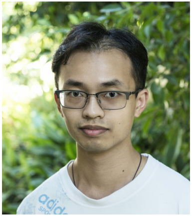
Uniqueness and Rapid Mixing
in the Bipartite Hardcore Model

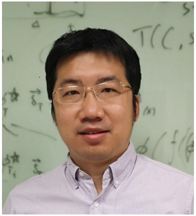
Xiaoyu Chen

Nanjing University

based on joint work with



Jingcheng Liu



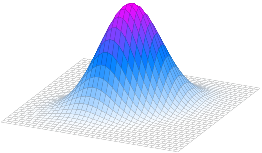
Yitong Yin

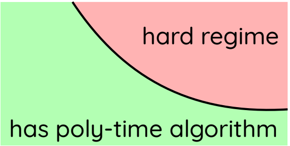
Sampling problem:

Draw (approximate) random samples from a distribution

Gibbs distribuiton:

- high-dimensional joint distribution
- described by few parameters and local interactions





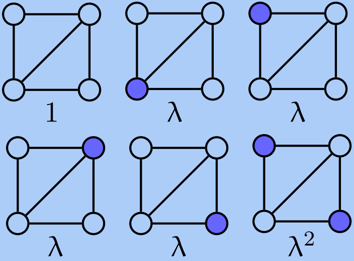
Computational phase transition:
computational complexity of sampling problem changes sharply around certain parameter values

Hardcore model

- $G = ([n], E)$ with n vertices and max degree Δ.
- Fugacity $\lambda > 0$ is a real number.
- $\mathrm{Ind}(G) = \{S \subseteq [n] \mid S \text{ is an independent set}\}$.
- Gibbs distribution

$$\forall S \in \mathrm{Ind}(G), \quad \mu(S) := \frac{\lambda^{|S|}}{Z}, \quad \text{where } Z_G(\lambda) := \sum_{I \in \mathrm{Ind}(G)} \lambda^{|I|}.$$

An example



Partition function:

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

This model is self-reducible

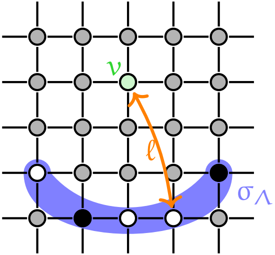
Computational phase transition

On general graph with maximum degree Δ:

uniqueness $\lambda_c(\Delta)$ non-uniqueness

$\lambda : 0$ ⟶ ∞

tree uniqueness threshold: $\lambda_c(\Delta) := (\Delta-1)^{(\Delta-1)}/(\Delta-2)^{\Delta} \approx \frac{e}{\Delta}$



Weak spatial mixing (WSM)

$\forall G, v, \mathbf{Pr}_{S\sim\mu}[v \in S \mid \sigma_\Lambda]$ does not depend on σ_Λ as $\ell \to +\infty$

WSM $\iff \lambda \leq \lambda_c(\Delta)$

σ_Λ : fixed configuration in Λ

easy $\lambda_c(\Delta)$ hard

$\lambda : 0$ ⟶ ∞

Computational phase transition:

- $\lambda < \lambda_c$: poly-time algorithm for approx. sampling [Wei06]
- $\lambda > \lambda_c$: no poly-time algorithm unless **NP = RP** [Sly10]

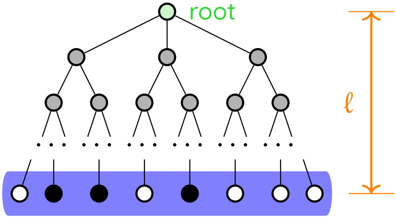
Computational phase transition

On general graph with maximum degree Δ:

uniqueness $\lambda_c(\Delta)$ non-uniqueness

$\lambda : 0 \longrightarrow \infty$

tree uniqueness threshold: $\lambda_c(\Delta) := (\Delta - 1)^{(\Delta-1)}/(\Delta-2)^{\Delta} \approx \frac{e}{\Delta}$



σ : boundary condition on level ℓ

Weak spatial mixing (WSM)

$\mathbf{Pr}_{S\sim\mu}[\text{root} \in S \mid \sigma]$ does not depend on σ when $\ell \to \infty$

WSM $\iff \lambda \le \lambda_c(\Delta)$

Δ-regular tree is the worst case [Wei06]

easy $\lambda_c(\Delta)$ hard

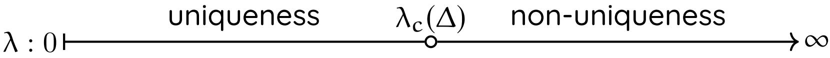
$\lambda : 0 \longrightarrow \infty$

Computational phase transition:

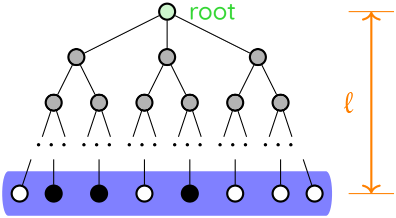
- $\lambda < \lambda_c$: poly-time algorithm for approx. sampling [Wei06]
- $\lambda > \lambda_c$: no poly-time algorithm unless **NP = RP** [Sly10]

Computational phase transition

On general graph with maximum degree Δ:



tree uniqueness threshold: $\lambda_c(\Delta) := (\Delta - 1)^{(\Delta-1)}/(\Delta-2)^{\Delta} \approx \frac{e}{\Delta}$



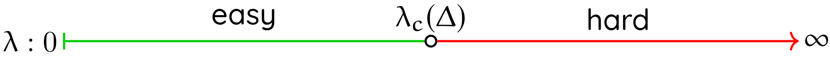
σ : boundary condition on level ℓ

Weak spatial mixing (WSM)

$\mathbf{Pr}_{S\sim\mu}[\text{root} \in S \mid \sigma]$ does not depend on σ when $\ell \to \infty$

WSM $\iff \lambda \le \lambda_c(\Delta)$

Δ-regular tree is the worst case [Wei06]



Computational phase transition:

- $\lambda < \lambda_c$: poly-time algorithm for approx. sampling [Wei06]
- $\lambda > \lambda_c$: no poly-time algorithm unless **NP = RP** [Sly10]

Hardcore model on bipartite graph (weighted #BIS)

It is easy: there is a poly-time algorithm to find a matximum independent set in the bipartite graph (Kőnig's theorem[1]).

It is hard: many important problems are proved to be #BIS-equivalent or #BIS-hard under AP-reductions.

Selected examples

- stable matchings (counting)
- ferro. Potts model (parti. func.)
- ferro. Ising with mixed external fields (parti. func.)

[DGGJ04, GJ07, DGJ10, CGM12 DGJR12, GJ12a, BDG+13, LLZ14, GJ15, CGG+16, GŠVY16, GGY21,]

Conjecture[DGGJ04]:
#BIS represents an intermediate complexity class:

- it has no FPRAS in general
- it is easier than #SAT

[1] In any bipartite graph, the number of edges in a maximum matching equals the number of vertices in a minimum vertex cover.

Hardcore model on bipartite graph (weighted #BIS)

It is easy: there is a poly-time algorithm to find a matximum independent set in the bipartite graph (Kőnig's theorem[1]).

It is hard: many important problems are proved to be #BIS-equivalent or #BIS-hard under AP-reductions.

Selected examples

- stable matchings (counting)
- ferro. Potts model (parti. func.)
- ferro. Ising with mixed external fields (parti. func.)

[DGGJ04, GJ07, DGJ10, CGM12 DGJR12, GJ12a, BDG+13, LLZ14, GJ15, CGG+16, GŠVY16, GGY21,]

Conjecture[DGGJ04]:
#BIS represents an intermediate complexity class:

- it has no FPRAS in general
- it is easier than #SAT

[1] In any bipartite graph, the number of edges in a maximum matching equals the number of vertices in a minimum vertex cover.

Hardcore model on bipartite graph (weighted #BIS)

It is easy: there is a poly-time algorithm to find a matximum independent set in the bipartite graph (Kőnig's theorem[1]).

It is hard: many important problems are proved to be #BIS-equivalent or #BIS-hard under AP-reductions.

Selected examples

- stable matchings (counting)
- ferro. Potts model (parti. func.)
- ferro. Ising with mixed external fields (parti. func.)

[DGGJ04, GJ07, DGJ10, CGM12 DGJR12, GJ12a, BDG+13, LLZ14, GJ15, CGG+16, GŠVY16, GGY21,]

Conjecture[DGGJ04]:
#BIS represents an intermediate complexity class:

- it has no FPRAS in general
- it is easier than #SAT

[1] In any bipartite graph, the number of edges in a maximum matching equals the number of vertices in a minimum vertex cover.

Hardcore model on bipartite graph (weighted #BIS)

It is easy: there is a poly-time algorithm to find a matximum independent set in the bipartite graph (Kőnig's theorem[1]).

It is hard: many important problems are proved to be #BIS-equivalent or #BIS-hard under AP-reductions.

Selected examples

- stable matchings (counting)
- ferro. Potts model (parti. func.)
- ferro. Ising with mixed external fields (parti. func.)

[DGGJ04, GJ07, DGJ10, CGM12 DGJR12, GJ12a, BDG+13, LLZ14, GJ15, CGG+16, GŠVY16, GGY21,]

Conjecture[DGGJ04]:

#BIS represents an intermediate complexity class:

- it has no FPRAS in general
- it is easier than #SAT

[1] In any bipartite graph, the number of edges in a maximum matching equals the number of vertices in a minimum vertex cover.

Previous algorithmic results

Non-uniqueness regime:

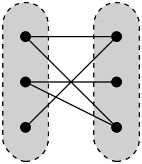
- α-expander bipartite graph:
  - $\lambda \geq (C_0\Delta)^{4/\alpha}$, an $n^{O(\log\Delta)}$ time sampler [JKP20]
  - $\lambda \geq (C_1\Delta)^{6/\alpha}$, an $O(n\log n)$ time sampler [CGG+21]
  - $\lambda \geq (C_2\Delta)^{2/\alpha}$, an $n^{O(\log\Delta)}$ time sampler [FGKP23]
- Δ-regular α-expander bipartite graph:
  - $\lambda \geq \frac{f(\alpha)\log\Delta}{\Delta^{1/4}}$, an $n^{O(\Delta)}$ time sampler [JPP22]
- random Δ-regular bipartite graph:
  - $\Delta \geq \Delta_0, \lambda \geq \frac{\log^4\Delta}{\Delta}$, an $n^{O(1)}$ time sampler [LLLM19]
  - $\Delta \geq \Delta_1, \lambda \geq \frac{50\log^2\Delta}{\Delta}$, an $n^{1+O(\frac{\log^2(\Delta)}{\Delta})}$ time sampler [JKP20]
  - $\Delta \geq \Delta_2, \lambda \geq \frac{100\log\Delta}{\Delta}$, an $O(n\log n)$ time sampler [CGŠV22]
- unbalanced bipartite graph:
  - $6\Delta_L\Delta_R\lambda \leq (1+\lambda)^{\frac{\delta_R}{\Delta_L}}$, an $n^{O(\log(\Delta_L\Delta_R))}$ time sampler [CP20]
  - $3.4\Delta_L\Delta_R\lambda \leq (1+\lambda)^{\frac{\delta_R}{\Delta_L}}$, an $n^{O(\log(\Delta_L\Delta_R))}$ time sampler [FGKP23]
  - $(1+e)\Delta_L\Delta_R\lambda \leq (1+\lambda)^{\frac{\delta_R}{\Delta_L}}$, an $O(n\log n)$ time sampler [BCP22]

Previous algorithmic results

Uniqueness regime:

- general graph: if $\lambda < \lambda_c(\Delta)$, there is an $O(n \log n)$ time sampler
- bipartite graph: if $\lambda = 1, \Delta_L \leq 5$, an $O(n^2)$ time sampler [LL15]
  ($\lambda = 1 \wedge \lambda < \lambda_c(\Delta) \Leftrightarrow \Delta \leq 5$)

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



Our results

For $\delta \in (0,1)$, $\Delta_L \geq 3$, if $\lambda \leq (1-\delta)\lambda_c(\Delta_L)$, then

- the system is in the uniqueness regime
- there is a sampler that runs in time

$$T := n\left(\frac{\Delta_L \log n}{\lambda}\right)^{O(1/\delta)}$$

- the **mixing time** of **Glauber dynamics** is bounded by $O(n^2) \cdot T$

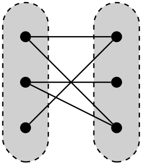
- When $\Delta_L = 1$, G is a forest, which is trivial.
- When $\Delta_L = 2$, this model becomes an Ising model. Our results still work, but since $\lambda_c(2) = \infty$, it is quite technical to state them here.

Our results

Uniqueness regime:

- general graph: if $\lambda < \lambda_c(\Delta)$, there is an $O(n \log n)$ time sampler
- bipartite graph: if $\lambda = 1, \Delta_L \le 5$, an $O(n^2)$ time sampler [LL15]
  $(\lambda = 1 \wedge \lambda < \lambda_c(\Delta) \Leftrightarrow \Delta \le 5)$

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



Our results

For $\delta \in (0,1)$, $\Delta_L \ge 3$, if $\lambda \le (1-\delta)\lambda_c(\Delta_L)$, then

- the system is in the uniqueness regime
- there is a sampler that runs in time
$$T := n \left(\frac{\Delta_L \log n}{\lambda} \right)^{O(1/\delta)}$$
- the **mixing time** of **Glauber dynamics** is bounded by $O(n^2) \cdot T$

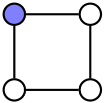
- When $\Delta_L = 1$, G is a forest, which is trivial.
- When $\Delta_L = 2$, this model becomes an Ising model. Our results still work, but since $\lambda_c(2) = \infty$, it is quite technical to state them here.

Our results

Glauber dynamics for Hardcore model:

start from an arbitrary independent set X_0;
for t from 1 to T **do**:

- pick a vertex $v \in V$ uniformly at random;
- with prob. $\frac{\lambda}{1+\lambda}$, let $S = X_{t-1} \cup \{v\}$;
  with prob. $\frac{1}{1+\lambda}$, let $S = X_{t-1} \setminus \{v\}$;
- **if** $S \in \mathrm{Ind}(G)$ **then** $X_t = S$ **else** $X_t = X_{t-1}$;



irreducible + aperiodic + reversible $\Longrightarrow X_t \sim \mu$ as $t \to \infty$

mixing time: essential running time of Glauber dynamics

$$T_{\mathrm{mix}} := \max_{X_0} \min\{t \mid D_{\mathrm{TV}}(X_t \parallel \mu) \le 1/100\}$$

total variation distance: conanical distance between distributions

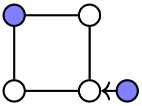
$$D_{\mathrm{TV}}(X_t \parallel \mu) := \frac{1}{2} \sum_{S \in \mathrm{Ind}(G)} |\mathbf{Pr}[X_t = S] - \mu(S)|$$

Our results

Glauber dynamics for Hardcore model:

start from an arbitrary independent set X_0;
for t from 1 to T **do**:

- pick a vertex $v \in V$ uniformly at random;
- with prob. $\frac{\lambda}{1+\lambda}$, let $S = X_{t-1} \cup \{v\}$;
  with prob. $\frac{1}{1+\lambda}$, let $S = X_{t-1} \setminus \{v\}$;
- **if** $S \in \mathrm{Ind}(G)$ **then** $X_t = S$ **else** $X_t = X_{t-1}$;



irreducible + aperiodic + reversible $\Longrightarrow X_t \sim \mu$ as $t \to \infty$

mixing time: essential running time of Glauber dynamics

$$T_{\mathrm{mix}} := \max_{X_0} \min\{t \mid D_{\mathrm{TV}}(X_t \parallel \mu) \le 1/100\}$$

total variation distance: conanical distance between distributions

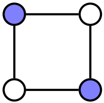
$$D_{\mathrm{TV}}(X_t \parallel \mu) := \frac{1}{2} \sum_{S \in \mathrm{Ind}(G)} |\mathbf{Pr}[X_t = S] - \mu(S)|$$

Our results

Glauber dynamics for Hardcore model:

start from an arbitrary independent set X_0;
for t from 1 to T **do**:

- pick a vertex $v \in V$ uniformly at random;
- with prob. $\frac{\lambda}{1+\lambda}$, let $S = X_{t-1} \cup \{v\}$;
  with prob. $\frac{1}{1+\lambda}$, let $S = X_{t-1} \setminus \{v\}$;
- **if** $S \in \mathrm{Ind}(G)$ **then** $X_t = S$ **else** $X_t = X_{t-1}$;



irreducible + aperiodic + reversible $\Longrightarrow X_t \sim \mu$ as $t \to \infty$

mixing time: essential running time of Glauber dynamics

$$T_{\text{mix}} := \max_{X_0} \min\{t \mid D_{\text{TV}}(X_t \parallel \mu) \le 1/100\}$$

total variation distance: conanical distance between distributions

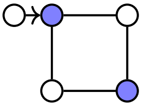
$$D_{\text{TV}}(X_t \parallel \mu) := \frac{1}{2} \sum_{S \in \mathrm{Ind}(G)} \left|\mathbf{Pr}\left[X_t = S\right] - \mu(S)\right|$$

Our results

Glauber dynamics for Hardcore model:

start from an arbitrary independent set X_0;
for t from 1 to T **do**:

- pick a vertex $v \in V$ uniformly at random;
- with prob. $\frac{\lambda}{1+\lambda}$, let $S = X_{t-1} \cup \{v\}$;
  with prob. $\frac{1}{1+\lambda}$, let $S = X_{t-1} \setminus \{v\}$;
- **if** $S \in \mathrm{Ind}(G)$ **then** $X_t = S$ **else** $X_t = X_{t-1}$;



irreducible + aperiodic + reversible $\Longrightarrow X_t \sim \mu$ as $t \to \infty$

mixing time: essential running time of Glauber dynamics

$$T_{\mathrm{mix}} := \max_{X_0} \min\{t \mid D_{\mathrm{TV}}(X_t \parallel \mu) \le 1/100\}$$

total variation distance: conanical distance between distributions

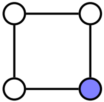
$$D_{\mathrm{TV}}(X_t \parallel \mu) := \frac{1}{2} \sum_{S \in \mathrm{Ind}(G)} |\mathbf{Pr}[X_t = S] - \mu(S)|$$

Our results

Glauber dynamics for Hardcore model:

start from an arbitrary independent set X_0;
for t from 1 to T **do**:

- pick a vertex $v \in V$ uniformly at random;
- with prob. $\frac{\lambda}{1+\lambda}$, let $S = X_{t-1} \cup \{v\}$;
  with prob. $\frac{1}{1+\lambda}$, let $S = X_{t-1} \setminus \{v\}$;
- **if** $S \in \mathrm{Ind}(G)$ **then** $X_t = S$ **else** $X_t = X_{t-1}$;



irreducible + aperiodic + reversible $\Longrightarrow X_t \sim \mu$ as $t \to \infty$

mixing time: essential running time of Glauber dynamics

$$T_{\mathrm{mix}} := \max_{X_0} \min\{t \mid D_{\mathrm{TV}}(X_t \parallel \mu) \le 1/100\}$$

total variation distance: conanical distance between distributions

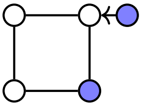
$$D_{\mathrm{TV}}(X_t \parallel \mu) := \frac{1}{2} \sum_{S \in \mathrm{Ind}(G)} \left|\mathbf{Pr}\left[X_t = S\right] - \mu(S)\right|$$

Our results

Glauber dynamics for Hardcore model:

start from an arbitrary independent set X_0;
for t from 1 to T **do**:

- pick a vertex $v \in V$ uniformly at random;
- with prob. $\frac{\lambda}{1+\lambda}$, let $S = X_{t-1} \cup \{v\}$;
  with prob. $\frac{1}{1+\lambda}$, let $S = X_{t-1} \setminus \{v\}$;
- **if** $S \in \mathrm{Ind}(G)$ **then** $X_t = S$ **else** $X_t = X_{t-1}$;



irreducible + aperiodic + reversible $\Longrightarrow X_t \sim \mu$ as $t \to \infty$

mixing time: essential running time of Glauber dynamics

$$T_{\mathrm{mix}} := \max_{X_0} \min\{t \mid D_{\mathrm{TV}}(X_t \parallel \mu) \le 1/100\}$$

total variation distance: conanical distance between distributions

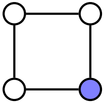
$$D_{\mathrm{TV}}(X_t \parallel \mu) := \frac{1}{2} \sum_{S \in \mathrm{Ind}(G)} \left| \mathbf{Pr}\left[X_t = S\right] - \mu(S) \right|$$

Our results

Glauber dynamics for Hardcore model:

start from an arbitrary independent set X_0;
for t from 1 to T **do**:

- pick a vertex $v \in V$ uniformly at random;
- with prob. $\frac{\lambda}{1+\lambda}$, let $S = X_{t-1} \cup \{v\}$;
  with prob. $\frac{1}{1+\lambda}$, let $S = X_{t-1} \setminus \{v\}$;
- **if** $S \in \mathrm{Ind}(G)$ **then** $X_t = S$ **else** $X_t = X_{t-1}$;



irreducible + aperiodic + reversible $\Longrightarrow X_t \sim \mu$ as $t \to \infty$

mixing time: essential running time of Glauber dynamics

$$T_{\mathrm{mix}} := \max_{X_0} \min\{t \mid D_{\mathrm{TV}}(X_t \parallel \mu) \le 1/100\}$$

total variation distance: conanical distance between distributions

$$D_{\mathrm{TV}}(X_t \parallel \mu) := \frac{1}{2} \sum_{S \in \mathrm{Ind}(G)} \left|\mathbf{Pr}\left[X_t = S\right] - \mu(S)\right|$$

Background

Proof outline

Fast sampler

Mixing of Glauber dynamics on $\mathbb{L} \cup \mathbb{R}$

Background

Let ν be a distribution over $\Omega = \{-1, +1\}^n$. $\forall \sigma \in \Omega, \quad \|\sigma\|_+ = |\{i \mid \sigma_i = 1\}|$

impose external field $\theta > 0$

$\theta * \nu$: a distribution on Ω:

$$\forall \sigma, \quad (\theta * \nu)(\sigma) \propto \nu(\sigma) \cdot \theta^{\|\sigma\|_+}$$

flip the distribution

$\overline{\nu}$: a distribution on Ω:

$$\forall \sigma, \quad \overline{\nu}(\sigma) = \nu(-\sigma)$$

- hardcore model: μ (fugacity λ) $\Longrightarrow \theta * \mu$ (fugacity $\theta\lambda$)

For $0 < \theta \neq 1$, Field dynamics $P^{\mathrm{FD}}_{\theta,\nu}$: Markov chain $(X_t)_{t \geq 0}$ on Ω:

X_0 is an arbitrary vector in Ω and let $s \in \{-1, +1\}$ so that $\theta^s < 1$;
for each $t > 0$:

1. **generate** $R \subseteq [n]$: **for** $i \in [n]$ with $X_{t-1}(i) = s$
   add i to R with prob. $1 - \theta^s$
2. **let** $X_t = \sigma$ with prob. $\mathbf{Pr}_{\sigma \sim \theta * \nu}[\sigma \mid \sigma_R = s]$

irreducible + aperiodic + reversible [CFYZ21] $\Longrightarrow \quad X_t \sim \nu \quad$ as $t \to \infty$ 🤔

rapid mixing of $P^{\mathrm{FD}}_{\theta,\nu}$ + sampler for $\theta * \nu$ = sampler for ν

Background

Let ν be a distribution over $\Omega = \{-1,+1\}^n$. $\forall \sigma \in \Omega, \quad \|\sigma\|_+ = |\{i \mid \sigma_i = 1\}|$

impose external field $\theta > 0$

$\theta * \nu$: a distribution on Ω:

$$\forall \sigma, \quad (\theta * \nu)(\sigma) \propto \nu(\sigma) \cdot \theta^{\|\sigma\|_+}$$

flip the distribution

$\overline{\nu}$: a distribution on Ω:

$$\forall \sigma, \quad \overline{\nu}(\sigma) = \nu(-\sigma)$$

- hardcore model: μ (fugacity λ) $\Longrightarrow \theta * \mu$ (fugacity $\theta\lambda$)

For $0 < \theta \neq 1$, Field dynamics $P^{\mathrm{FD}}_{\theta,\nu}$: Markov chain $(X_t)_{t\geq 0}$ on Ω:

X_0 is an arbitrary vector in Ω and let $s \in \{-1,+1\}$ so that $\theta^s < 1$;
for each $t > 0$:

1. **generate** $R \subseteq [n]$: **for** $i \in [n]$ with $X_{t-1}(i) = s$
   add i to R with prob. $1 - \theta^s$
2. **let** $X_t = \sigma$ with prob. $\mathbf{Pr}_{\sigma\sim\theta*\nu}[\sigma \mid \sigma_R = s]$

irreducible + aperiodic + reversible [CFYZ21] $\Longrightarrow$ $X_t \sim \nu$ as $t \to \infty$ 🤔

rapid mixing of $P^{\mathrm{FD}}_{\theta,\nu}$ + sampler for $\theta * \nu$ = sampler for ν

Background

Let ν be a distribution over $\Omega = \{-1, +1\}^n$. $\forall \sigma \in \Omega, \quad \|\sigma\|_+ = |\{i \mid \sigma_i = 1\}|$

impose external field $\theta > 0$

$\theta * \nu$: a distribution on Ω:

$$\forall \sigma, \quad (\theta * \nu)(\sigma) \propto \nu(\sigma) \cdot \theta^{\|\sigma\|_+}$$

flip the distribution

$\overline{\nu}$: a distribution on Ω:

$$\forall \sigma, \quad \overline{\nu}(\sigma) = \nu(-\sigma)$$

- hardcore model: μ (fugacity λ) $\Longrightarrow$ $\theta * \mu$ (fugacity $\theta\lambda$)

For $0 < \theta \neq 1$, Field dynamics $P^{\mathrm{FD}}_{\theta,\nu}$: Markov chain $(X_t)_{t \geq 0}$ on Ω:

X_0 is an arbitrary vector in Ω and let $s \in \{-1, +1\}$ so that $\theta^s < 1$;
for each $t > 0$:

1. **generate** $R \subseteq [n]$: **for** $i \in [n]$ with $X_{t-1}(i) = s$
   add i to R with prob. $1 - \theta^s$
2. **let** $X_t = \sigma$ with prob. $\mathbf{Pr}_{\sigma \sim \theta * \nu}[\sigma \mid \sigma_R = s]$

irreducible + aperiodic + reversible [CFYZ21] $\Longrightarrow$ $X_t \sim \nu$ as $t \to \infty$ 🤔

rapid mixing of $P^{\mathrm{FD}}_{\theta,\nu}$ + sampler for $\theta * \nu$ = sampler for ν

Background

Theorem ([CFYZ21, AJKPV22, CFYZ22, CE22])

Let $0 < \theta \neq 1$ and ν be a distribution over $\{-1, +1\}^n$ that

1. $\lambda * \nu$ is K-**marginally stable** for all λ between θ, 1,
2. $\lambda * \nu$ is η-**spectrally independent** for all λ between θ, 1,
3. the Glauber dynamics on $\theta * \nu$ mixes in time $\widetilde{O}(n)$,

then

$$1 \wedge 2 \implies T_{\text{mix}}(P^{\text{FD}}_{\theta,\nu}) \approx \max\{\theta, 1/\theta\}^{\eta \cdot \text{poly}(K)}.$$

$$1 \wedge 2 \wedge 3 \implies \text{sampler for } \nu \text{ in time } \widetilde{O}(n) \cdot \max\{\theta, 1/\theta\}^{\eta \cdot \text{poly}(K)}$$

$$1 \wedge 2 \wedge 3 \overset{\text{Var}}{\implies} T_{\text{mix}}(P^{\text{GD}}_{\nu}) \approx \widetilde{O}(n) \cdot \underbrace{n \cdot \max\{\theta, 1/\theta\}^{\eta \cdot \text{poly}(K)}}_{\text{relaxation time}}$$

Background

Let ν be a distribution over $\{-1,+1\}^n$ and $X \sim \nu$ be a random vector.

influence matrix $\Psi_\nu \in \mathbb{R}^{n\times n}$

$$\Psi_\nu(i,j) := \begin{cases} 0, & \text{if } \mathbf{Pr}_\nu[i] \in \{0,1\} \\ \mathbf{Pr}_\nu[j \mid i] - \mathbf{Pr}_\nu\left[j \mid \bar{i}\right] \end{cases}$$

$$i = \{X_i = +1\}, \bar{i} = \{X_i = -1\}$$

$\mathrm{Corr}(X) \in \mathbb{R}^{n\times n}$

$$\mathrm{Corr}(X)_{ij} = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}(X_i)\mathrm{Var}(X_j)}}$$

$$\Psi_\nu(i,j) = \frac{\mathrm{Cov}(X_i, X_j)}{\mathrm{Var}(X_i)}$$

- Ψ_ν is similar to $\mathrm{Corr}(X)$

η-spectral independence (in ∞-norm)

$\forall \Lambda \subseteq [n]$ with $|\Lambda| \le n-2$, and $\forall \tau \in \Omega(\nu_\Lambda)$, $\|\Psi_{\nu^\tau}\|_\infty \le \eta$

K-marginal stability

there is $\rho \in \{\nu, \bar{\nu}\}$ that for $i \in [n]$, $S \subseteq \Lambda \subseteq [n] \setminus \{i\}$, $\tau \in \Omega(\rho_\Lambda)$,

$$R_i^\tau \le K \cdot R_i^{\tau_S} \text{ and } \rho_i^\tau(-1) \ge K^{-1}$$

- marginal ratio $R_i^\tau = \frac{\rho_i^\tau(+1)}{\rho_i^\tau(-1)}$

Background

Let ν be a distribution over $\{-1,+1\}^n$ and $X \sim \nu$ be a random vector.

influence matrix $\Psi_\nu \in \mathbb{R}^{n\times n}$

$$\Psi_\nu(i,j) := \begin{cases} 0, & \text{if } \mathbf{Pr}_\nu[i] \in \{0,1\} \\ \mathbf{Pr}_\nu[j \mid i] - \mathbf{Pr}_\nu[j \mid \bar{i}] \end{cases}$$

$i = \{X_i = +1\}, \bar{i} = \{X_i = -1\}$

$\mathrm{Corr}(X) \in \mathbb{R}^{n\times n}$

$$\mathrm{Corr}(X)_{ij} = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}(X_i)\mathrm{Var}(X_j)}}$$

$$\Psi_\nu(i,j) = \frac{\mathrm{Cov}(X_i, X_j)}{\mathrm{Var}(X_i)}$$

- Ψ_ν is similar to $\mathrm{Corr}(X)$

η-spectral independence (in ∞-norm)

$\forall \Lambda \subseteq [n]$ with $|\Lambda| \le n-2$, and $\forall \tau \in \Omega(\nu_\Lambda)$, $\|\Psi_{\nu^\tau}\|_\infty \le \eta$

K-marginal stability

there is $\rho \in \{\nu, \bar{\nu}\}$ that for $i \in [n]$, $S \subseteq \Lambda \subseteq [n] \setminus \{i\}$, $\tau \in \Omega(\rho_\Lambda)$,

$$R_i^\tau \le K \cdot R_i^{\tau_S} \text{ and } \rho_i^\tau(-1) \ge K^{-1}$$

- marginal ratio $R_i^\tau = \dfrac{\rho_i^\tau(+1)}{\rho_i^\tau(-1)}$

Background

Let ν be a distribution over $\{-1,+1\}^n$ and $X \sim \nu$ be a random vector.

influence matrix $\Psi_\nu \in \mathbb{R}^{n\times n}$

$$\Psi_\nu(i,j) := \begin{cases} 0, & \text{if } \mathbf{Pr}_\nu[i] \in \{0,1\} \\ \mathbf{Pr}_\nu[j \mid i] - \mathbf{Pr}_\nu\left[j \mid \bar{i}\right] \end{cases}$$

$$i = \{X_i = +1\}, \bar{i} = \{X_i = -1\}$$

$\mathrm{Corr}(X) \in \mathbb{R}^{n\times n}$

$$\mathrm{Corr}(X)_{ij} = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}(X_i)\mathrm{Var}(X_j)}}$$

$$\Psi_\nu(i,j) = \frac{\mathrm{Cov}(X_i, X_j)}{\mathrm{Var}(X_i)}$$

- Ψ_ν is similar to $\mathrm{Corr}(X)$

η-spectral independence (in ∞-norm)

$\forall \Lambda \subseteq [n]$ with $|\Lambda| \le n-2$, and $\forall \tau \in \Omega(\nu_\Lambda)$, $\|\Psi_{\nu^\tau}\|_\infty \le \eta$

K-marginal stability

there is $\rho \in \{\nu, \overline{\nu}\}$ that for $i \in [n]$, $S \subseteq \Lambda \subseteq [n] \setminus \{i\}$, $\tau \in \Omega(\rho_\Lambda)$,

$$R_i^\tau \le K \cdot R_i^{\tau_S} \text{ and } \rho_i^\tau(-1) \ge K^{-1}$$

- marginal ratio $R_i^\tau = \dfrac{\rho_i^\tau(+1)}{\rho_i^\tau(-1)}$

Background

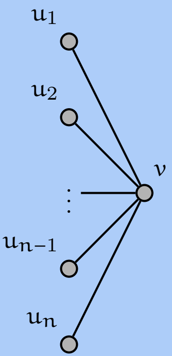
Proof outline

Fast sampler

Mixing of Glauber dynamics on $L \cup R$

Proof outline

example



Let $\lambda = 1$ be the fugacity
μ: Gibbs distribution of the hardcore model

$\|\Psi_\mu\|_\infty$ is unbounded

- $\forall i, \quad |\Psi_\mu(v, u_i)| = \frac{\lambda}{\lambda+1} = \frac{1}{2}$
- $\|\Psi_\mu\|_\infty \geq \sum_i |\Psi_\mu(v, u_i)| = \frac{n}{2}$

What could we do? 🤔

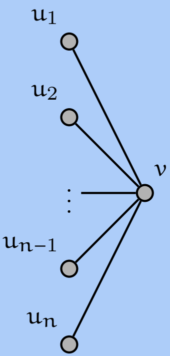
$\|\Psi_{\mu_L}\|_\infty$ is bounded

- $|\Psi_\mu(u_1, u_2)| = \frac{\lambda}{1+\lambda} - \frac{\lambda(1+\lambda)^{n-2}}{\lambda+(1+\lambda)^{n-1}} = \frac{1}{2^n+2}$
- $\|\Psi_{\mu_L}\|_\infty = \sum_{i\geq 2} |\Psi_\mu(u_1, u_i)| = O(1)$

Maybe we could take $\nu = \mu_L$.

Proof outline

example



Let $\lambda = 1$ be the fugacity
μ: Gibbs distribution of the hardcore model

$\|\Psi_\mu\|_\infty$ is unbounded

- $\forall i, \quad |\Psi_\mu(v, u_i)| = \frac{\lambda}{\lambda+1} = \frac{1}{2}$
- $\|\Psi_\mu\|_\infty \geq \sum_i |\Psi_\mu(v, u_i)| = \frac{n}{2}$

What could we do? 🤔

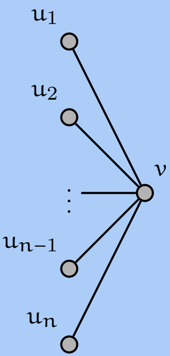
$\|\Psi_{\mu_L}\|_\infty$ is bounded

- $|\Psi_\mu(u_1, u_2)| = \frac{\lambda}{1+\lambda} - \frac{\lambda(1+\lambda)^{n-2}}{\lambda+(1+\lambda)^{n-1}} = \frac{1}{2^n+2}$
- $\|\Psi_{\mu_L}\|_\infty = \sum_{i \geq 2} |\Psi_\mu(u_1, u_i)| = O(1)$

Maybe we could take $\nu = \mu_L$.

Proof outline

example



Let $\lambda = 1$ be the fugacity
μ: Gibbs distribution of the hardcore model

$\|\Psi_\mu\|_\infty$ is unbounded

- $\forall i, \quad |\Psi_\mu(v, u_i)| = \frac{\lambda}{\lambda+1} = \frac{1}{2}$
- $\|\Psi_\mu\|_\infty \geq \sum_i |\Psi_\mu(v, u_i)| = \frac{n}{2}$

What could we do? 🤔

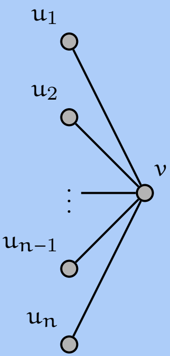
$\|\Psi_{\mu_L}\|_\infty$ is bounded

- $|\Psi_\mu(u_1, u_2)| = \frac{\lambda}{1+\lambda} - \frac{\lambda(1+\lambda)^{n-2}}{\lambda+(1+\lambda)^{n-1}} = \frac{1}{2^n+2}$
- $\|\Psi_{\mu_L}\|_\infty = \sum_{i\geq 2} |\Psi_\mu(u_1, u_i)| = O(1)$

Maybe we could take $\nu = \mu_L$.

Proof outline

example



Let $\lambda = 1$ be the fugacity
μ: Gibbs distribution of the hardcore model

$\|\Psi_\mu\|_\infty$ is unbounded

- $\forall i, \quad |\Psi_\mu(v, u_i)| = \frac{\lambda}{\lambda+1} = \frac{1}{2}$
- $\|\Psi_\mu\|_\infty \geq \sum_i |\Psi_\mu(v, u_i)| = \frac{n}{2}$

What could we do? 🤔

$\|\Psi_{\mu_L}\|_\infty$ is bounded

- $|\Psi_\mu(u_1, u_2)| = \frac{\lambda}{1+\lambda} - \frac{\lambda(1+\lambda)^{n-2}}{\lambda+(1+\lambda)^{n-1}} = \frac{1}{2^n+2}$
- $\|\Psi_{\mu_L}\|_\infty = \sum_{i\geq 2} |\Psi_\mu(u_1, u_i)| = O(1)$

Maybe we could take $\nu = \mu_L$.

Proof outline: fast sampler for μ

μ is the Gibbs distribution of the hardcore model and ν is μ_L

$$\nu \xrightarrow[\qquad P^{FD}_{\theta,\nu} \text{ with } \theta = \Theta(\frac{\Delta \log n}{\lambda}) > 1 \qquad]{O(1/\delta)\text{-spectrally independent} \qquad O(1)\text{-marginally stable}} \theta * \nu$$

BHC(λ, λ) — BHC$(\theta\lambda, \lambda)$

Glauber dynamics mixes in $\widetilde{O}(n)$

- fast sampler for ν in time $n \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$ ($\Rightarrow$ fast sampler for μ)
- Glauber dynamics on ν mixes in time $n^2 \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$

For $\nu = \mu_L$ on BHC(λ, α): δ-uniqueness $\Longrightarrow$ $O(1/\delta)$-spectral independence

BHC(λ, α)

uniqueness (boundary)

Let $d = \Delta_L - 1$, this parametric curve is the boundary of our uniqueness regime, for $w > d^{-1}$:

$$\begin{cases} \alpha(w) = \frac{d^w (w+1)^{w+1}}{(dw-1)^{w+1}} \\ \lambda(w) = \frac{w^d (d+1)^{d+1}}{(dw-1)^{d+1}} \end{cases}$$

Proof outline: fast sampler for μ

μ is the Gibbs distribution of the hardcore model and ν is μ_L

$$\nu \xrightarrow[\text{}]{O(1/\delta)\text{-spectrally independent} \qquad\qquad O(1)\text{-marginally stable}} \theta * \nu$$

BHC(λ, λ) $\qquad$ $P^{\mathrm{FD}}_{\theta,\nu}$ with $\theta = \Theta(\frac{\Delta \log n}{\lambda}) > 1$ $\qquad$ BHC$(\theta\lambda, \lambda)$

Glauber dynamics mixes in $\widetilde{O}(n)$

- fast sampler for ν in time $n \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$ ($\Rightarrow$ fast sampler for μ)
- Glauber dynamics on ν mixes in time $n^2 \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$

For $\nu = \mu_L$ on BHC(λ, α): $\qquad$ δ-uniqueness $\Longrightarrow$ $O(1/\delta)$-spectral independence

BHC(λ, α)

$\Delta_L = 3$

uniqueness (boundary)

Let $d = \Delta_L - 1$, this parametric curve is the boundary of our uniqueness regime, for $w > d^{-1}$:

$$\begin{cases} \alpha(w) = \frac{d^w (w+1)^{w+1}}{(dw-1)^{w+1}} \\ \lambda(w) = \frac{w^d (d+1)^{d+1}}{(dw-1)^{d+1}} \end{cases}$$

Proof outline: fast sampler for μ

μ is the Gibbs distribution of the hardcore model and ν is μ_L

$$\nu \xrightarrow{\ O(1/\delta)\text{-spectrally independent} \qquad\qquad O(1)\text{-marginally stable ✅}\ } \theta * \nu$$

BHC(λ, λ) $\qquad$ $P^{\mathsf{FD}}_{\theta,\nu}$ with $\theta = \Theta(\frac{\Delta \log n}{\lambda}) > 1$ $\qquad$ BHC$(\theta\lambda, \lambda)$

Glauber dynamics mixes in $\widetilde{O}(n)$✅

- fast sampler for ν in time $n \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$ ($\Rightarrow$ fast sampler for μ)
- Glauber dynamics on ν mixes in time $n^2 \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$

For $\nu = \mu_L$ on BHC(λ, α): $\qquad$ δ-uniqueness $\Longrightarrow$ $O(1/\delta)$-spectral independence

uniqueness (boundary)

Let $d = \Delta_L - 1$, this parametric curve is the boundary of our uniqueness regime, for $w > d^{-1}$:

$$\begin{cases} \alpha(w) = \dfrac{d^{w}(w+1)^{w+1}}{(dw-1)^{w+1}} \\[2ex] \lambda(w) = \dfrac{w^{d}(d+1)^{d+1}}{(dw-1)^{d+1}} \end{cases}$$

Proof outline: fast sampler for μ

μ is the Gibbs distribution of the hardcore model and ν is μ_L

$$\nu \xrightarrow[]{O(1/\delta)\text{-spectrally independent} \qquad\qquad\qquad O(1)\text{-marginally stable ✅}} \theta * \nu$$

BHC(λ, λ) $\qquad$ $P^{\mathrm{FD}}_{\theta,\nu}$ with $\theta = \Theta(\frac{\Delta \log n}{\lambda}) > 1$ $\qquad$ BHC$(\theta\lambda, \lambda)$

Glauber dynamics mixes in $\widetilde{O}(n)$✅

- fast sampler for ν in time $n \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$ ($\Rightarrow$ fast sampler for μ)
- Glauber dynamics on ν mixes in time $n^2 \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$

For $\nu = \mu_L$ on BHC(λ, α): $\qquad$ δ-uniqueness $\Longrightarrow$ $O(1/\delta)$-spectral independence

uniqueness (boundary)

Let $d = \Delta_L - 1$, this parametric curve is the boundary of our uniqueness regime, for $w > d^{-1}$:

$$\begin{cases} \alpha(w) = \frac{d^{w}(w+1)^{w+1}}{(dw-1)^{w+1}} \\ \lambda(w) = \frac{w^{d}(d+1)^{d+1}}{(dw-1)^{d+1}} \end{cases}$$

Proof outline: fast sampler for μ

μ is the Gibbs distribution of the hardcore model and ν is μ_L

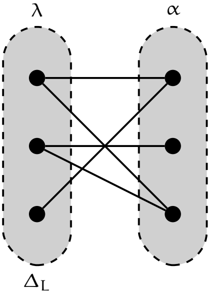
$$\nu \xrightarrow[\qquad]{O(1/\delta)\text{-spectrally independent} \qquad\qquad O(1)\text{-marginally stable ✅}} \theta * \nu$$

BHC(λ, λ) $\qquad$ $P^{\mathrm{FD}}_{\theta,\nu}$ with $\theta = \Theta(\frac{\Delta \log n}{\lambda}) > 1$ $\qquad$ BHC$(\theta\lambda, \lambda)$

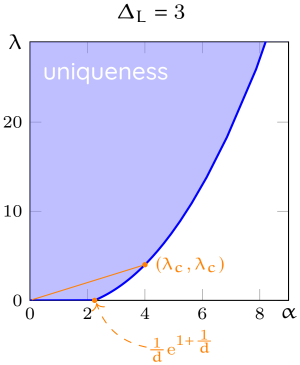
Glauber dynamics mixes in $\widetilde{O}(n)$✅

- fast sampler for ν in time $n \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$ ($\Rightarrow$ fast sampler for μ)
- Glauber dynamics on ν mixes in time $n^2 \cdot (\frac{\Delta \log n}{\lambda})^{O(1/\delta)}$

For $\nu = \mu_L$ on BHC(λ, α): $\qquad$ δ-uniqueness $\Longrightarrow$ $O(1/\delta)$-spectral independence



BHC(λ, α)



uniqueness (boundary)

Let $d = \Delta_L - 1$, this parametric curve is the boundary of our uniqueness regime, for $w > d^{-1}$:

$$\begin{cases} \alpha(w) = \dfrac{d^{w}(w+1)^{w+1}}{(dw-1)^{w+1}} \\ \lambda(w) = \dfrac{w^{d}(d+1)^{d+1}}{(dw-1)^{d+1}} \end{cases}$$

Background

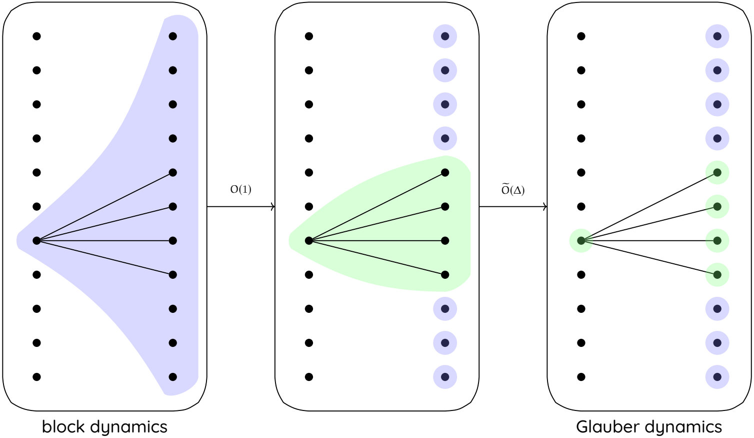
Proof outline

Fast sampler

Mixing of Glauber dynamics on $L \cup R$

Proof outline: mixing of GD on μ

- Glauber dynamics on $\nu = \mu_L$ is rapidly mixing.
- It works like a block dynamics that update a random vertex on the left and all the vertices on the right in each step.
- We finish the proof by comparing it and the Glauber dynamics on μ via the **block factorization** [CMT15, CP20, CLV21].



Thank you

arXiv:2305.00186

Summary

For $\delta \in (0,1)$, $\Delta_L \geq 3$, if $\lambda \leq (1-\delta)\lambda_c(\Delta_L)$, then

- the system is in the uniqueness regime
- there is a sampler that runs in time
$$T := n\left(\frac{\Delta_L \log n}{\lambda}\right)^{O(1/\delta)}$$
- the mixing time of Glauber dynamics is bounded by $O(n^2) \cdot T$

Open problems

- Remove the depedency on Δ_L in the running time of the sampler.
- Better mixing time for the Glauber dynamics.
- Bipartite hardcore model for negative λ.