Optimal mixing for two-state anti-ferromagnetic spin systems

Weiming Feng
University of Edinburgh

joint work with

Xiaoyu Chen
(Nanjing University)

Yitong Yin
(Nanjing University)

Xinyuan Zhang
(Nanjing University)

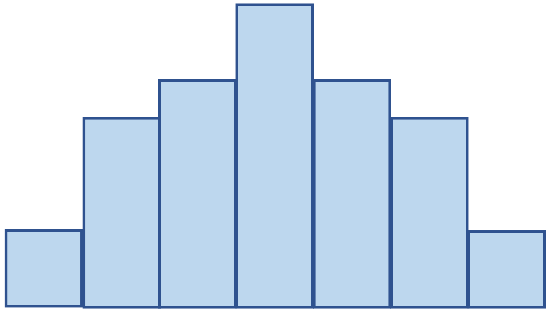
FOCS, Denver, CO, US, 2nd Nov 2022

Sampling, counting and phase transition

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

joint distribution μ:

$$\forall X = (X_v)_{v \in V} \in \{-,+\}^V, \qquad \mu(X) \propto w(X)$$

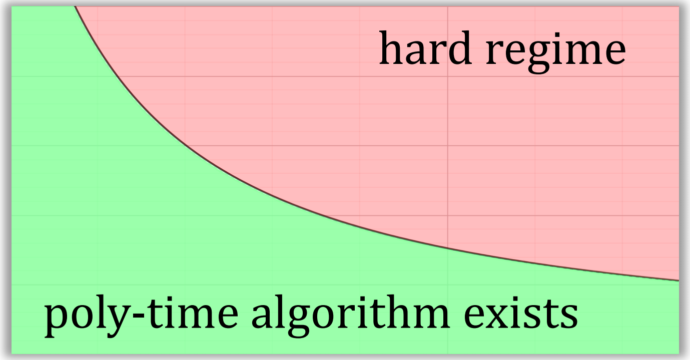


Sampling problem

Draw (approximate) random samples from distribution μ

Computational phase transition

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

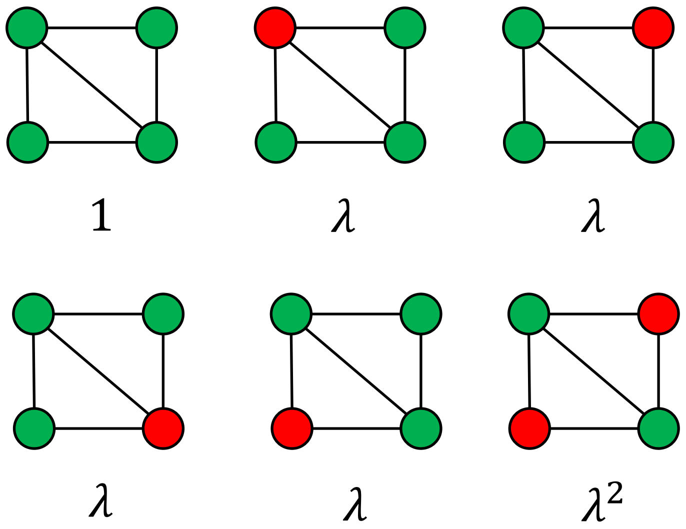


Hardcore gas model

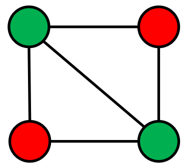
- Graph $G = (V, E)$: n-vertex and max degree Δ;
- Fugacity parameter $\lambda \in \mathbb{R}_{\geq 0}$;
- Configuration $X \in \{-, +\}^V$
  - $X_v = +$: vertex v is **occupied**
  - $X_v = -$: vertex v is **unoccupied**
- $X \in \Omega$ if occupied vertices form an independent set
- Gibbs distribution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

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

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



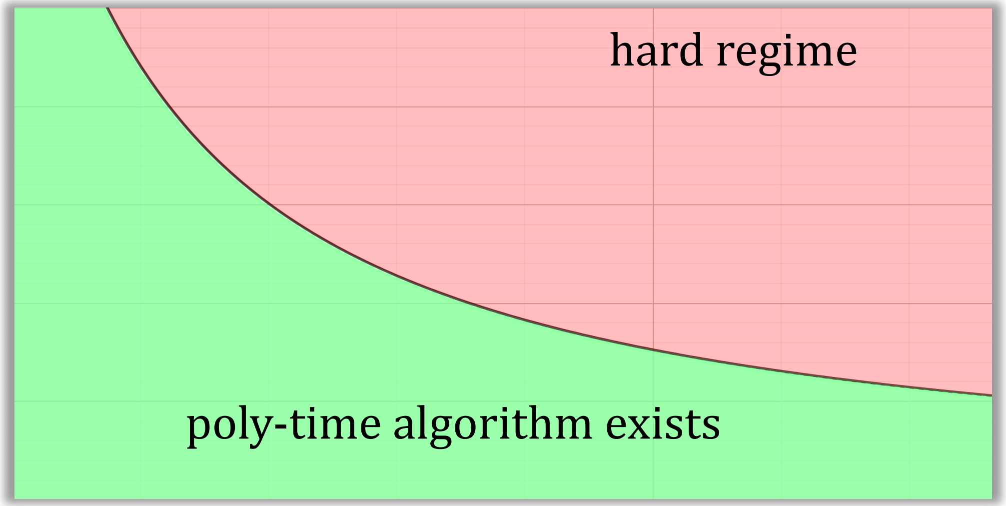
Partition function
$Z = 1 + 4\lambda + \lambda^2$

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

Uniqueness Threshold

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

Δ: maximum degree



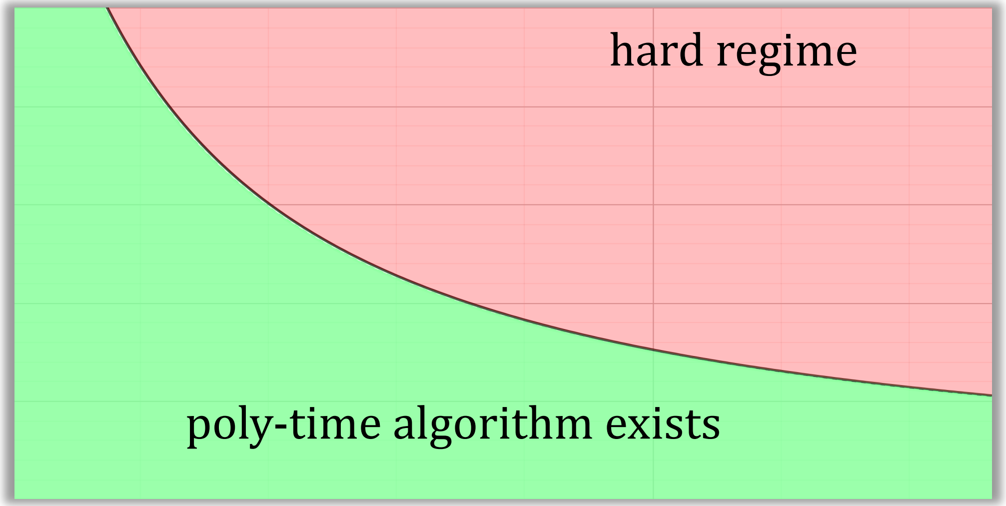
Computational phase transition

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

Uniqueness Threshold

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

Δ: maximum degree



Computational phase transition

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

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

Glauber dynamics for hardcore model

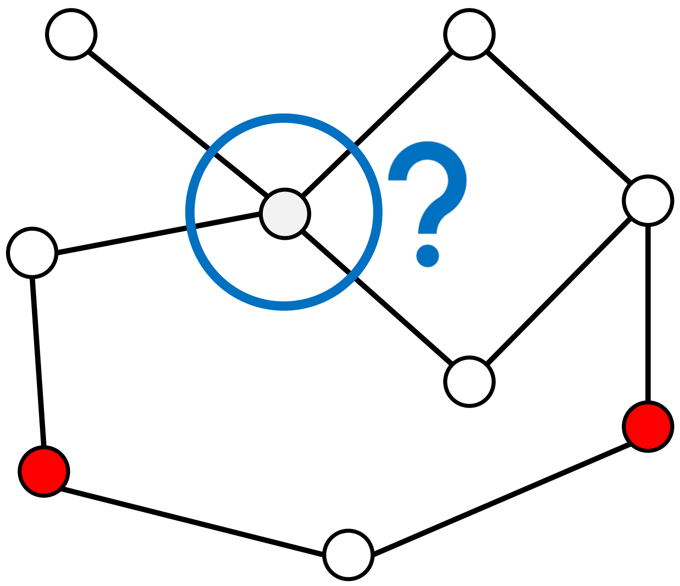
Start from an arbitrary independent set X;

For each transition step **do**

- Lazy w.p. $\frac{1}{2}$, otherwise do as follows:
- 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: $T_{\text{mix}} = \max_{X_0 \in \Omega} \min \left\{ t \mid d_{TV}(X_t, \mu) \le \frac{1}{4e} \right\}$,

$d_{TV}(X_t, \mu)$: 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	$O\left(\frac{1}{\delta} n \log n\right)$

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	$\lambda \leq (1-\delta)\lambda_c(\Delta)$	$\Delta^{O(\Delta^2/\delta)} n \log n$
Chen, F. Yin, Zhang, 2021	$\lambda \leq (1-\delta)\lambda_c(\Delta)$	$e^{O(1/\delta)} n^2 \log n$

Open question: Can we prove the optimal $O(n \log n)$ mixing for all degrees ?

Work	**Mixing Time when $\lambda \leq (1-\delta)\lambda_c(\Delta)$**
Anari, Jain, Koehler, Pham, Vuong, 2021	$e^{O(1/\delta)} n \log n$ Balanced Glauber dynamics

Work	Mixing Time when $\lambda \leq (1-\delta)\lambda_c(\Delta)$
Anari, Jain, Koehler, Pham, Vuong, 2021	$e^{O(1/\delta)} n \log n$ Balanced Glauber dynamics
Chen, F. Yin, Zhang (this work), 2022	$e^{O(1/\delta)} n \log n$
Chen, Eldan (this conference), 2022	$e^{O(1/\delta)} n \log n$

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: $C(\delta)\, n \log n$.

Optimal mixing for two-state anti-ferro spin systems in the uniqueness regime

- Ising model
- general spin systems on regular graphs
- strictly anti-ferro spin systems (both parameters $\beta, \gamma \leq 1$)

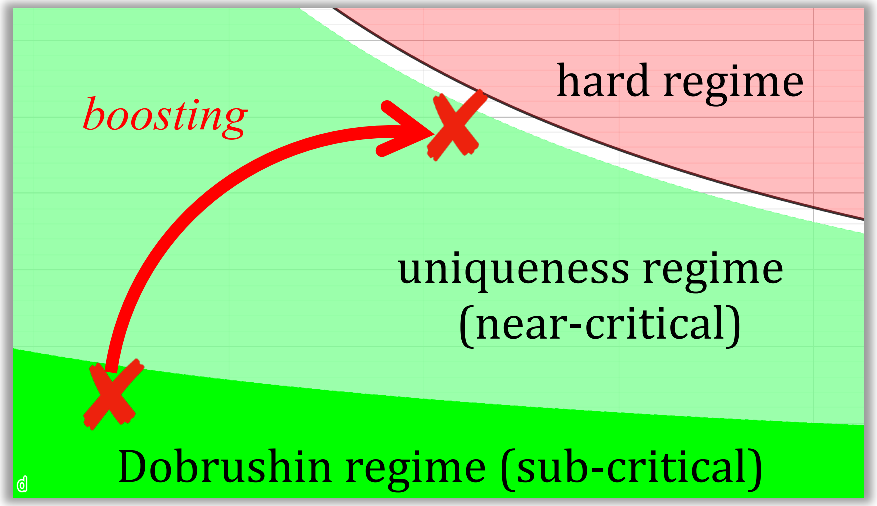
Hardcore model in uniqueness regime

- If λ is ***close*** to $\lambda_c(\Delta)$, e.g., $\lambda = 0.999\lambda_c$ (**near-critical**)

  analyzing mixing time is ***hard***

- If λ is ***far-away*** from $\lambda_c(\Delta)$, e.g., $\lambda \leq 0.1\lambda_c$ (**sub-critical**)

  analyzing mixing time is ***easy***



Results for general joint distributions

A **boosting result** of **modified log-Sobolev constant**

for distributions satisfying

complete spectral independence and complete marginal stability

Anti-ferro 2-spin systems: guaranteed by the ***uniqueness condition***

Modified log-Sobolev contant

μ: a joint distribution over $\Omega \subseteq \{-,+\}^n$, e.g., the Gibbs distribution of hardcore model

P: transition matrix of (lazy) Glauber dynamics on μ

Modified log-Sobolev (MLS) constant of Glauber dynamics : $\rho = \rho(P, \mu)$ such that

$$T_{\text{mix}} \le O\left(\frac{1}{\rho}\left(\log\log\frac{1}{\mu_{\min}}\right)\right) = O\left(\frac{\log n}{\rho}\right), \quad \mu_{\min} = \min_{\sigma\in\{-,+\}^V} \mu(\sigma) = \frac{1}{2^{O(n)}}$$

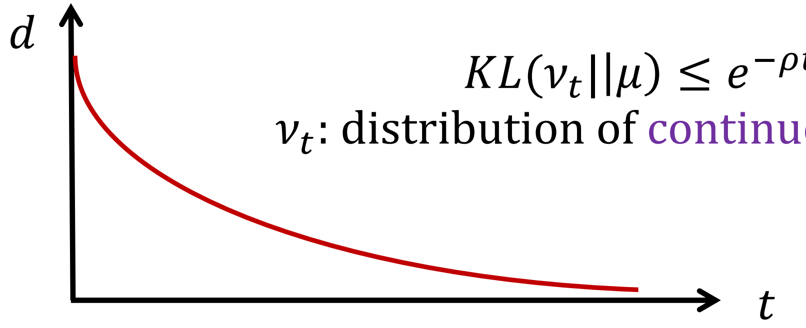
$\rho = \Omega(1/n)$ → $T_{\text{mix}} = O(n \log n)$

$$\rho = \inf\left\{\frac{\mathcal{E}_P(f, \log f)}{\text{Ent}_\mu[f]} \mid f: \Omega \to \mathbb{R}_{>0}, \text{Ent}_\mu[f] > 0\right\}$$

$$\mathcal{E}_P(f, \log f) = \sum_{x,y\in\Omega} \mu(x) P(x,y)\,(f_x - f_y)(\log f_x - \log f_y)$$

$$\text{Ent}_\mu[f] = \sum_x \mu(x) f_x \log f_x - \sum_x \mu(x) f_x \log \sum_y \mu(y) f_y$$

formal definition MLS constant



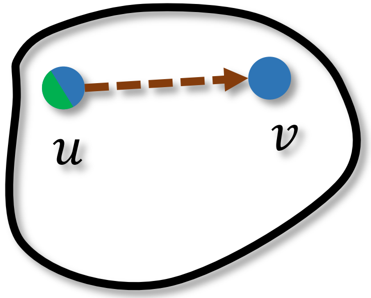
$KL(\nu_t || \mu) \le e^{-\rho t} KL(\nu_0 || \mu)$

ν_t: distribution of continuous-time at time t

t: time in **continuous-time** Glauber dynamics

d: **KL-divergence** between current and stationary distribution

Influence matrix and spectral independence



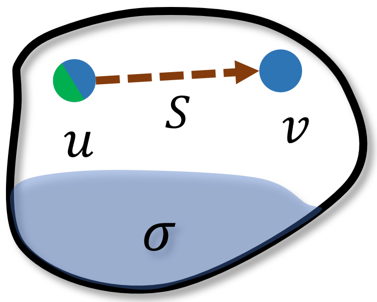
influence on v caused by a **disagreement** on u

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,

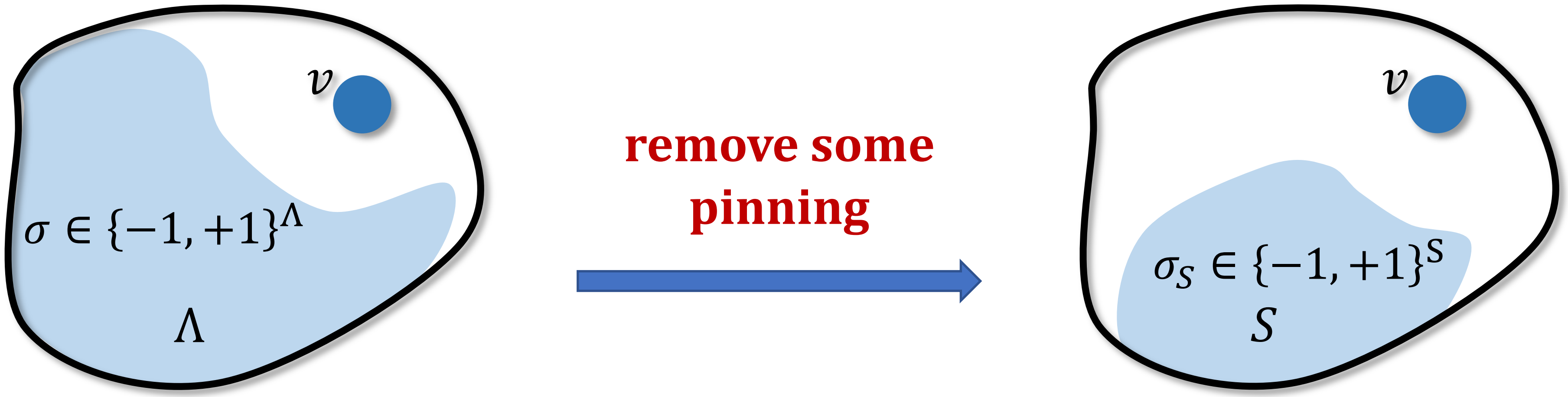
spectral radius of influence matrices $\rho(\Psi_S^\sigma) \le C$.

Marginal stability [This work]

For any pinning $\sigma \in \{-,+\}^{\Lambda}$ and $v \notin \Lambda$, let

marginal ratio $$R_v^{\sigma} = \frac{\mu_v^{\sigma}(+)}{\mu_v^{\sigma}(-)},$$

- $R_v^{\sigma} \leq \zeta$
- $R_v^{\sigma} \leq \zeta R_v^{\sigma_S}$ for any $S \subseteq \Lambda$



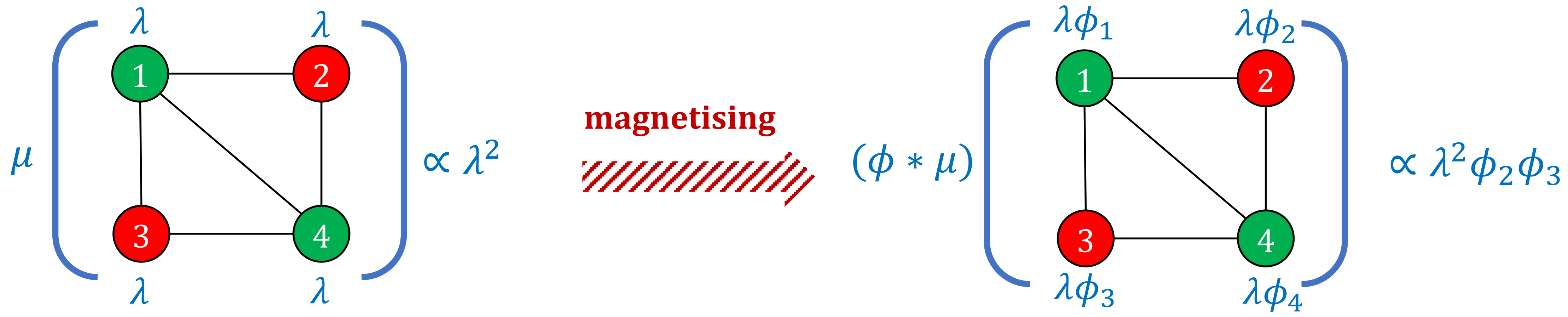
marginal ratio R_v^{σ} marginal ratio $R_v^{\sigma_S}$

Distribution with local fields

Magnetising 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

There is constants $C > 0$ and $\epsilon > 0$ s.t.

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

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

Complete marginal stability

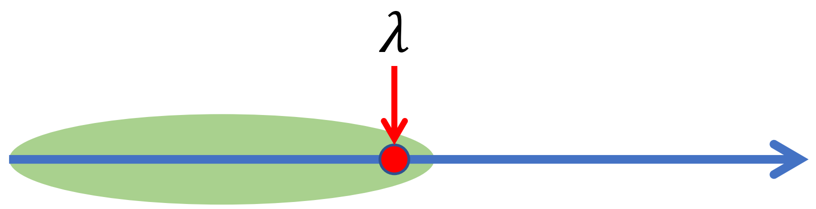
There is constants $\zeta > 0$ and $\epsilon > 0$ s.t.

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

$(\boldsymbol{\phi} * \mu)$ is *marginally stable* with parameter ζ

Example: hardcore model (G, λ):

any hardcore models $(G, (\lambda_v)_{v \in V})$ with $\lambda_v \leq (1+\epsilon)\lambda$ are ***spectrally independent*** ***and*** ***marginally stable***



spectral independence & marginal stability for ***all*** subcritical local fields

Boosting result for modified log-Sobolev constant [This work]

If μ is **completely spectrally independent** with parameter $C, \epsilon > 0$

and **completely marginally stable** with parameter $\zeta > 0$

then for any $\theta \in (0,1)$

$$\rho_{\text{mls}}^{\text{GD}}(\mu) \geq f(\theta, C, \epsilon, \zeta) \cdot \rho_{\text{minmls}}^{\text{GD}}(\boldsymbol{\theta} * \mu), \qquad \boldsymbol{\theta}_v = \theta \text{ for all } v \in V$$

$\rho_{\text{minmls}}^{\text{GD}}(\boldsymbol{\theta} * \mu)$: *minimum MLS constant of Glauber dynamics for all conditional distributions induced by* $\boldsymbol{\theta} * \mu$.

Boosting result for modified log-Sobolev constant [This work]

If μ is **completely spectrally independent** with parameter $C, \epsilon > 0$

and **completely marginally stable** with parameter $\zeta > 0$

then for any $\theta \in (0,1)$

$$\rho_{\text{mls}}^{\text{GD}}(\mu) \geq f(\theta, C, \epsilon, \zeta) \cdot \rho_{\text{minmls}}^{\text{GD}}(\boldsymbol{\theta} * \mu), \qquad \boldsymbol{\theta}_v = \theta \text{ for all } v \in V$$

Boosting modified log-Sobolev constant
with cost $\Theta(1)$

$$\rho_{\text{minmls}}^{\text{GD}}(\boldsymbol{\theta} * \mu) = \Omega\left(\frac{1}{n}\right)$$

⇒ **optimal** mixing time bound

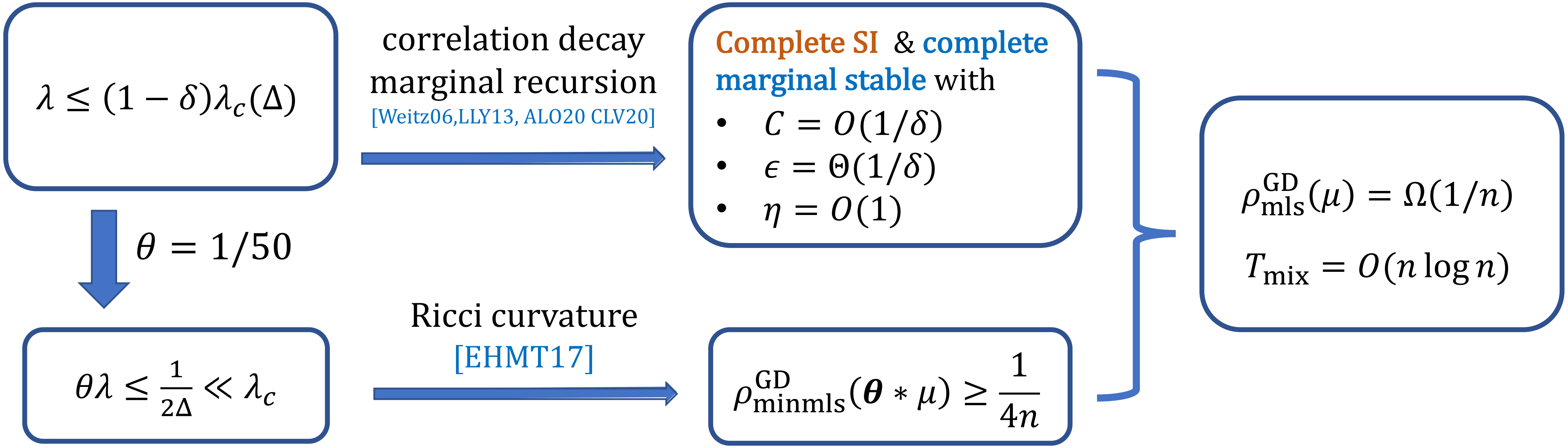
Boosting result for modified log-Sobolev constant [This work]

If μ is **completely spectrally independent** with parameter $C, \epsilon > 0$

and **completely marginally stable** with parameter $\zeta > 0$

then for any $\theta \in (0,1)$

$$\rho_{\text{mls}}^{\text{GD}}(\mu) \geq f(\theta, C, \epsilon, \zeta) \cdot \rho_{\text{minmls}}^{\text{GD}}(\boldsymbol{\theta} * \mu), \qquad \boldsymbol{\theta}_v = \theta \text{ for all } v \in V$$

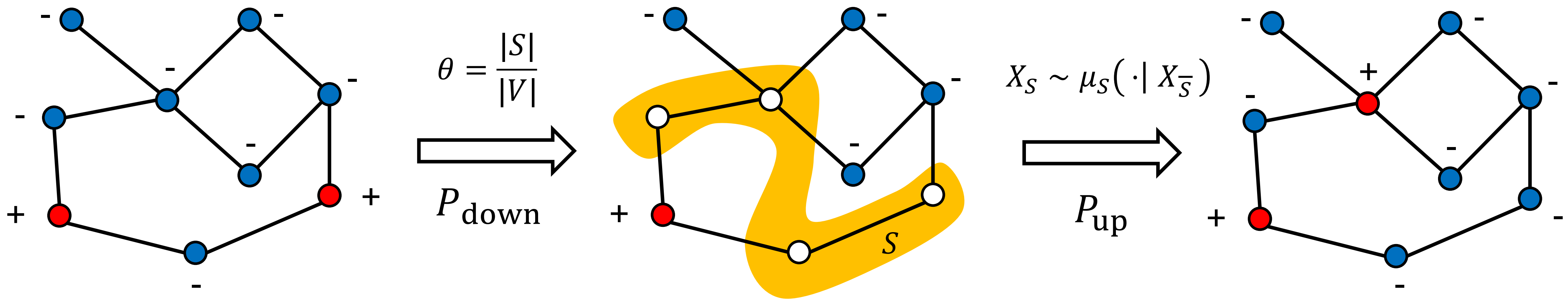


Proof Overview

θ-down up walk on μ

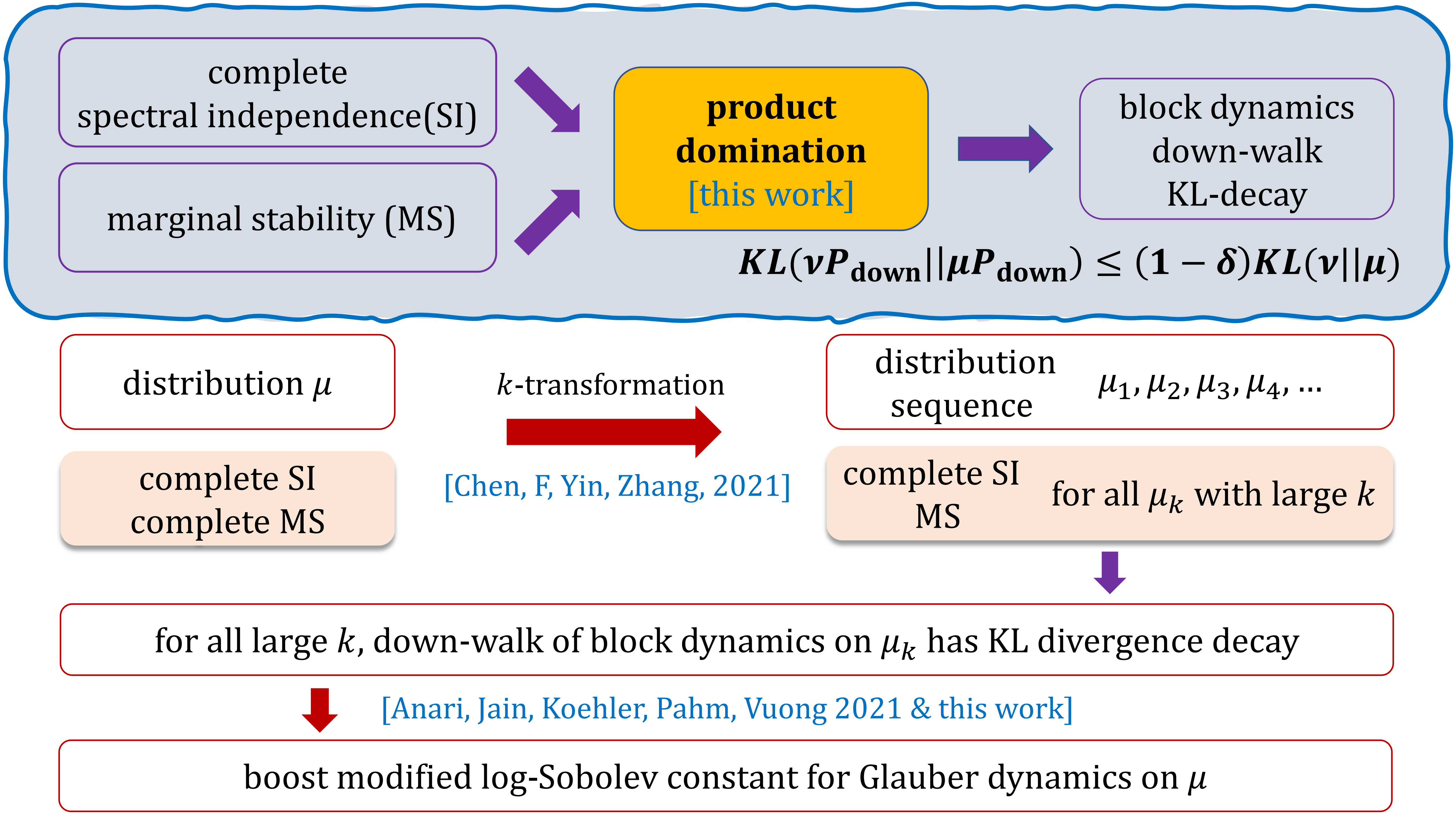
Transition step: given configuration $X \in \{-,+\}^V$

- pick θ fraction of variables $S \subseteq V$ uniformly at random — **down walk P_{down}**
- resample $X_S \sim \mu_S(\cdot \mid X_{\overline{S}})$ — **up walk P_{up}**



Glauber dynamics on μ: $\theta = \frac{1}{n}$

Block dynamics: $\theta = \Theta(1)$


complete spectral independence(SI)
marginal stability (MS)
product domination
[this work]
block dynamics down-walk KL-decay
$KL(\nu P_{\text{down}} || \mu P_{\text{down}}) \le (1-\delta) KL(\nu || \mu)$
distribution μ
k-transformation
distribution sequence
$\mu_1, \mu_2, \mu_3, \mu_4, \ldots$
complete SI
complete MS
[Chen, F, Yin, Zhang, 2021]
complete SI
MS
for all μ_k with large k
for all large k, down-walk of block dynamics on μ_k has KL divergence decay
[Anari, Jain, Koehler, Pahm, Vuong 2021 & this work]
boost modified log-Sobolev constant for Glauber dynamics on μ

μ: distribution over $\{-,+\}^{[n]}$; probability generating function (PGF):

$$g_\mu(z_1, z_2, \ldots, z_n) = \sum_{X \in \{-,+\}^{[n]}} \mu(X) \prod_{i \in [n]: X(i)=+} z_i$$

Product domination (PD): there exists a constant $0 < \alpha < 1$ such that

$$\forall (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n_{>0}, \qquad g_\mu(z_1^\alpha, z_2^\alpha, \ldots, z_n^\alpha)^{\frac{1}{\alpha}} \le \prod_{i=1}^n (\mu_i(+1) z_i + \mu_i(-1))$$

α-fractional PGF

PGF of a **product distribution**, $X_i \sim \mu_i$ for each $i \in [n]$

product domination

$\forall$ conditional distributions

Entropic independence

[Anari, Jain, Koehler, Pahm, Vuong 2021]

block factorisation of entropy

[Caputo,Parisi, 2020]

block dynamics

down-walk

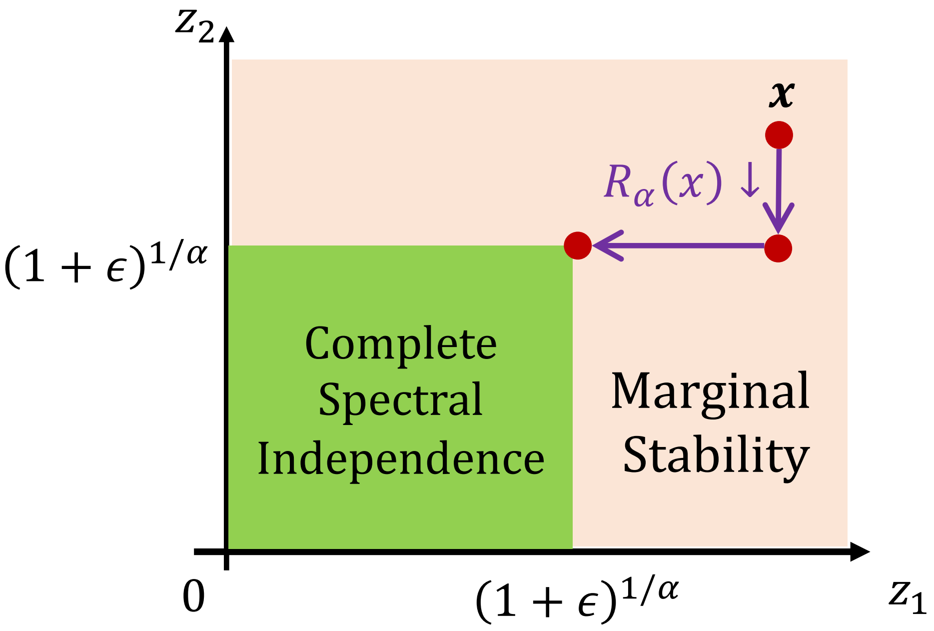
KL-decay

complete spectral independence(SI)

marginal stability (MS)

product domination
[this work]

$$\forall \boldsymbol{z} > 0\,, R_\alpha(\boldsymbol{z}) = \frac{g_\mu(z_1^\alpha, z_2^\alpha, \ldots, z_n^\alpha)^{\frac{1}{\alpha}}}{\prod_{i\in[n]}\left(\mu_i(+)z_i + \mu_i(-)\right)} \leq 1$$



- ***Complete SI***: $\forall \boldsymbol{x} > 0$ with $|\boldsymbol{x}|_\infty \leq (1+\epsilon)^{1/\alpha}$
$$R_\alpha(\boldsymbol{x}) \leq 1$$
- **MS:** $\forall \boldsymbol{x} > 0, \forall i \in [n]$ with $x_i \geq (1+\epsilon)^{1/\alpha}$
$$\left.\frac{\partial R_\alpha}{\partial z_i}\right|_{\boldsymbol{z}=\boldsymbol{x}} \leq 0$$
⇒ $R_\alpha(\boldsymbol{x}) \leq 1$ for all $x > 0$ with $|\boldsymbol{x}|_\infty > (1+\epsilon)^{1/\alpha}$
- **Complete SI & MS** ⇒ **product domination**

Summary

- ***Optimal*** $O(n \log n)$ ***mixing time*** for Glauber dynamics on
  - hardcore / anti-ferro Ising model in the uniqueness regime
  - some general anti-ferro 2-spin systems in the uniqueness regime
- ***Boosting modified log-Sobolev constant*** for distributions satisfying
  - complete spectral independence
  - complete marginal stability
- Technique: ***product domination***

Open problems

- Optimal $O(n \log n)$ mixing time for ***all*** 2-spin systems in the uniqueness regime
  - potential way: ***other sufficient condition*** for product domination?
- Beyond the ***Boolean*** distributions
- More applications of product domination