Rapid mixing of Glauber dynamics via spectral independence for all degrees

Weiming Feng

University of Edinburgh

Joint work with: Xiaoyu Chen (Nanjing University)

Yitong Yin (Nanjing University)

Xinyuan Zhang (Nanjing University)

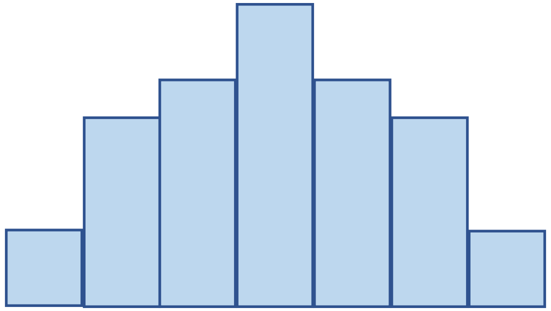
FOCS 2021

Sampling, counting and phase transition

Boolean variables set V, weight function $w: \{-,+\}^V \to \mathbb{R}_{\geq 0}$

joint distribution μ: $\forall X \in \{-,+\}^V,\ \mu(X) = \frac{w(X)}{Z}$

partition function $Z = \sum_{\{-,+\}^V} w(X)$

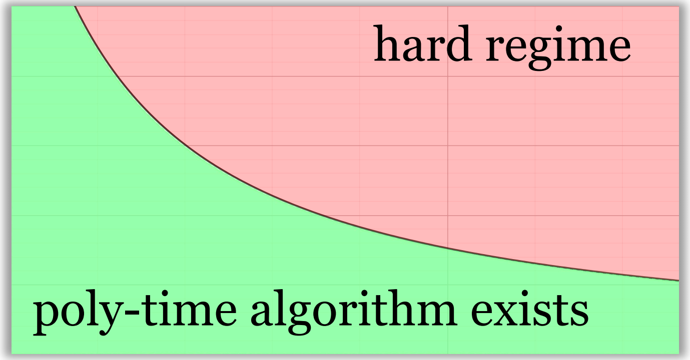


Sampling problem

Draw (approximate) random samples from distribution μ

Computational phase transition

computational complexity of sampling problem
changes sharply around some parameters of μ

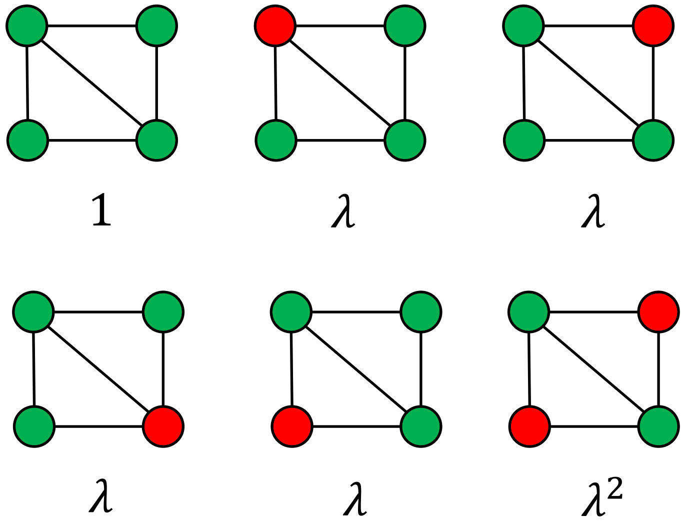


Hardcore gas model

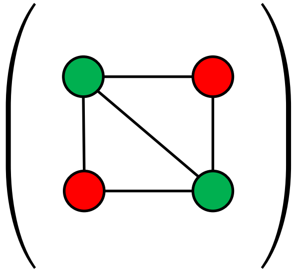
- Graph $G = (V, E)$: n-vertex and max degree Δ;
- Fugacity parameter $\lambda \in \mathbb{R}_{\geq 0}$;
- Configuration $X \in \{-,+\}^V$
  - $X_v = +$: vertex v is **occupied**
  - $X_v = -$: vertex v is **unoccupied**
- $X \in \Omega$ if occupied vertices form an independent set
- Gibbs distribution μ:

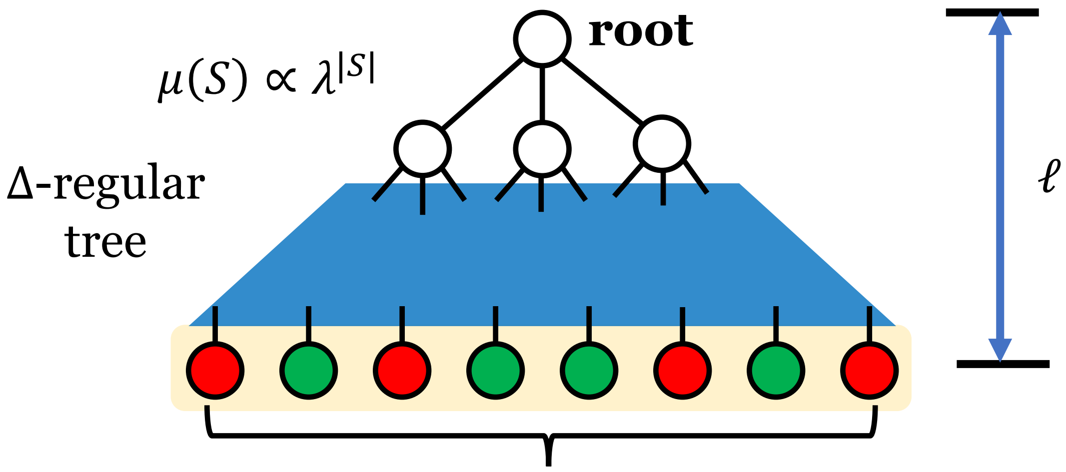
$$\forall X \in \Omega, \qquad \mu(X) \propto w(X) = \lambda^{|X|_+}.$$

$|X|_+ =$ *number of occupied vertices* $(X_v = +)$



Partition function
$Z = 1 + 4\lambda + \lambda^2$

μ  $= \dfrac{\lambda^2}{1 + 4\lambda + \lambda^2}$



σ: boundary configuration at level ℓ
each leaf $\sigma_v \in +$ or $\sigma_v = -$

Uniqueness Threshold

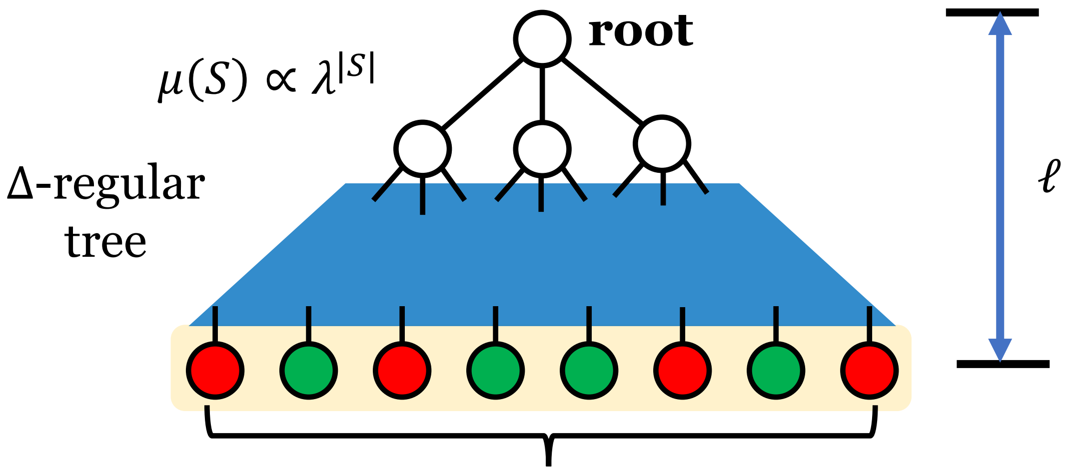
$\Pr[X(\text{root}) = + \mid \sigma]$ is independent of σ if $\ell \to \infty$

iff $\lambda \le \lambda_c(\Delta) = \dfrac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \dfrac{e}{\Delta}$

Δ: maximum degree

Computational phase transition

- $\lambda < \lambda_c$: poly-time algorithm for sampling[Weitz06]
- $\lambda > \lambda_c$: no poly-time algorithm unless $\boldsymbol{NP} = \boldsymbol{RP}$ [Sly10]



σ: boundary configuration at level ℓ
each leaf $v \in S$ or $v \notin S$

Uniqueness Threshold

$\Pr[X(\text{root}) = + \mid \sigma]$ is independent of σ if $\ell \to \infty$

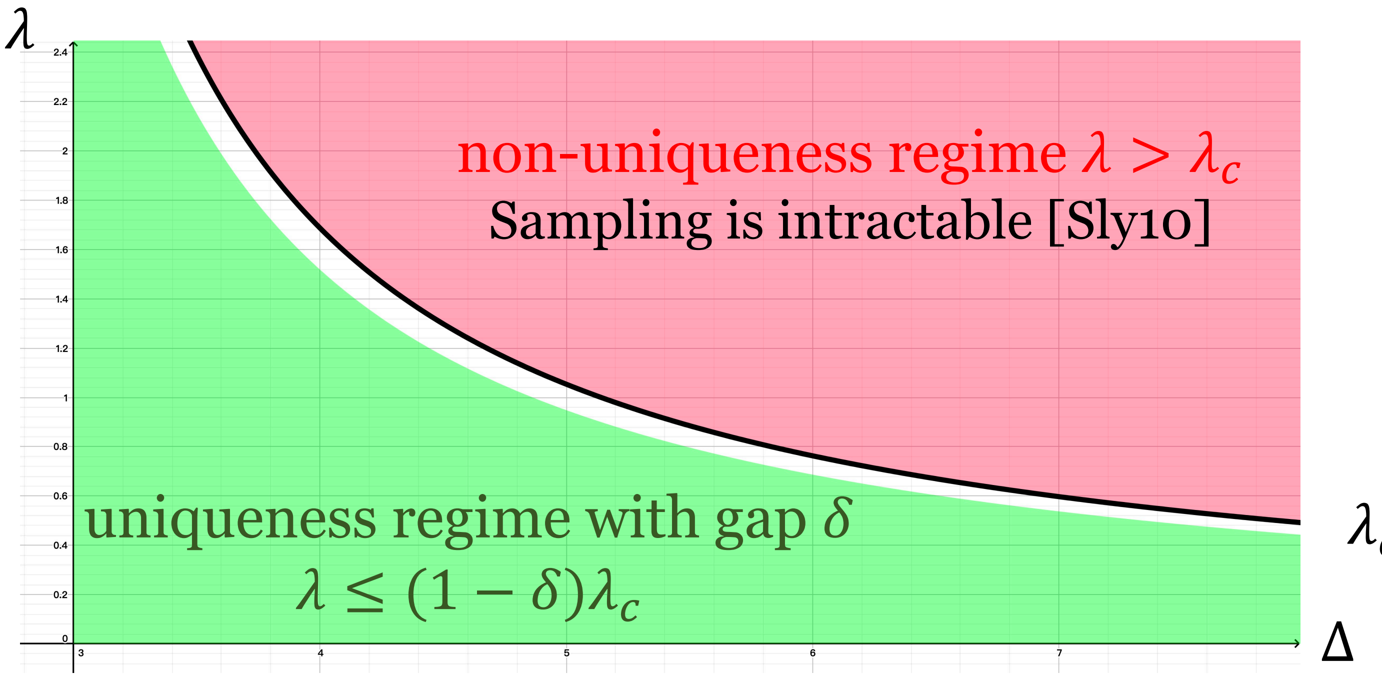
iff $\lambda \le \lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{e}{\Delta}$

Δ: maximum degree

Computational phase transition

- bounded degree $\Delta = O(1)$
- δ in the exponent of n

- $\lambda \le (1-\delta)\lambda_c$: $n^{O\left(\frac{\log \Delta}{\delta}\right)}$-time algorithms for sampling (via approx. counting) [Weitz06]
- $\lambda > \lambda_c$: no poly-time algorithm unless $\boldsymbol{NP} = \boldsymbol{RP}$ [Sly10]



$$\lambda_c(\Delta) = \frac{(\Delta - 1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}}$$

Problem : *fixed parameter trackable* sampling algorithm for hardcore model

Let $\delta > 0$ be an *arbitrary gap*. For any hardcore model with $\lambda \le (1-\delta)\lambda_c(\Delta)$,

can we sample from Gibbs distribution in time $C(\delta) \cdot \mathrm{poly}(n)$?

Glauber dynamics for hardcore model

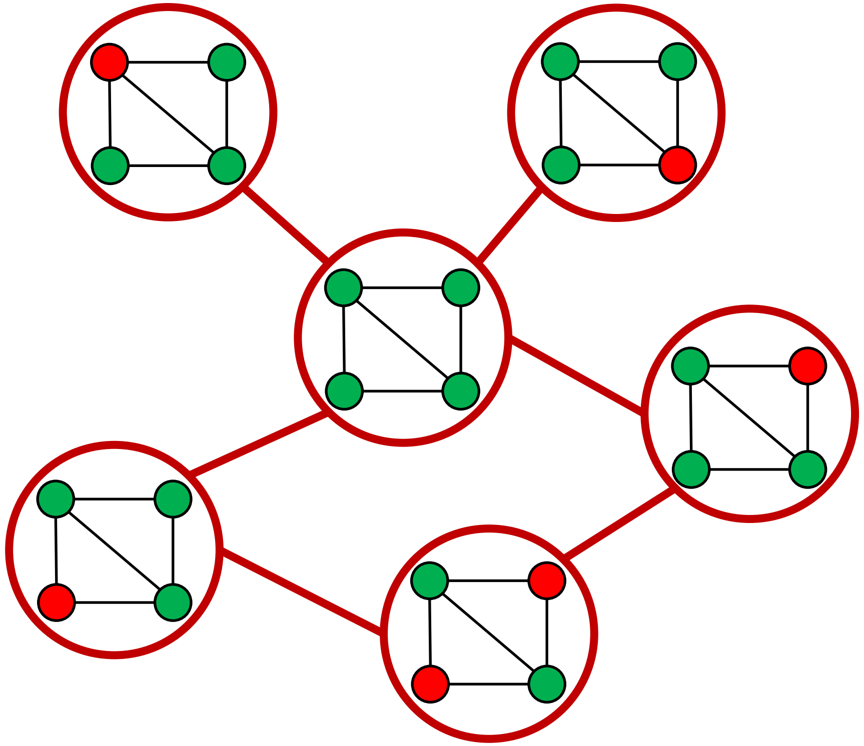
Start from an arbitrary independent set X;

For each transition step **do**

- Pick a vertex v uniformly at random;
- **If** $X_u = -$ for all neighbors u **then**

$$X_v = \begin{cases} + & \text{w.p.} \quad \lambda/(1+\lambda) \\ - & \text{w.p.} \quad 1/(1+\lambda) \end{cases}$$

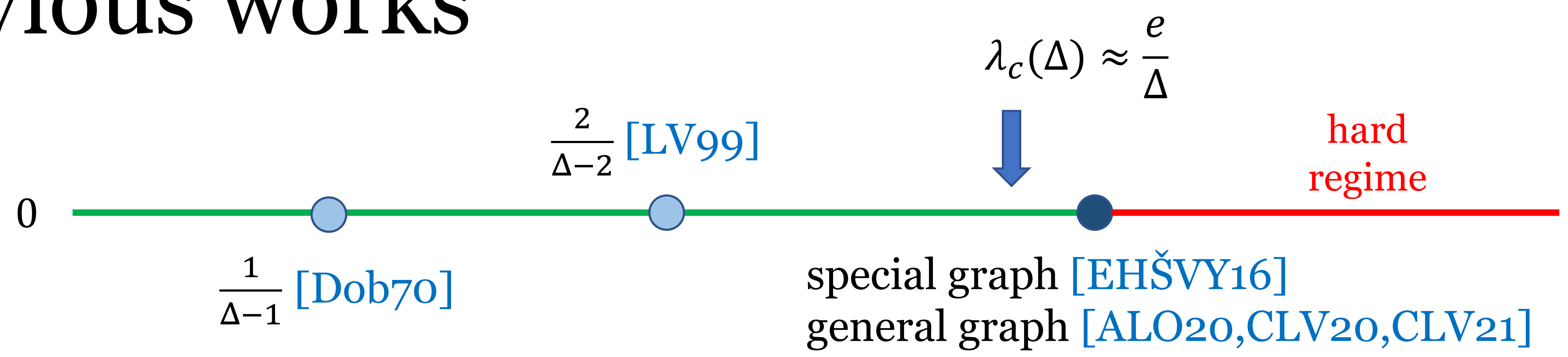
- **Else** $X_v \leftarrow -$



Mixing time: $T_{\text{mix}} = \max_{X_0 \in \Omega} \min \left\{ t \;\middle|\; d_{TV}(X_t, \mu) \le \frac{1}{4e} \right\}$,

$d_{TV}(X_t, \mu)$: the *total variation distance* between X_t and μ.

Previous works



Work	Condition	Mixing Time
Dobrushin 1970	$\lambda \leq \frac{1-\delta}{\Delta-1}$	$O\left(\frac{1}{\delta} n \log n\right)$
Luby, Vigoda 1999	$\lambda \leq \frac{2(1-\delta)}{\Delta-2}$	$O\left(\frac{1}{\delta} n \log n\right)$
Efthymiou *et al* 2016	$\lambda \leq (1-\delta)\lambda_c(\Delta)$ $\Delta \geq \Delta_0(\delta)$, girth ≥ 7	$O\left(\frac{1}{\delta} n \log n\right)$
Anari, Liu, Oveis Gharan 2020 improved by Chen, Liu, Vigoda 2020	$\lambda \leq (1-\delta)\lambda_c(\Delta)$	$n^{O(1/\delta)}$
Chen, Liu, Vigoda 2021	$\lambda \leq (1-\delta)\lambda_c(\Delta)$	$\Delta^{O(\Delta^2/\delta)} n \log n$

Our results

Following results holds for all $\delta \in (0,1)$

Work	Condition	Mixing Time
Anari, Liu, Oveis Gharan *2020* Improved by Chen, Liu, Vigoda 2020	$\lambda \le (1-\delta)\lambda_c(\Delta)$	$n^{O(1/\delta)}$
Chen, Liu, Vigoda 2021	$\lambda \le (1-\delta)\lambda_c(\Delta)$	$\Delta^{O(\Delta^2/\delta)} n \log n$
Our Result	$\boldsymbol{\lambda \le (1-\delta)\lambda_c(\Delta)}$	$\boldsymbol{\exp\left(O\left(\frac{1}{\delta}\right)\right) \cdot n^2 \log n}$

Theorem (hardcore model) [this work]

For any $\delta \in (0,1)$, any hardcore model satisfying $\lambda \le (1-\delta)\lambda_c(\Delta)$

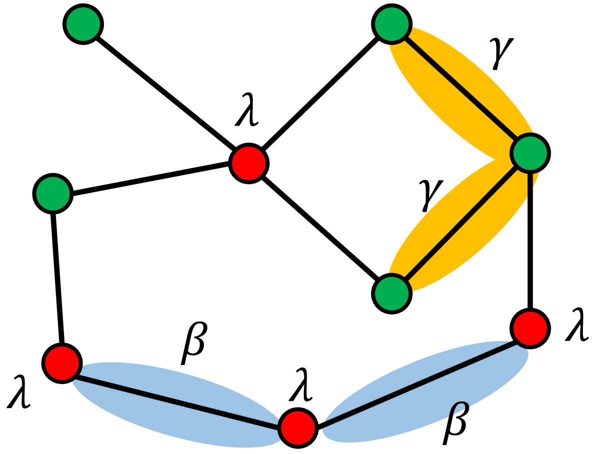
Glauber dynamics mixing time: $C(\delta)\, n^2 \log n$. (FPT w.r.t. δ)

Our results

Anti-ferro two-spin systems [this work]

For anti-ferro two-spin system that is up-to-Δ unique,

Glauber dynamics mixing time: $O(n^3)$.



$\mu(\sigma) \propto \lambda^4\beta^2\gamma^2$

Anti-ferro two-spin systems

- Hardcore model
- Ising model
- …

Joint distribution defined by external fields and local interactions

Results for general joint distributions

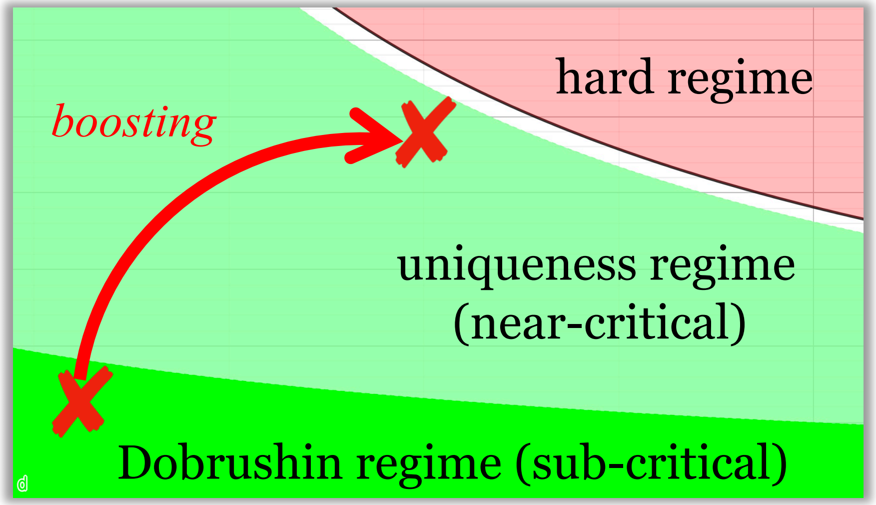
A **boosting result** of **spectral gap**

for **completely spectrally independent** distributions

Result for hardcore model: a corollary of general result

uniqueness condition → complete spectral independence

mixing result in ***sub-critical*** regime → (*boosting*) → mixing result in **near-critical** regime



Spectral gap and mixing time

Transition matrix of Glauber dynamics : $P: \Omega \times \Omega \to \mathbb{R}_{\geq 0}$

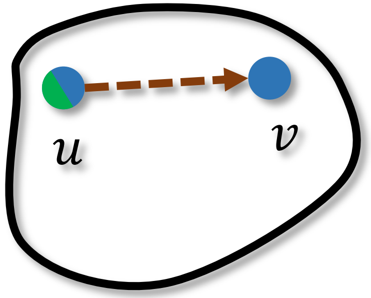
Eigenvalues : P has $|\Omega|$ non-negative real eigenvalues

$$1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\Omega|} \geq 0$$

Spectral gap $\lambda_{\text{gap}}(\mu) = 1 - \lambda_2$

$$T_{\text{mix}} = O\left(\frac{1}{\lambda_{\text{gap}}} \log \frac{1}{\mu_{\min}}\right), \qquad \mu_{\min} = \min_{\sigma \in \Omega} \mu(\sigma)$$

Influence matrix and spectral independence



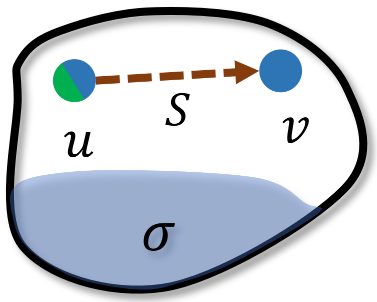
influence on v caused by a **disagreement** on u

μ: a distribution over $\Omega \subseteq \{-1,+1\}^V$

$|V|\times|V|$ **influence matrix** $\Psi \in \mathbb{R}^{V\times V}$ such that

$$\Psi(u,v) = \left|\Pr_{\mu}[\, v = + \mid u = + \,] - \Pr_{\mu}[\, v = + \mid u = - \,]\right|$$

Influence matrix and spectral independence



Influence from u to v
for **conditional distribution**

For any subset $S \subseteq V$, any feasible $\sigma \in \{-1,+1\}^{V \setminus S}$

μ_S^σ distribution on S conditional on σ

influence matrix $\Psi_S^\sigma \in \mathbb{R}^{S \times S}$ for **conditional distribution**

$$\Psi_S^\sigma(u,v) = \left| \Pr_{\mu_S^\sigma}[\, v = + \mid u = + \,] - \Pr_{\mu_S^\sigma}[\, v = + \mid u = - \,] \right|$$

Spectral independence (SI) [ALO20, CGŠV21, FGYZ21]

There is a constant $C > 0$ s.t. for all conditional distribution μ_S^σ,

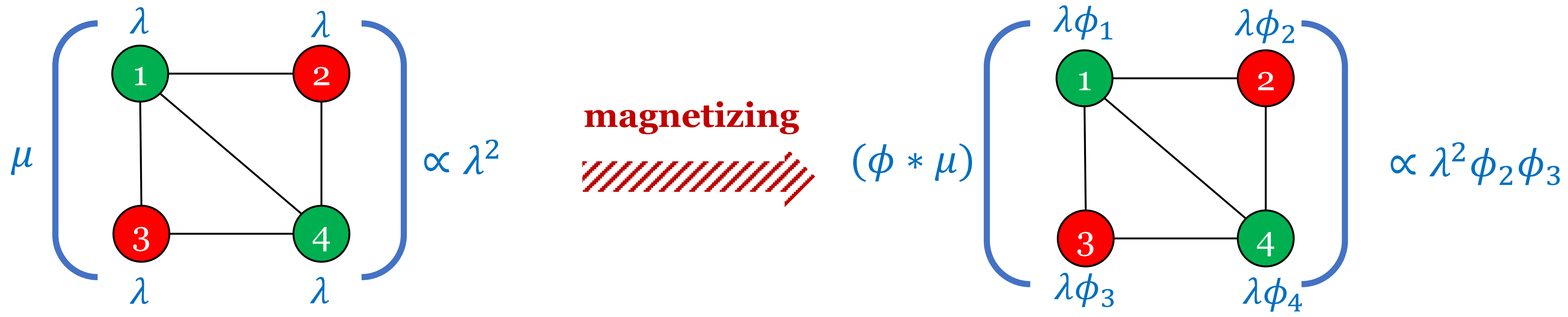
spectral radius of influence matrices $\rho(\Psi_S^\sigma) \leq C$.

Complete spectral independence

Magnetizing joint distribution with local fields

Joint distribution μ over $\{-,+\}^V$, local fields $\phi = (\phi_v)_{v\in V} \in \mathbb{R}^V_{>0}$

$$(\boldsymbol{\phi} * \mu)(\sigma) \propto \mu(\sigma) \prod_{v\in V:\sigma_v=+} \phi_v$$



Hardcore model: $\mu(S) \propto \lambda^{|S|}$

Hardcore mode with local fields

$\mu^{(\phi)}(S) \propto \lambda^{|S|} \prod_{v\in S} \phi_v = \prod_{v\in S} \lambda\phi_v$

Complete spectral independence

Complete Spectral independence [This work]

There is a constant $C > 0$ s.t.

for all local fields $\phi \in (0,1]^V$ (for all $v \in V$, $0 < \phi_v \leq 1$),

$(\boldsymbol{\phi} * \mu)$ is *spectrally independent* with parameter C

Example: hardcore model (G, λ) is ***completely spectrally independent*** if

any hardcore models $(G, (\lambda_v)_{v \in V})$ with $\lambda_v \leq \lambda$

are ***spectrally independent***

Boosting result of spectral gap [This work]

If μ is C-completely spectrally independent, for any $\theta \in (0,1)$

$$\lambda_{\text{gap}}(\mu) \geq \left(\frac{\theta}{2}\right)^{2C+7} \lambda^*_{\text{gap}}(\boldsymbol{\theta} * \mu), \qquad \boldsymbol{\theta}_v = \theta \text{ for all } v \in V$$

$\lambda^*_{gap}(\boldsymbol{\theta} * \mu)$: minimum spectral gap of Glauber dynamics for all conditional distributions induced by $\boldsymbol{\theta} * \mu$.

Boosting result of spectral gap [This work]

If μ is C-completely spectrally independent, for any $\theta \in (0,1)$

$$\lambda_{\text{gap}}(\mu) \geq \left(\frac{\theta}{2}\right)^{2C+7} \lambda^*_{\text{gap}}(\boldsymbol{\theta} * \mu), \qquad \boldsymbol{\theta}_v = \theta \text{ for all } v \in V$$

Near-Critical Regime

Boosting with cost $O(1)$

Impose local fields → Easy Regime

Boosting result of spectral gap [This work]

If μ is C-completely spectrally independent, for any $\theta \in (0,1)$

$$\lambda_{\text{gap}}(\mu) \geq \left(\frac{\theta}{2}\right)^{2C+7} \lambda^*_{\text{gap}}(\boldsymbol{\theta} * \mu), \qquad \boldsymbol{\theta}_v = \theta \text{ for all } v \in V$$

Application on hardcore model

$\lambda \leq (1-\delta)\lambda_c(\Delta)$

correction decay [Weitz06,LLY13, ALO20 CLV20] →

$O\left(\frac{1}{\delta}\right)$-completely SI

$\theta = 1/25$ ↓

$\theta\lambda \leq \frac{1}{2\Delta} \ll \lambda_c$

Dobrushin condition

- path coupling [BD97]
- coupling v.s. spectral gap [Chen98]

→

$\lambda^*_{\text{gap}}(\boldsymbol{\theta} * \mu) \geq \frac{1}{2n}$

$\lambda_{\text{gap}}(\mu) = \Omega_\delta(1/n)$

$T_{\text{mix}} = O_\delta(n^2 \log n)$

Proof of boosting result

New Markov chain: *field dynamics*

Field Dynamics

Input: a distribution μ over $\{-1,+1\}^V$, a parameter $\theta \in (0,1)$

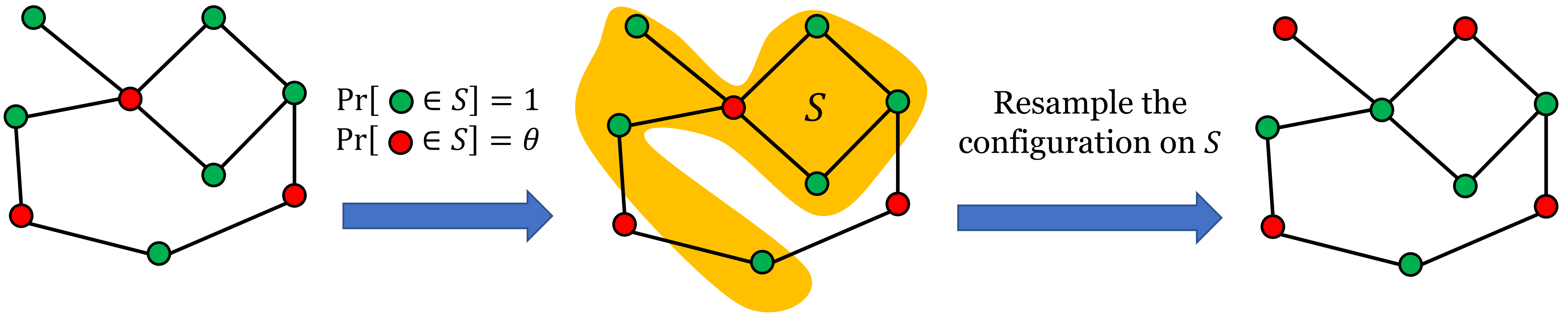
Start from an arbitrary feasible configuration $X \in \{-,+\}^V$

For each t from 1 to T **do**

- Construct $S \subseteq V$ be selecting each $v \in V$ independently with probability

$$p_v = \begin{cases} 1 & \text{if } X_v = - \\ \theta & \text{if } X_v = + \end{cases}$$

- Resample $X_S \sim (\boldsymbol{\theta} * \mu)_S(\cdot \mid X_{V \setminus S})$ conditional distribution induced from $(\boldsymbol{\theta} * \mu)$



Field Dynamics

Input: a distribution μ over $\{-1,+1\}^V$, a parameter $\theta \in (0,1)$

Start from an arbitrary feasible configuration $X \in \{-,+\}^V$

For each t from 1 to T **do**

- Construct $S \subseteq V$ be selecting each $v \in V$ independently with probability

$$p_v = \begin{cases} 1 & \text{if } X_v = - \\ \theta & \text{if } X_v = + \end{cases}$$

- Resample $X_S \sim (\boldsymbol{\theta} * \mu)_S(\cdot \mid X_{V \setminus S})$ conditional distribution induced from $(\boldsymbol{\theta} * \mu)$

Proposition (Field Dynamics): for any $\theta \in (0,1)$

The Field Dynamics $P_{FD}(\theta)$ is irreducible, aperiodic and reversible with respect to μ.

$P_{FD}(\theta)$ has the unique stationary distribution μ.

Comparison lemma

Proved by
a calculation

For any distribution μ over $\{-,+\}^V$

$$\lambda_{\text{gap}}(\mu) \geq \lambda_{\text{gap}}^{\text{Field}}(\mu,\theta) \cdot \lambda_{\text{gap}}^{*}(\boldsymbol{\theta} * \mu), \qquad \boldsymbol{\theta}_v = \theta \text{ for all } v \in V$$

Mixing lemma of field dynamics

If μ is C-completely spectrally independent, for any $\theta \in (0,1)$

$$\lambda_{\text{gap}}^{\text{Field}}(\mu,\theta) \geq \left(\frac{\theta}{2}\right)^{2C+7}$$

Comparison lemma + Mixing lemma → Boosting result

$$\lambda_{\text{gap}}(\mu) \geq \left(\frac{\theta}{2}\right)^{2C+7} \lambda_{\text{gap}}^{*}(\boldsymbol{\theta} * \mu)$$

Mixing of block dynamics [Chen, Liu and Vigoda 2021]

update of θ-*fraction block dynamics*

- sample θ fraction of variables R u.a.r.
- resample the value of R conditional on others

for any distribution π that is C-spectrally independent
$\lambda_{\text{gap}}^{\text{block}}(\pi) \geq \theta^{O(C)}$

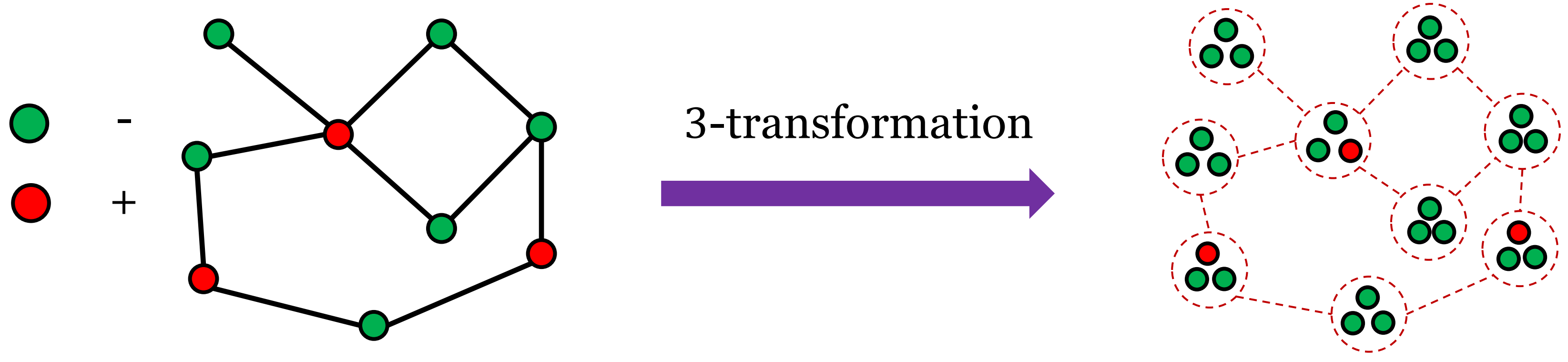
k-transformation [This work]

μ over $\{-,+\}^V$ → μ_k over $\{-,+\}^{V_k}$

$V_k = \{u_1, u_2, \dots, u_k \mid u \in V\}$

generate $Y \sim \mu_k$

- sample $X \sim \mu$;
- if $X(u) = -$, then $Y(u_i) = -$ for all $i \in [k]$
- if $X(u) = +$, then
  - sample $i \in [k]$ u.a.r.
  - $Y(u_i) = +$ and $Y(u_j) = -$ for all $j \neq i$



$X \sim \mu$ $\qquad$ $Y \sim \mu_3$

Mixing of block dynamics [Chen, Liu and Vigoda 2021]

update of θ-*fraction block dynamics*

- sample θ fraction of variables R u.a.r.
- resample the value of R conditional on others

for any distribution π that is C-spectrally independent
$\lambda_{\text{gap}}^{\text{block}}(\pi) \geq \theta^{O(C)}$

k-transformation [This work]

μ over $\{-,+\}^V$ → μ_k over $\{-,+\}^{V_k}, V_k = \{u_1, u_2, \dots, u_k \mid u \in V\}$

Lemma I field dynamics on μ is the *limit instance* of block dynamics on μ_k

$$\lambda_{\text{gap}}^{\text{field}}(\mu,\theta) \geq \limsup_{k\to\infty} \lambda_{\text{gap}}^{\text{block}}(\mu_k)$$

Lemma II mixing of block dynamics:

μ is C-completely-SI → μ_k is $(C+2)$-SI for all k →[CLV21] $\lambda_{\text{gap}}^{\text{block}}(\mu_k) \geq \theta^{O(C)}$

Lemma I + Lemma II → **Mixing of field dynamics** $\lambda_{\text{gap}}^{\text{field}}(\mu,\theta) \geq \theta^{O(C)}$

Summary

- Optimal $\Omega(1/n)$ spectral gap for anti-ferro two-spin systems in the **uniqueness regime**
  - Example of applications: $O(n^2 \log n)$ mixing time for hardcore model
- A **boosting result** of spectral gap for completely spectrally independent distributions.
- A new Markov chain **field dynamics**
  - draw samples from target distribution
  - analyze Glauber dynamics

Thank you!

Open problem

- Prove the ***optimal*** $O(n \log n)$ mixing time for two spin systems in the uniqueness regime
  - $O(n\log n)$ mixing time for Ising model [CLV21,AJKPV21,CFYZ21]
  - $\tilde{O}(n)$-time sampling algorithm for hardcore model [AJKPV21]
- Extend our technique to ***general distributions*** beyond the Boolean domain i.e. q-coloring