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

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

**Sampling problem**

Draw (approximate) random samples from distribution $\mu$

**Computational phase transition**

Computational complexity of sampling problem changes sharply around some parameters of $\mu$
Hardcore gas model

- Graph $G = (V, E)$: $n$-vertex and max degree $\Delta$;
- 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 $\mu$:
  $$\forall X \in \Omega, \quad \mu(X) \propto w(X) = \lambda^{|X|_+}.$$  
  $|X|_+ = \text{number of occupied vertices } (X_v = +)$

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

$$\mu \left( \begin{array}{c}
\text{vertex} \\
\text{configuration}
\end{array} \right) = \frac{\lambda^2}{1 + 4\lambda + \lambda^2}$$
Uniqueness Threshold

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

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

\( \Delta \): maximum degree

---

**Computational phase transition**

- \( \lambda < \lambda_c \): **poly-time algorithm** for sampling [Weitz06]
- \( \lambda > \lambda_c \): **no poly-time algorithm** unless \( NP = RP \) [Sly10]
Computational phase transition

• $\lambda \leq (1 - \delta)\lambda_c$: $n^{O\left(\frac{\log \Delta}{\delta}\right)}$-time algorithms for sampling (via approx. counting) [Weitz06]

• $\lambda > \lambda_c$: no poly-time algorithm unless $NP = RP$ [Sly10]

Uniqueness Threshold

Pr[$X(root) = + | \sigma$] is independent of $\sigma$ if $\ell \to \infty$

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

$\Delta$: maximum degree
Problem: *fixed parameter trackable* sampling algorithm for hardcore model

*Let* $\delta > 0$ *be an arbitrary gap*. For any hardcore model with $\lambda \leq (1 - \delta)\lambda_c(\Delta)$, can we sample from Gibbs distribution in time $C(\delta) \cdot \text{poly}(n)$?
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
  $X_v = \begin{cases} 
  + & \text{w. p. } \lambda/(1 + \lambda) \\
  - & \text{w. p. } 1/(1 + \lambda) 
  \end{cases}$
- Else $X_v \leftarrow -$ 

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

$d_{TV}(X_t, \mu)$: the total variation distance between $X_t$ and $\mu$. 
## Previous works

### Work and Condition

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

---

### Diagram

- $\lambda_c(\Delta) \approx \frac{e}{\Delta}$
- Hard regime

- $\frac{2}{\Delta - 2}$ [LV99]
- $\frac{1}{\Delta - 1}$ [Dob70]
- Special graph [EHŠVY16]
- General graph [ALO20, CLV20, CLV21]
## Our results

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

<table>
<thead>
<tr>
<th>Work</th>
<th>Condition</th>
<th>Mixing Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anari, Liu, Oveis Gharan 2020</td>
<td>$\lambda \leq (1 - \delta) \lambda_c(\Delta)$</td>
<td>$n^{O(1/\delta)}$</td>
</tr>
<tr>
<td>Improved by Chen, Liu, Vigoda 2020</td>
<td>$\lambda \leq (1 - \delta) \lambda_c(\Delta)$</td>
<td>$\Delta^{O(\Delta^2/\delta)} n \log n$</td>
</tr>
<tr>
<td>Chen, Liu, Vigoda 2021</td>
<td>$\lambda \leq (1 - \delta) \lambda_c(\Delta)$</td>
<td>$\Delta^{O(\Delta^2/\delta)} n \log n$</td>
</tr>
</tbody>
</table>

**Our Result**

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

$\exp\left(\Theta\left(\frac{1}{\delta}\right)\right) \cdot n^2 \log n$

---

**Theorem (hardcore model)** [this work]

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

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

Anti-ferro two-spin systems [this work]

For anti-ferro two-spin system that is up-to-$\Delta$ unique,

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

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 \[\rightarrow\] complete spectral independence

mixing result in \textit{sub-critical} regime \[\rightarrow\] mixing result in \textit{near-critical} regime

\textit{boosting} hard regime

uniqueness regime (near-critical)

Dobrushin regime (sub-critical)
Spectral gap and mixing time

**Transition matrix of Glauber dynamics** : \( P : \Omega \times \Omega \rightarrow \mathbb{R}_{\geq 0} \)

**Eigenvalues** : \( P \) has \(|\Omega|\) non-negative real eigenvalues

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

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

\[ T_{\text{mix}} = O\left( \frac{1}{\lambda_{\text{gap}}} \log \frac{1}{\mu_{\text{min}}} \right), \quad \mu_{\text{min}} = \min_{\sigma \in \Omega} \mu(\sigma) \]
Influence matrix and spectral independence

\[ \mu: \text{a distribution over } \Omega \subseteq \{-1, +1\}^V \]

|\(V|\times|V|\) influence matrix \(\Psi \in \mathbb{R}^{V \times V}\) such that

\[
\Psi(u, v) = \left| \Pr_\mu[v = +| u = +] - \Pr_\mu[v = +| u = -] \right|
\]
Influence matrix and spectral independence

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

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

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

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

There is a constant $C > 0$ s.t. for all conditional distribution $\mu_S^\sigma$, 

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

Complete spectral independence

Magnetizing joint distribution with local fields

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

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

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

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

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

$(\phi \ast \mu)$ is spectrally independent with parameter $C$

**Example:** hardcore model $(G, \lambda)$ 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 $\mu$ is $C$-completely spectrally independent, for any $\theta \in (0,1)$

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

$\lambda^*_\text{gap}(\theta \ast \mu)$: minimum spectral gap of Glauber dynamics for all conditional distributions induced by $\theta \ast \mu$. 
Boosting result of spectral gap [This work]

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

\[
\lambda_{\text{gap}}(\mu) \geq \left( \frac{\theta}{2} \right)^{2C+7} \lambda^*_{\text{gap}}(\theta \ast \mu), \quad \theta_v = \theta \text{ for all } v \in V
\]

Near-Critical Regime  Boosting with cost \( O(1) \)  Impose local fields \( \rightarrow \) Easy Regime
Boosting result of spectral gap [This work]

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

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

Application on hardcore model

\[ \lambda \leq (1 - \delta)\lambda_c(\Delta) \]

\[ \theta = 1/25 \]

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

**Dobrushin condition**

- correlation decay [Weitz06, LLY13, ALO20 CLV20]
- $O\left(\frac{1}{\delta}\right)$-completely SI
- path coupling [BD97]
- coupling v.s. spectral gap [Chen98]

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

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

\[ \lambda_{\text{gap}}(\theta \ast \mu) \geq \frac{1}{2n} \]
Proof of boosting result

New Markov chain: *field dynamics*
Field Dynamics

**Input:** a distribution $\mu$ 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 $X_S \sim (\theta \ast \mu)_S(\cdot | X_{V \setminus S})$ conditional distribution induced from $(\theta \ast \mu)$
**Field Dynamics**

**Input**: a distribution $\mu$ 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 $X_S \sim (\theta \ast \mu)_S|_{X_{V \setminus S}}$ conditional distribution induced from $(\theta \ast \mu)$

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

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

$P_{FD}(\theta)$ has the unique stationary distribution $\mu$. 
**Comparison lemma**

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

$$\lambda_{\text{gap}}(\mu) \geq \lambda_{\text{Field}}(\mu, \theta) \cdot \lambda^*_\text{gap}(\theta * \mu), \quad \theta_v = \theta \text{ for all } v \in V$$

**Mixing lemma of field dynamics**

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

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

**Comparison lemma + Mixing lemma**  →  **Boosting result**

$$\lambda_{\text{gap}}(\mu) \geq \left(\frac{\theta}{2}\right)^{2C+7} \lambda^*_\text{gap}(\theta * \mu)$$
Mixing of block dynamics [Chen, Liu and Vigoda 2021]

update of \( \theta \)-fraction block dynamics
- sample \( \theta \) fraction of variables \( R \) u.a.r.
- resample the value of \( R \) conditional on others

\( k \)-transformation [This work]

- \( \mu \) over \( \{-, +\}^V \)
- \( \mu_k \) over \( \{-, +\}^{V_k} \)
- \( V_k = \{u_1, u_2, ..., u_k \mid u \in V\} \)

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

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 \( j \neq i \)

\[ V_9 = \{u \mid u \in V\} \]

3-transformation

\[ X \sim \mu \]

\[ Y \sim \mu_3 \]
Mixing of block dynamics [Chen, Liu and Vigoda 2021]

update of $\theta$-fraction block dynamics

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

for any distribution $\pi$ that is $C$-spectrally independent

$$\lambda_{\text{gap}}^{\text{block}}(\pi) \geq \theta^{O(C)}$$

$k$-transformation [This work]

$$\mu \text{ over } \{-, +\}^V \rightarrow \mu_k \text{ over } \{-, +\}^{V_k}, V_k = \{u_1, u_2, ..., u_k \mid u \in V\}$$

Lemma I field dynamics on $\mu$ is the limit instance of block dynamics on $\mu_k$

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

Lemma II mixing of block dynamics:

$$\mu \text{ is } C\text{-completely-SI} \rightarrow \mu_k \text{ is } (C + 2)\text{-SI for all } k \rightarrow \lambda_{\text{gap}}^{\text{block}}(\mu_k) \geq \theta^{O(C)}$$

Lemma I + Lemma II $\rightarrow$ 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,AJPKV21,CFYZ21]
  - $\tilde{O}(n)$-time sampling algorithm for hardcore model [AJPKV21]
- Extend our technique to **general distributions** beyond the Boolean domain i.e. **q-coloring**