu_{J,h} \equiv H$ is quadratic

Parameters: interaction matrix $J \in \mathbb{R}^{n\times n}$, external field h $\in \mathbb{R}^n$

$$H^{J,h}(\sigma) = \frac{1}{2}\langle \sigma, J\sigma \rangle + \langle h, \sigma \rangle = \sum_{i\leq j} J_{ij}\sigma_i\sigma_j + \sum_i h_i\sigma_i$$

$$\mu_{J,h}(\sigma) = \frac{\exp\left(H^{J,h}(\sigma)\right)}{Z}$$

Ising model

Probability distribution $\mu_{J,h}\colon \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$

where $\log \mu_{J,h} \equiv H$ is quadratic

Parameters: interaction matrix $J \in \mathbb{R}^{n\times n}$, external field h $\in \mathbb{R}^n$

$$H^{J,h}(\sigma) = \frac{1}{2}\langle \sigma, J\sigma \rangle + \langle h, \sigma \rangle = \sum_{i\leq j} J_{ij}\sigma_i\sigma_j + \sum_i h_i\sigma_i$$

Sherrington-Kirkpatrick (SK):

J is (scaled) GOE matrix i.e.

$\forall i < j\colon J_{ji} = J_{ij} \sim \frac{\beta}{\sqrt{n}}\mathcal{N}(0,1)$

Ising model

Probability distribution with density function:

$$\mu \propto \exp(H): \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$$

H is quadratic

$$H^{J,h}(\sigma) = \frac{1}{2}\langle x, Jx \rangle + \langle h, x \rangle = \sum_{i \leq j} J_{ij}\sigma_i\sigma_j + \sum_i h_i\sigma_i$$

Motivation:

- Physics [Ising-Lenz'1920-1925, Sherrington-Kirkpatrick'1975, Parisi'1979]
- Bayesian inference in high-dim regression [DAM17, LM19, MV21,MW24]
- Spiked matrix model from PCA [Joh01],
- Stochastic block model [Sin11,DAM17,AMM+18]

Ising model

Probability distribution with density function:

$$\mu \propto \exp(H) \colon \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$$

H is quadratic

$$H^{J,h}(\sigma) = \frac{1}{2}\langle x, Jx \rangle + \langle h, x \rangle = \sum_{i \leq j} J_{ij}\sigma_i\sigma_j + \sum_i h_i\sigma_i$$

Motivation:

- Physics [Ising-Lenz’1920-1925, Sherrington-Kirkpatrick’1975, Parisi’1979,…]
- Bayesian inference in high-dim regression [DAM17, LM19, MV21,MW24]
- Spiked matrix model from PCA [Joh01],
- Stochastic block model [Sin11,DAM17,AMM+18]

Computer Science

Motivation: Bayesian inference in high-dim regression [DAM17, LM19, MV21,MW24]

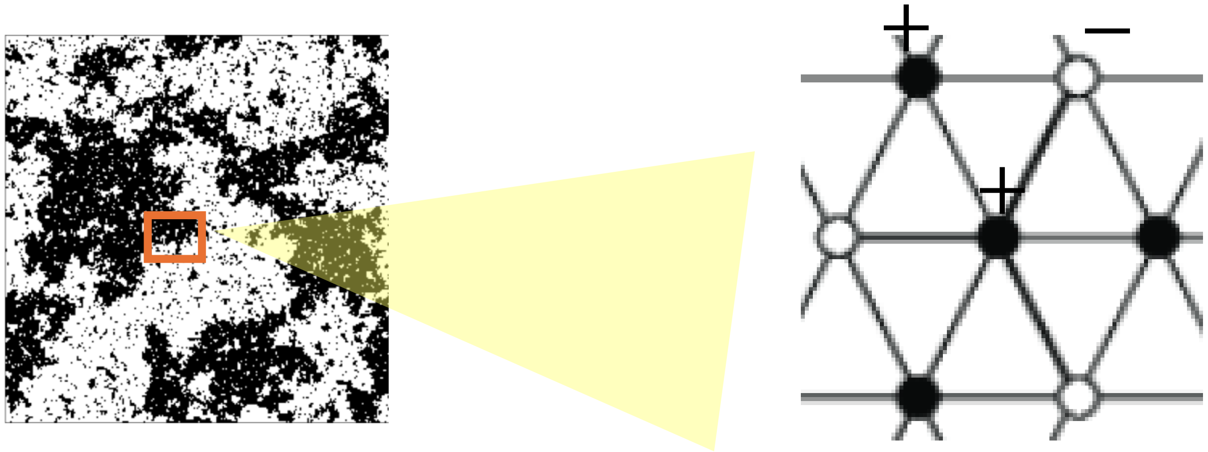
Given observation $y_0 = X\Theta + Gaussian(0, \sigma^2 I)$,
the Bayesian estimator for Θ with prior Uniform($\{\pm 1\}^n$) is

$$\pi(\theta) \propto \exp\left(-\frac{||y_0 - X\Theta||^2}{2\sigma^2}\right) = \text{Ising with } J = XX^T/\sigma^2 \text{ and } h = y_0^T X/2\sigma^2$$

Bayesian approach requires sampling from this Ising model.
Closely related to the SK case

Motivation: Physics [Ising-Lenz'1920-1925, Sherrington-Kirkpatrick'1975, Parisi'1979,…]

$\mu_{J,h}$: distribution over charged ($\pm$) configurations of particles, where matrix J represents the interaction between particles



Sampling $\equiv$ Simulation

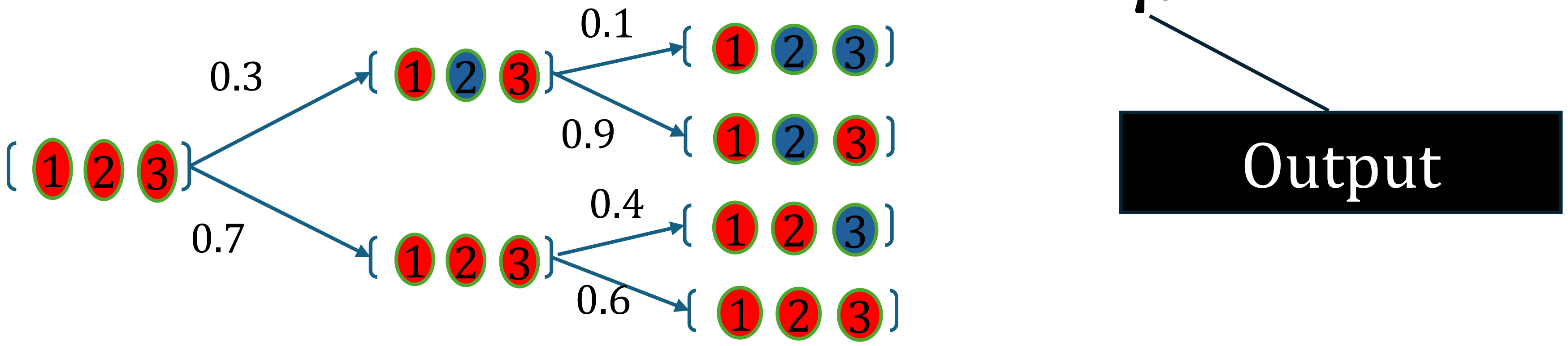
Glauber, a natural algorithm for sampling

Target: $\mu : \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$
Start at arbitrary σ
Repeat for "some" steps:

- Choose a random location i and resample $\sigma_i \sim \mu(.\,|\sigma_{-i})$

We obtain a sequence of random variables:

$$\color{red}{X^0} \longrightarrow \color{green}{X^1} \longrightarrow X^2 \to \ \ldots \to \boldsymbol{X^T} \approx \boldsymbol{\mu}$$



Goal: bounding the mixing time

$$t_{mix}(\epsilon) = \underset{t}{\operatorname{argmin}}\, d_{TV}(X^t, \mu) \leq \epsilon$$

Target: $\mu: \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$

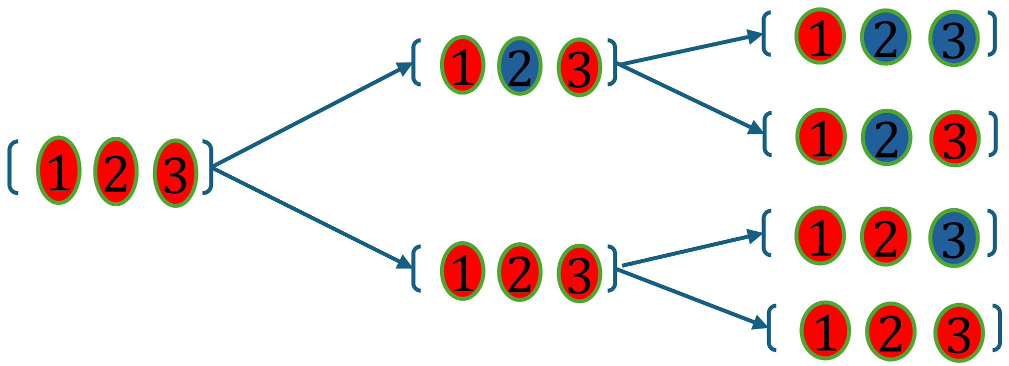
Start at arbitrary σ

Repeat for "some" steps:

- Choose a random location i and resample $\sigma_i \sim \mu(.|\sigma_{-i})$

We obtain a sequence of random variables:

$$X^0 \dashrightarrow X^1 \dashrightarrow X^2 \to \ \dots \to \boldsymbol{X^T} \approx \boldsymbol{\mu}$$

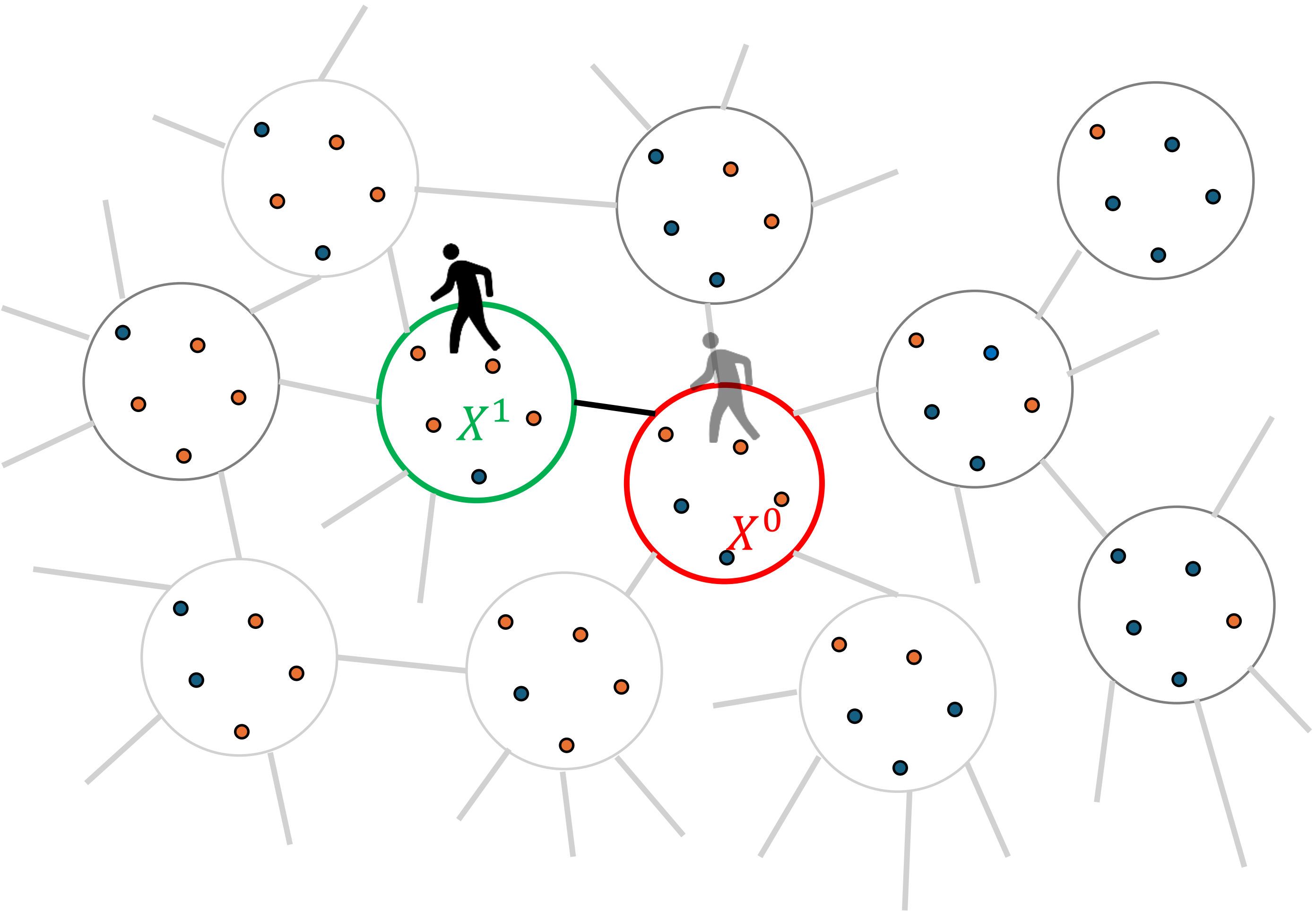


Output

Fast mixing ≡ no sparse cut in meta-graph

Meta graph:

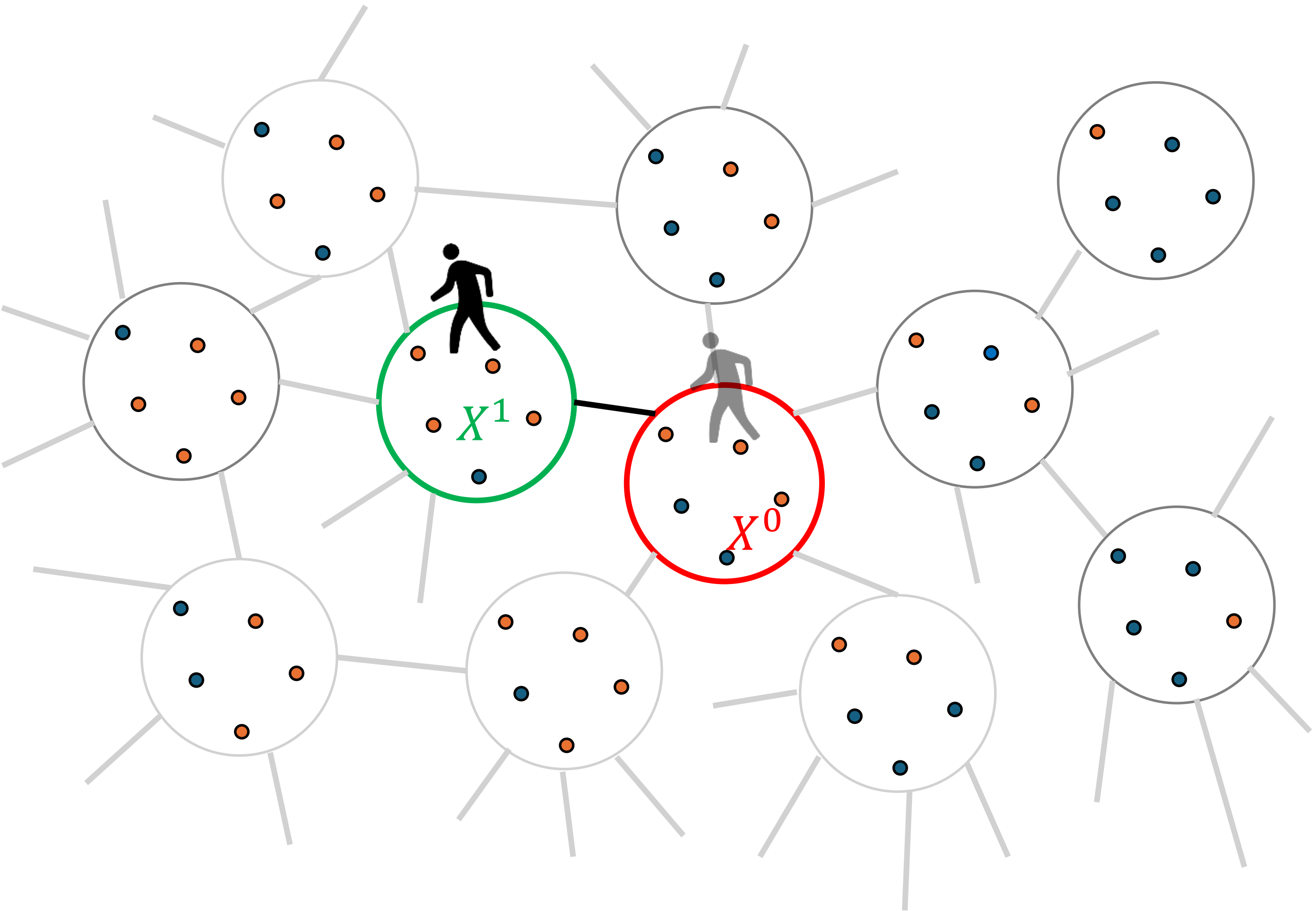
- Vertices $V = \{\pm 1\}^n$
- Vertex weight: $\mu(\sigma)$
- Edge weight: $\mathbb{P}[\sigma \to \sigma']$



Fast mixing ≡ no sparse cut in meta-graph

Meta graph:

- Vertices $V = \{\pm 1\}^n$
- Vertex weight: $\mu(\sigma)$
- Edge weight: $\mathbb{P}[\sigma \to \sigma']$

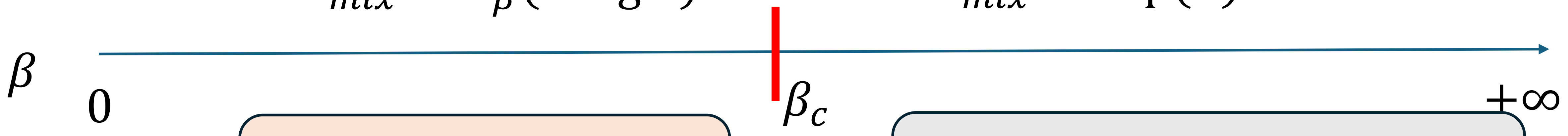


<u>For intuition</u>:
J "big" ↔ ∃ sparse cut
J "small" ↔ NO sparse cut

A sharp phase transition

$$J = \beta\, J^{(0)}$$

$t_{mix} = O_\beta(n \log n)$ $\qquad$ $t_{mix} = \exp(n)$



Near-linear and optimally fast

Exponentially slow

<u>Conj</u>: no poly-time sampler (proved* in special cases)

Example: mean-field/Curies-Weiss model

$J = \beta\, \mathbf{1}$

$t_{mix} = O_\beta(n \log n)$

$t_{mix} = \exp(n)$

β

0

β_c

$+\infty$

Near-linear and optimally fast

Exponentially slow

$U\left(\{\pm 1\}^n\right)$

$U(\{(+1)^n, (-1)^n\})$

Ising model: $\mu^{J,h}(x) = \exp(\frac{1}{2}\langle x, Jx \rangle + \langle h, x \rangle)$

Glauber's mixing time	J	h
$O(n \log n)$ [Dobrushin'68,Marton'15]	$\left\|\|J_{i,\cdot}\|\right\|_1 < 1$	0
$\exp(O(\sqrt{n}))$ [Bauerschmidt- Bodineau'19]	$\lambda_{\max}(J) - \lambda_{\min}(J) < 1$	0
$O(n^2)$ [Eldan-Koehler-Zertouni'21]	$\lambda_{\max}(J) - \lambda_{\min}(J) < 1$	arbitrary
$O(n \log n)$ [Anari-Jain-Koehler-Pham-V.'21a] [Chen-Eldan'22]: shorter proof	$\lambda_{\max}(J) - \lambda_{\min}(J) < 1$	arbitrary
$\exp(\Omega(n))$ No polytime algorithm [Kun'23, Galanis-Kalavakis-Kandiros'24]	$\lambda_{\max}(J) - \lambda_{\min}(J) > 1$	0

Tight example: $J = \frac{c}{n} 1\, 1^T, \lambda_{\max}(J) = c, \lambda_{\min}(J) = 0$

SK model: $\mu^{J,h}(x) = \exp(\frac{1}{2}\langle x, Jx \rangle + \langle h, x \rangle), J_{ji} = J_{ij} \sim \mathcal{N}\left(0, \frac{\beta^2}{n}\right)$ i.i.d.

Glauber's mixing time	J	h
$O(n^2)$ [Eldan-Koehler-Zertouni'21]	$\beta < 1/4$	arbitrary
$O(n \log n)$ [Anari-Jain-Koehler-Pham-**V.**'21a]	$\beta < 1/4$	arbitrary

$$\lambda_{\max}(J) \approx 2\beta, \lambda_{\min}(J) \approx -2\beta \Rightarrow \lambda_{\max}(J) - \lambda_{\min}(J) \leq 4\beta \leq 1$$

SK model: $\mu^{J,h}(x) = \exp(\frac{1}{2}\langle x, Jx \rangle + \langle h, x \rangle), J_{ji} = J_{ij} \sim \mathcal{N}\left(0, \frac{\beta^2}{n}\right)$ i.i.d.

Glauber's mixing time	J	h
$O(n^2)$ [Eldan-Koehler-Zertouni'21]	$\beta < 1/4$	All external fields
$O(n \log n)$ [Anari-Jain-Koehler-Pham-**V.**'21a]	$\beta < 1/4$	All external fields

$$\lambda_{\max}(J) \approx 2\beta, \lambda_{\min}(J) \approx -2\beta \Rightarrow \lambda_{\max}(J) - \lambda_{\min}(J) \leq 4\beta < 1$$

Conjecture: Glauber mixes fast for $\beta < 1$

SK model: $\mu^{J,h}(x) = \exp(\frac{1}{2}\langle x, Jx \rangle + \langle h, x \rangle), J_{ji} = J_{ij} \sim \mathcal{N}\left(0, \frac{\beta^2}{n}\right)$ i.i.d.

Sampling time	J	h
$O(n^2)$ [Eldan-Koehler-Zertouni'21]	$\beta < 1/4$	All external fields
$O(n \log n)$ [Anari-Jain-Koehler-Pham-V.'21a]	$\beta < 1/4$	All external fields
$O(n^2)$ **but** very **weak guarantee**: $W_2(\mu, \hat{\mu}) = o(n)$ [El Alaoui-Montanari-Sellke'22, Brennecke-Xu-Yau'23, Celentano'23]	$\beta < 1$	0
$t_{mix,\text{Glauber}} \geq \exp(\Omega(n))$ No polytime "stable" algorithm [AMS'22]	$\beta > 1$	0

SDE-based sampler + marginal estimation by AMP

→ large error. Much weaker than $d_{TV}(\mu, \hat{\mu}) < \frac{1}{poly(n)}$ achievable by Glauber/Markov chains

SK model: $\mu^{J,h}(x) = \exp(\frac{1}{2}\langle x, Jx \rangle + \langle h, x \rangle), J_{ji} = J_{ij} \sim \mathcal{N}\left(0, \frac{\beta^2}{n}\right)$ i.i.d.

Glauber's mixing time	J	h
$O(n^2)$ [Eldan-Koehler-Zertouni'21]	$\beta < 1/4$	All external fields
$O(n \log n)$ [Anari-Jain-Koehler-Pham-**V.**'21a]	$\beta < 1/4$	All external fields
$O(n \log n)$ [Anari-Koehler-**V.**'24]	$\beta \le 0.295$	All external fields

SK model: $\mu^{J,h}(x) = \exp(\frac{1}{2}\langle x, Jx \rangle + \langle h, x \rangle), J_{ji} = J_{ij} \sim \mathcal{N}\left(0, \frac{\beta^2}{n}\right)$ i.i.d.

Glauber's mixing time	J	h
$O(n^2)$ [Eldan-Koehler-Zertouni'21]	$\beta < 1/4$	All external fields
$O(n \log n)$ [Anari-Jain-Koehler-Pham-**V.**'21a]	$\beta < 1/4$	All external fields
$O(n \log n)$ [Anari-Koehler-**V.**'24]	$\beta \leq 0.295$	All external fields

Open: $0.295 < \beta < 1$?

Results for related models

- Langevin dynamics in O(N) model (replacing $\{\pm 1\}$ with S_N)

$$\mu: S_N^n \to \mathbb{R}_{\geq 0}, \mu(\sigma) \propto \exp\left(\sum_{i \leq j} J_{ij} \langle \sigma_i, \sigma_j \rangle + \sum_i \langle h_i, \sigma_i \rangle\right)$$

Results for related models

- Langevin dynamics in O(N) model (replacing $\{\pm 1\}$ with S_N)
- Anti-ferromagnetic Ising on random d-regular graphs: $J = -\beta A(\mathrm{G})$ with $G \sim \mathcal{G}_d, \beta = \Theta(1/\sqrt{d})$

Correct order on d.
Open: getting the correct constant?

Results for related models

- Langevin dynamics in O(N) model (replacing $\{\pm 1\}$ with S_N)
- Anti-ferromagnetic Ising on random d-regular graphs
- Global version of Kawasaki dynamics for fixed-magnetization Ising --concurrent with [Bauerschmidt-Bordineau-Dagallier23]

Ising restricted to $\{\sigma \in \{\pm 1\}^n \mid \sum \sigma_i = k\}$

Results for related models

- Langevin dynamics in O(N) model (replacing $\{\pm 1\}$ with S_N)
- Anti-ferromagnetic Ising on random d-regular graphs
- Global version of Kawasaki dynamics for fixed-magnetization Ising --concurrent with [Bauerschmidt-Bordineau-Dagallier23]
- P-spin model: $\log \mu(\sigma)$ is degree-p polynomial (using different technique—[AJKPV24], but can we unify the two techniques?)

Proof technique

Goal: bounding mixing time of Glauber

Common method to bound mixing time:

- Quantifying progress in each time-step
- More precisely, compare $dist(X^{t+1}, \mu)$ with $dist(X^t, \mu)$ for some measure of distance between distributions
- Hard to directly work with $dist = \mathrm{d}_{TV}$, instead, work with

$$dist \in \{\mathcal{D}_{KL}, \chi^2\}$$

$$\mathcal{D}_{KL}(\nu||\mu) = Ent\left(\frac{\nu}{\mu}\right)$$

$$\chi^2(\nu||\mu) = Var\left(\frac{\nu}{\mu}\right)$$

Entropy/variance contraction & connection to mixing time

Variance contraction

- $\forall \nu$: $\chi^2(\nu P||\mu P) \leq \left(1-\rho_{\chi^2}\right)\chi^2(\nu||\mu)$
- $T_{mix} \leq \rho_{\chi^2}^{-1} \log\left(\min \mu(x)\right)^{-1} \approx \rho_{\chi^2}^{-1} n$
- Easier to bound but worse control on mixing time

Entropy contraction

- $\forall \nu$: $\mathcal{D}_{KL}(\nu P||\mu P) \leq (1-\rho_{KL})\mathcal{D}_{KL}(\nu||\mu)$
- $T_{mix} \leq \rho_{KL}^{-1} \log\log\ \left(\min \mu(x)\right)^{-1} \approx \rho_{KL}^{-1} \log n$
- $\rho_{\chi^2} \geq \rho_{KL}$
- Tight control on mixing time

  +concentration bounds…

Entropic independence-[AJKP**V**21]:
Lossless boost variance contraction
into KL contraction

For Glauber on Ising $\mu^{J,h}: \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$, we expect: $\rho_{\chi^2} \approx \rho_{KL} \approx \frac{1}{n}$

Decomposition & induction

- Old idea
- Decompose into **disjoint** subspaces
- Often loose bounds
- How to decompose?

Solve the (easier) problem on subspace

Large state space $\Omega = \{\pm 1\}^n$

Decomposition into chains on smaller space

$\Omega_a \equiv \{\pm 1\}^{n/3}$

Decomposition & induction

Spectral/entropic ind/[CE22]'s stochastic localization

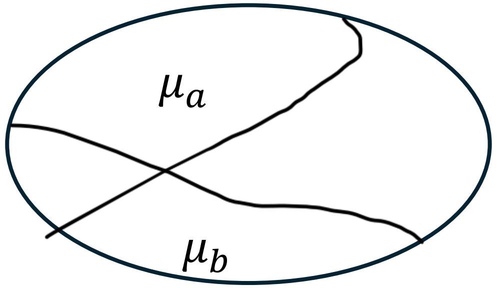
- Canonical decomposition
- Main difference: allow overlapping
- Tight/optimal bound

Solve the (easier) problem on subspace

Large state space

$\Omega = \{\pm 1\}^n$

Decomposition into chains on smaller space



Decomposition

Spectral/entropic independence

Decomposition via conditioning/pinning:

$$\mu = \mathbb{E}_{i,c}[\mu(.|x_i = c)]$$

Reduces entropy/variance contraction of μ to $\mu(.|x_i = \mathrm{c})$?

[ALOV19,ALO20,CLV21,22,AJKPV21,..]:

Yes, if $Cov(\mu) \preccurlyeq ..$

Linear-tilt stochastic localization

Decomposition into “soft pinning”:

$$\mu = \mathbb{E}_{\theta\sim\pi}[\tilde{\mu}_\theta]$$

Reduces entropy/variance contraction of μ to $\tilde{\mu}_\theta$?

[CE22]:

Yes, if $Cov(\mu) \preccurlyeq ..$

But only for quadratic Hamiltonian

i.e. Ising

What we have learned

Bounded covariance $\Rightarrow$ entropy contraction & mixing time bound

via reducing to simpler systems

- Intuition: zero covariance $\Leftrightarrow$ product distribution

  $\Rightarrow$ optimal mixing

- Formalized in [AJKPV21-22,CE22,AKV24]

What we have learned

Bounded covariance ⇒ entropy contraction & mixing time bound

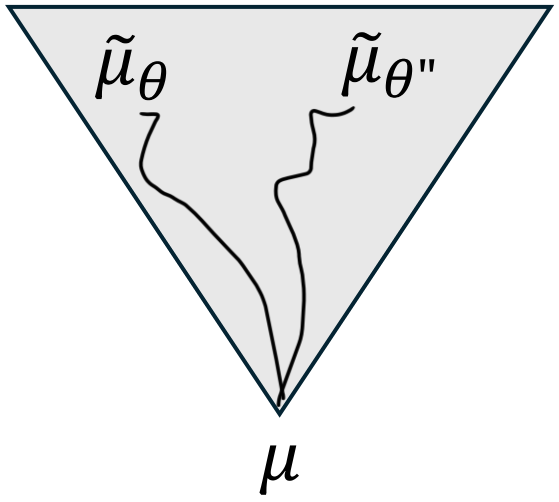
via reducing to simpler systems

- Intuition: zero covariance ⇔ product distribution

  ⇒ optimal mixing

- Formalized in [AJKP**V**21-22,CE22,AK**V**24]
- Note for experts: bounding covariance of the target/stationary distribution is not enough.

Our contribution:
a new way to bound the covariance



Given [CE22] decomposition:

$$\mu = \mathbb{E}_{\theta\sim\pi}[\tilde{\mu}_\theta]$$

We can bound $\mathrm{Cov}(\mu)$ using $Cov(\tilde{\mu}_\theta)$!

A generalization of Oppenheim's trickle-down theorem, which uses the decomposition via pinning.

Decomposition via stochastic/linear-tilt localization [Chen-Eldan'22]

Decomposition via stochastic localization

[Chen-Eldan'22]

Sequence of densities (μ_t), $F_t = \frac{d\mu_t}{d\mu}$

Driving matrices: C

$$F_0 = 1$$

$$dF_t(x) = F_t(x)\langle x - mean(\mu_t), CdB_t\rangle$$

Decomposition:

$$\mu = \mu_0 = \mathbb{E}_{\mathcal{F}_t}[\mu_t]$$

$$F_t(x) \propto \exp(\langle C^2 y_t, x\rangle - \frac{t}{2}\langle x, C^2 x\rangle),$$

for $\mathrm{y_t} = \mathrm{t}\,\mathrm{X}^* + C^{-1}B_t$ for $X^* \sim \mu_0$

Decomposition via stochastic localization

[Chen-Eldan'22]

Sequence of densities (μ_t), $F_t = \frac{d\mu_t}{d\mu}$

Driving matrices: C

$$F_0 = 1$$

$$dF_t(x) = F_t(x)\langle x - mean(\mu_t), CdB_t\rangle$$

Decomposition:

$$\mu = \mu_0 = \mathbb{E}_{\mathcal{F}_t}[\mu_t]$$

$$F_t(x) \propto \exp(\langle C^2 y_t, x\rangle - \frac{t}{2}\langle x, C^2 x\rangle),$$

for $\mathrm{y_t} = \mathrm{t\,X^*} + C^{-1}B_t$ for $X^* \sim \mu_0$

<u>Application</u>:

- μ_0: Ising model with interaction matrix $J \succcurlyeq 0$
- $C = J^{1/2}$
- $\mu_t \equiv \mu_{Ising}^{(1-t)J,h}$

$$\mu^{J,h} \equiv \mu^{J+\alpha I,h}$$

Decomposition via stochastic localization

[Chen-Eldan'22]

Sequence of densities (μ_t), $F_t = \frac{d\mu_t}{d\mu}$

Driving matrices: C

$$F_0 = 1$$

$$dF_t(x) = F_t(x)\langle x - mean(\mu_t), C dB_t \rangle$$

Decomposition:

$$\mu = \mu_0 = \mathbb{E}_{\mathcal{F}_t}[\mu_t]$$

$$F_t(x) \propto \exp(\langle C^2 y_t, x\rangle - \frac{t}{2}\langle x, C^2 x\rangle),$$

for $\mathrm{y_t} = \mathrm{t}\,\mathrm{X}^* + C^{-1}B_t$ for $X^* \sim \mu_0$

Application:

- μ_0: Ising model with interaction matrix $J \succcurlyeq 0$
- $C = J^{1/2}$
- $\mu_t \equiv \mu_{Ising}^{(1-t)J, h_t}$
- $T = 1$
- $\mu_1(x) \propto \exp(\langle C^2 y_t + h, x\rangle)$:
  - product distribution
  - easy to bound $\Sigma_1 = \mathrm{Cov}(\mu_1)$ because Σ_1 is diagonal

Bounded covariance implies entropy contraction

<u>Thm</u>[CE22] : If $\forall t, w$: $\left|\left|\mathrm{cov}\left(\mu_{Ising}^{(1-t)J,h_t+w}\right)\right|\right|_{OP} \leq \gamma_t$ then

$$\rho_{KL}\left(P_{Glauber}, \mu_{Ising}^{J,h}\right) \geq \frac{\exp\left(-||J||_{OP}\int \gamma_t dt\right)}{\mathrm{n}}$$

$$\Rightarrow t_{mix} \leq \exp\left(||J||_{OP}\int \gamma_t dt\right) O(n \log n)$$

Bounding covariance using trickle-down

Trickle down for stochastic localization

[Anari-Koehler-V'24]

Sequence of densities (μ_t), $F_t = \frac{d\mu_0}{d\mu}$

Driving matrices: C

$$F_0 = 1$$
$$dF_t(x) = F_t(x)\langle x - mean(\mu_t), CdB_t\rangle$$

Thm 2 [ACV24]: for $0 \leq t \leq 1$: $\Sigma_t := \mathrm{cov}(\mu_t)$

$$\Sigma_t = \mathbb{E}[\Sigma_1|\mathcal{F}_t] + \int \mathbb{E}[\Sigma_s C^2 \Sigma_s|\mathcal{F}_t]\, ds$$

Bounding $\mathrm{cov}(\mu_1) \Rightarrow$ bounding $\mathrm{cov}(\mu_t)$

Trickle down for stochastic localization

[Anari-Koehler-V'24]

Sequence of densities (μ_t), $F_t = \frac{d\mu_0}{d\mu}$

Driving matrices: C

$$F_0 = 1$$

$$dF_t(x) = F_t(x)\langle x - mean(\mu_t), C dB_t\rangle$$

Thm 2 [ACV24]: for $0 \le t \le 1$: $\Sigma_t := \mathrm{cov}(\mu_t)$

$$\Sigma_t = \mathbb{E}[\Sigma_1 | \mathcal{F}_t] + \int \mathbb{E}[\Sigma_s C^2 \Sigma_s | \mathcal{F}_t]\, ds$$

By Ito calculus

$$d\Sigma_t = -\Sigma_t^2 C^2 \Sigma_t^2 dt + martingale$$

Taking everything together

<u>Thm</u>: Let $\mu_1^{(w)} := \mu_{Ising}^{0,h_1+w}$

if $\forall w$: $Cov\left(\mu_1^{(w)}\right) \preccurlyeq A$ then $\rho_{KL}\left(P_{Glauber}, \mu_{Ising}^{J,h}\right) \geq q(A)/n$

Naïve bound: $Cov\left(\mu_1^{(w)}\right) \preccurlyeq I$ $\Rightarrow$

Assume $J \succcurlyeq 0$

$\rho_{KL}\left(P_{Glauber}, \mu_{Ising}^{J,h}\right) \geq \frac{\left(1-||J||_{OP}\right)}{n}$

$t_{mix} \leq \frac{1}{1-||J||_{OP}} O(n \log n)$

Tight if $J = \frac{c}{n}\mathbf{1}\mathbf{1}^T$!

Recover [AJKP**V**21]

To do better for SK

<u>Thm</u>: Let $\mu_1^{(w)} \coloneqq \mu_{Ising}^{0,h_1+w}$

if $\forall w$: $Cov\left(\mu_1^{(w)}\right) \preccurlyeq A$ then $\rho_{KL}\left(P_{Glauber}, \mu_{Ising}^{J,h}\right) \geq q(A)/n$

μ_0: Ising model with $J^{(0)} \sim GOE\left(\frac{\beta^2}{n}\right), J = \underbrace{J^{(0)} + |\lambda_{min}(J^{(0)})| I}_{\text{PSD transformation}} \succcurlyeq 0$

- $\Sigma_1 := Cov\left(\mu_1^{(w)}\right)$ is a diagonal matrix
- Exploit the structure of J to bound the diagonal entries of Σ_1

J has large diagonal entries

To do better for SK

Thm: Let $\mu_1^{(w)} := \mu_{Ising}^{0,h_1+w}$

if $\forall w$: $Cov\left(\mu_1^{(w)}\right) \preccurlyeq A$ then $\rho_{KL}\left(P_{Glauber}, \mu_{Ising}^{J,h}\right) \geq q(A)/n$

μ_0: Ising model with $J^{(0)} \sim GOE\left(\frac{\beta^2}{n}\right), J = J^{(0)} + |\lambda_{min}(J^{(0)})| I \succcurlyeq 0$

- $\Sigma_1 := Cov\left(\mu_1^{(w)}\right)$ is a diagonal matrix
- Exploit the structure of J to bound the diagonal entries of Σ_1
- $(\Sigma_1)_{\mathrm{i,i}} = 1 - \tanh^2(\theta_i)$, $\theta_i = \langle J_i, X^* \rangle + \tilde{h}_i + \sqrt{J_{i,i}}\, g)$

$\mathrm{X}^* \sim \mu_{Ising}^{J,\tilde{h}}$

Lem15,58[AKV24]:
$\mathbb{E}[\tanh^2(\theta_i)] \geq r(J_{i,i})$

Main theorem for Ising

Thm3[AKV24]: $\left\|\operatorname{cov}\left(\mu_{Ising}^{J,h}\right)\right\|_{OP} \leq q_\eta(\Delta), \Delta := \lambda_{max}(J) - \lambda_{min}(J)), \eta := \frac{1}{1+|\frac{\lambda_{max}(J)}{\lambda_{min}(J)}|}$

$$q_\eta(z) = r(\eta z) + \int_0^z q_\eta(y)^2 dy$$

$$\rho_{KL}\left(P_{Glauber}, \mu_{Ising}^{J,h}\right) \geq \exp\left(\int_0^\Delta q_\eta(z)\, dz\right)/n$$

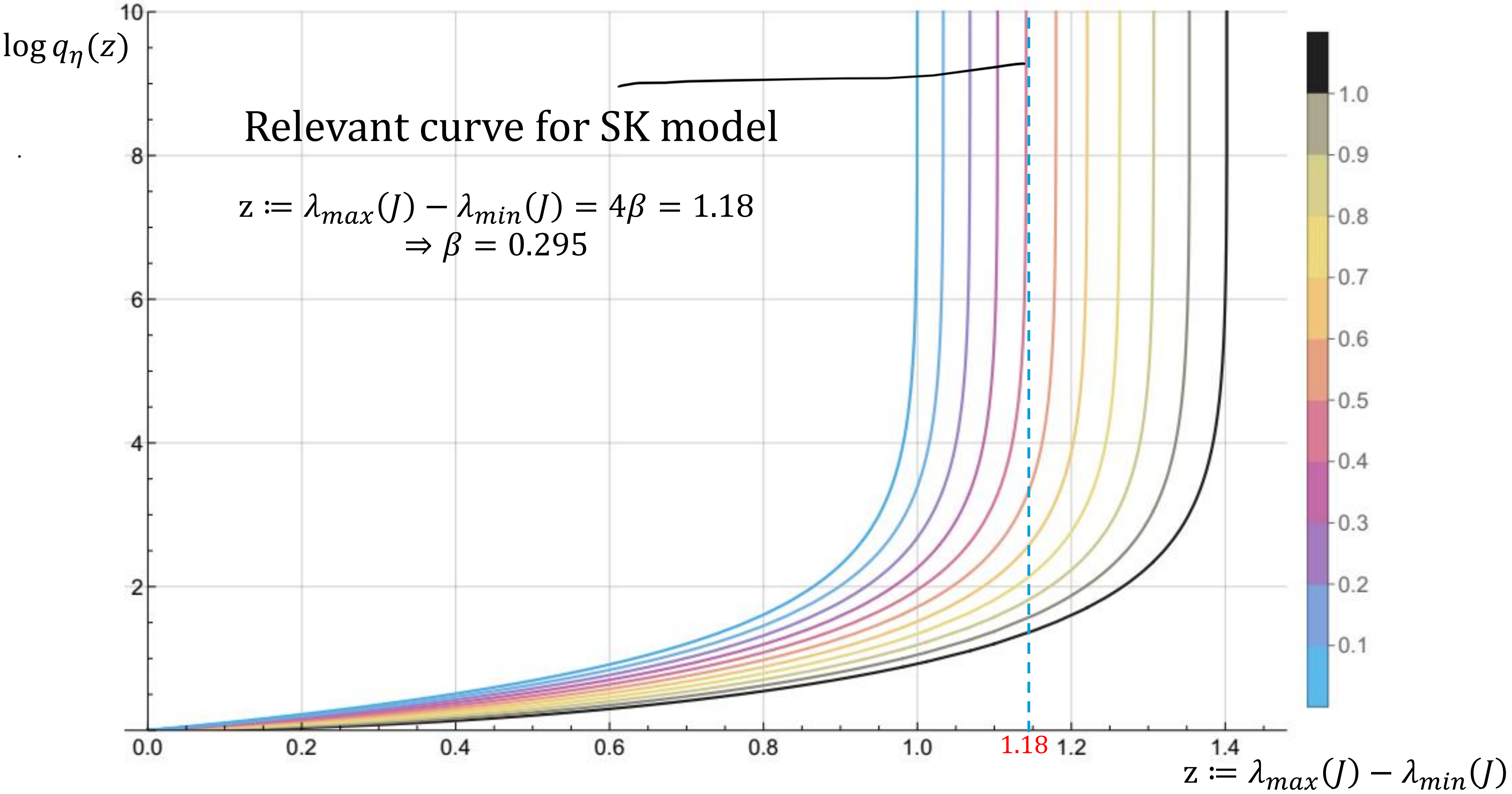


Figure 1: $\log q_\eta(z)$ from Theorem 3 plotted using numerical integration in Mathematica for $\eta = 0, 0.1, \ldots, 0.9, 1$. The curve for $\eta = 0$ was, before this work, the best known bound for all values of η in the Ising model; it asymptotes to ∞ at $z = 1$, which is tight due to the phase transition in the Curie-Weiss model [Ell06]. For $\eta = 0.5$ the function $\log q_\eta$ asymptotes to ∞ at z slightly above 1.18 and for $\eta = 1.0$ it asymptotes around 1.40.

SK model: $\mu^{J,h}(x) = \exp(\frac{1}{2}\langle x, Jx \rangle + \langle h, x \rangle), J_{ji} = J_{ij} \sim \mathcal{N}\left(0, \frac{\beta^2}{n}\right)$ i.i.d.

Glauber's mixing time	J	h
$O(n^2)$ [Eldan-Koehler-Zertouni'21]	$\beta < 1/4$	All external fields
$O(n \log n)$ [Anari-Jain-Koehler-Pham-V.'21a]	$\beta < 1/4$	All external fields
$O(n \log n)$ [Anari-Koehler-V.'24]	$\beta \leq 0.295$	All external fields
$\mathrm{t}_{mix,\mathrm{Glauber}} \geq \exp(\Omega(n))$ No polytime "stable" algorithm [AMS'22]	$\beta > 1$	0

Open Problem 1:
$0.295 < \beta < 1$?

Open Problem 2:
Stronger evidence for hardness when $\beta > 1$