Rapid mixing of Gibbs samplers: Coupling, Spectral Independence, and Entropy factorization

Pietro Caputo

March 18, 2021

Based on joint works with Antonio Blanca, Zonchen Chen, Daniel Parisi, Alistair Sinclair, Daniel Stefankovic, Eric Vigoda

Plan of the talk

- Spin systems on a graph G
- A general class of Gibbs samplers (heat bath dynamics)
- Entropy Factorization and Mixing time bounds
- Spectral Independence
- Spectral Independence $\Rightarrow$ Entropy Factorization
- Contractive coupling $\Rightarrow$ Spectral Independence

[CP20] PC, D. Parisi, *Block factorization of the relative entropy via spatial mixing* arXiv:2004.10574

[BCPSV20] A. Blanca, PC, D. Parisi, A. Sinclair, E. Vigoda,
Entropy decay in the Swendsen-Wang dynamics arXiv:2007.06931

[BCCPSV21] A. Blanca, PC, Z. Chen, D. Parisi, D. Stefankovic, E. Vigoda,
On Mixing of Markov Chains: Coupling, Spectral Independence, and Entropy Factorization arXiv:2103.07459

[ALO20] N. Anari, K. Liu, and S. Oveis Gharan.
Spectral Independence in High-Dimensional Expanders and Applications to the Hardcore Model, arXiv:2001.00303

[CLV20] Z. Chen, K. Liu, and E. Vigoda.
Optimal Mixing of Glauber Dynamics: Entropy Factorization via High-Dimensional Expansion, arXiv:2011.02075

Spin systems on a graph G

$G = (V, E)$ is a finite graph with maximum degree Δ.

A *spin system* on G is a Gibbs measure μ on $\Omega = [q]^V$, $[q] = \{1, \ldots, q\}$ for some $q \in \mathbb{N}$, associated with some interaction along the edges of G.

Spin systems on a graph G

$G = (V, E)$ is a finite graph with maximum degree Δ.

A *spin system* on G is a Gibbs measure μ on $\Omega = [q]^V$, $[q] = \{1, \ldots, q\}$ for some $q \in \mathbb{N}$, associated with some interaction along the edges of G.

Some examples of spin systems on G:

Potts Model: $\mu(\sigma) = \frac{\exp(\beta M(\sigma))}{Z(G,\beta)}$, $\quad M(\sigma) = \sum_{xy \in E} \mathbf{1}(\sigma_x = \sigma_y)$

Here $q \geqslant 2$. When $q = 2$ it is known as the **Ising Model**.
When $\beta \geqslant 0$ the Potts model is called ferromagnetic.

Spin systems on a graph G

$G = (V, E)$ is a finite graph with maximum degree Δ.

A *spin system* on G is a Gibbs measure μ on $\Omega = [q]^V$, $[q] = \{1, \ldots, q\}$ for some $q \in \mathbb{N}$, associated with some interaction along the edges of G.

Some examples of spin systems on G:

Potts Model: $\mu(\sigma) = \frac{\exp(\beta M(\sigma))}{Z(G,\beta)}$, $\quad M(\sigma) = \sum_{xy \in E} \mathbf{1}(\sigma_x = \sigma_y)$

Here $q \geqslant 2$. When $q = 2$ it is known as the **Ising Model**. When $\beta \geqslant 0$ the Potts model is called ferromagnetic.

q-Colorings: $\mu(\sigma) = \frac{\mathbf{1}(\sigma \in \Omega_{G,q})}{|\Omega_{G,q}|}$, $\Omega_{G,q} = \{\text{proper } q\text{-colorings of } G\}$

Spin systems on a graph G

$G = (V, E)$ is a finite graph with maximum degree Δ.

A *spin system* on G is a Gibbs measure μ on $\Omega = [q]^V$, $[q] = \{1, \ldots, q\}$ for some $q \in \mathbb{N}$, associated with some interaction along the edges of G.

Some examples of spin systems on G:

Potts Model: $\mu(\sigma) = \frac{\exp(\beta M(\sigma))}{Z(G,\beta)}$, $\quad M(\sigma) = \sum_{xy \in E} \mathbf{1}(\sigma_x = \sigma_y)$

Here $q \geqslant 2$. When $q = 2$ it is known as the **Ising Model**. When $\beta \geqslant 0$ the Potts model is called ferromagnetic.

q-Colorings: $\mu(\sigma) = \frac{\mathbf{1}(\sigma \in \Omega_{G,q})}{|\Omega_{G,q}|}$, $\Omega_{G,q} = \{\text{proper } q\text{-colorings of } G\}$

Hard-core gas: $\mu(\sigma) = \frac{\lambda^{|\sigma|}\mathbf{1}(\sigma \in \mathcal{I})}{Z(G,\lambda)}$, $\quad \mathcal{I} = \{\text{independent sets of } G\}$

($\lambda > 0$, $q = 2$, $\sigma_x = 2$ if x is *empty* and $\sigma_x = 1$ if x is *occupied*).

Gibbs samplers

Notation: μ_Λ^τ is the conditional distribution $\mu(\cdot|\sigma_{\Lambda^c} = \tau)$, $\Lambda \subset V$ and τ is called a boundary condition or a pinning. For $f : \Omega \mapsto \mathbb{R}$ we write $\mu_\Lambda f$ for the conditional expectation

$$\mu_\Lambda f : \Omega \mapsto \mathbb{R}\,, \qquad \mu_\Lambda f(\sigma) := \mu_\Lambda^{\sigma_{\Lambda^c}}[f]$$

Gibbs samplers

Notation: μ_Λ^τ is the conditional distribution $\mu(\cdot|\sigma_{\Lambda^c} = \tau)$, $\Lambda \subset V$ and τ is called a boundary condition or a pinning. For $f : \Omega \mapsto \mathbb{R}$ we write $\mu_\Lambda f$ for the conditional expectation

$$\mu_\Lambda f : \Omega \mapsto \mathbb{R}\,, \qquad \mu_\Lambda f(\sigma) := \mu_\Lambda^{\sigma_{\Lambda^c}}[f]$$

Let $\alpha = \{\alpha_\Lambda, \Lambda \subset V\}$ be a probability over 2^V and consider the Markov chain where at each step a subset $\Lambda \subset V$ is picked according to α and its spins σ_Λ are updated according to $\mu_\Lambda^{\sigma_{\Lambda^c}}$. This chain has transition operator

$$P_\alpha f = \sum_{\Lambda \subset V} \alpha_\Lambda \mu_\Lambda f$$

and will be called the α-weighted block dynamics.

Gibbs samplers

Notation: μ_Λ^τ is the conditional distribution $\mu(\cdot|\sigma_{\Lambda^c} = \tau)$, $\Lambda \subset V$ and τ is called a boundary condition or a pinning. For $f : \Omega \mapsto \mathbb{R}$ we write $\mu_\Lambda f$ for the conditional expectation

$$\mu_\Lambda f : \Omega \mapsto \mathbb{R}\,, \quad \mu_\Lambda f(\sigma) := \mu_\Lambda^{\sigma_{\Lambda^c}}[f]$$

Let $\alpha = \{\alpha_\Lambda, \Lambda \subset V\}$ be a probability over 2^V and consider the Markov chain where at each step a subset $\Lambda \subset V$ is picked according to α and its spins σ_Λ are updated according to $\mu_\Lambda^{\sigma_{\Lambda^c}}$. This chain has transition operator

$$P_\alpha f = \sum_{\Lambda \subset V} \alpha_\Lambda \mu_\Lambda f$$

and will be called the α-weighted block dynamics.
Note that choosing $\alpha_\Lambda = \frac{1}{|V|}\mathbf{1}(|\Lambda| = 1)$, gives the usual Glauber dynamics. In general, P_α is reversible and has Dirichlet form:

$$\mathcal{D}_\alpha(f, g) = \langle f, (1 - P_\alpha)g\rangle = \sum_{\Lambda \subset V} \alpha_\Lambda \mu\left[\mathrm{Cov}_\Lambda(f, g)\right]$$

Gibbs samplers

Notation: μ_Λ^τ is the conditional distribution $\mu(\cdot|\sigma_{\Lambda^c} = \tau)$, $\Lambda \subset V$ and τ is called a boundary condition or a pinning. For $f : \Omega \mapsto \mathbb{R}$ we write $\mu_\Lambda f$ for the conditional expectation

$$\mu_\Lambda f : \Omega \mapsto \mathbb{R}\,, \qquad \mu_\Lambda f(\sigma) := \mu_\Lambda^{\sigma_{\Lambda^c}}[f]$$

Let $\alpha = \{\alpha_\Lambda,\, \Lambda \subset V\}$ be a probability over 2^V and consider the Markov chain where at each step a subset $\Lambda \subset V$ is picked according to α and its spins σ_Λ are updated according to $\mu_\Lambda^{\sigma_{\Lambda^c}}$. This chain has transition operator

$$P_\alpha f = \sum_{\Lambda \subset V} \alpha_\Lambda \mu_\Lambda f$$

and will be called the α-weighted block dynamics.
Note that choosing $\alpha_\Lambda = \frac{1}{|V|}\mathbf{1}(|\Lambda| = 1)$, gives the usual Glauber dynamics. In general, P_α is reversible and has Dirichlet form:

$$\mathcal{D}_\alpha(f,g) = \langle f, (1 - P_\alpha)g\rangle = \sum_{\Lambda \subset V} \alpha_\Lambda \mu\left[\mathrm{Cov}_\Lambda(f,g)\right]$$

where $\mathrm{Cov}_\Lambda(f,g) = \mu_\Lambda((f - \mu_\Lambda f)(g - \mu_\Lambda g))$.

Mixing Time and Entropy

For a Markov chain with transition P and stationary distribution μ:

$$T_{\mathrm{mix}}(P) := \inf\left\{ t \in \mathbb{N} : \max_{\sigma\in\Omega} \|P^t(\sigma,\cdot) - \mu\|_{TV} \leqslant 1/4 \right\} .$$

The entropy of $f : \Omega \mapsto \mathbb{R}_+$ w.r.t. μ or μ_Λ is defined by

$$\mathrm{Ent}(f) = \mu\left[f \log(f/\mu[f])\right] , \quad \mathrm{Ent}_\Lambda(f) = \mu_\Lambda\left[f \log(f/\mu_\Lambda[f])\right].$$

Recall that $\mathrm{Ent}(f) = H(f\mu|\mu)$ when $\mu[f] = 1$.

Mixing Time and Entropy

For a Markov chain with transition P and stationary distribution μ:

$$T_{\mathrm{mix}}(P) := \inf\left\{ t \in \mathbb{N} : \max_{\sigma\in\Omega} \|P^t(\sigma,\cdot) - \mu\|_{TV} \leqslant 1/4 \right\} .$$

The entropy of $f : \Omega \mapsto \mathbb{R}_+$ w.r.t. μ or μ_Λ is defined by

$$\mathrm{Ent}(f) = \mu\left[f \log(f/\mu[f])\right] , \quad \mathrm{Ent}_\Lambda(f) = \mu_\Lambda\left[f \log(f/\mu_\Lambda[f])\right] .$$

Recall that $\mathrm{Ent}(f) = H(f\mu|\mu)$ when $\mu[f] = 1$.

Pinsker's inequality $\|\nu - \mu\|_{TV}^2 \leqslant \frac{1}{2} H(\nu|\mu)$ shows that

$$\mathrm{Ent}(Pf) \leqslant (1-\delta)\mathrm{Ent}(f) \quad \Rightarrow \quad T_{\mathrm{mix}}(P) \leqslant 4\,\delta^{-1} \log\log(1/\mu_*) ,$$

where $\mu_* = \min_\sigma \mu(\sigma)$.

Mixing Time and Entropy

For a Markov chain with transition P and stationary distribution μ:

$$T_{\mathrm{mix}}(P) := \inf\left\{ t \in \mathbb{N} : \max_{\sigma\in\Omega} \|P^t(\sigma,\cdot) - \mu\|_{TV} \leqslant 1/4 \right\} .$$

The entropy of $f : \Omega \mapsto \mathbb{R}_+$ w.r.t. μ or μ_Λ is defined by

$$\mathrm{Ent}(f) = \mu\left[f \log(f/\mu[f])\right] , \quad \mathrm{Ent}_\Lambda(f) = \mu_\Lambda\left[f \log(f/\mu_\Lambda[f])\right].$$

Recall that $\mathrm{Ent}(f) = H(f\mu|\mu)$ when $\mu[f] = 1$.

Pinsker's inequality $\|\nu - \mu\|_{TV}^2 \leqslant \frac{1}{2} H(\nu|\mu)$ shows that

$$\mathrm{Ent}(Pf) \leqslant (1-\delta)\mathrm{Ent}(f) \;\Rightarrow\; T_{\mathrm{mix}}(P) \leqslant 4\,\delta^{-1} \log\log(1/\mu_*) ,$$

where $\mu_* = \min_\sigma \mu(\sigma)$. The entropy contraction above is a discrete time analog of the *Modified log-Sobolev inequality*

$$\mathcal{D}(f, \log f) \geqslant \delta\, \mathrm{Ent}(f).$$

Mixing Time and Entropy

For a Markov chain with transition P and stationary distribution μ:

$$T_{\rm mix}(P) := \inf\left\{ t \in \mathbb{N} : \max_{\sigma\in\Omega} \|P^t(\sigma,\cdot) - \mu\|_{TV} \leqslant 1/4 \right\} .$$

The entropy of $f : \Omega \mapsto \mathbb{R}_+$ w.r.t. μ or μ_Λ is defined by

$$\mathrm{Ent}(f) = \mu\left[f\log(f/\mu[f])\right] , \quad \mathrm{Ent}_\Lambda(f) = \mu_\Lambda\left[f\log(f/\mu_\Lambda[f])\right].$$

Recall that $\mathrm{Ent}(f) = H(f\mu|\mu)$ when $\mu[f] = 1$.

Pinsker's inequality $\|\nu - \mu\|_{TV}^2 \leqslant \frac{1}{2}H(\nu|\mu)$ shows that

$$\mathrm{Ent}(Pf) \leqslant (1-\delta)\mathrm{Ent}(f) \;\Rightarrow\; T_{\rm mix}(P) \leqslant 4\,\delta^{-1}\log\log(1/\mu_*),$$

where $\mu_* = \min_\sigma \mu(\sigma)$. The entropy contraction above is a discrete time analog of the *Modified log-Sobolev inequality*

$$\mathcal{D}(f, \log f) \geqslant \delta\,\mathrm{Ent}(f).$$

All our upper bounds on the mixing time will follow from the entropy contraction. Note: $\log\log(1/\mu_*) = O(\log n)$, $n = |V|$.

Entropy factorization

We say that μ satisfies the Block Factorization (BF) of entropy with constant C if for all $f : \Omega \mapsto \mathbb{R}_+$, and for all weights α,

$$\gamma(\alpha)\,\mathrm{Ent} f \;\leqslant\; C \sum_{\Lambda \subset V} \alpha_\Lambda \, \mu[\mathrm{Ent}_\Lambda f],$$

where $\gamma(\alpha) = \min_{x \in V} \sum_{\Lambda \ni x} \alpha_\Lambda$.

Entropy factorization

We say that μ satisfies the Block Factorization (BF) of entropy with constant C if for all $f : \Omega \mapsto \mathbb{R}_+$, and for all weights α,

$$\gamma(\alpha)\,\mathrm{Ent} f \;\leqslant\; C \sum_{\Lambda\subset V} \alpha_\Lambda\, \mu[\mathrm{Ent}_\Lambda f],$$

where $\gamma(\alpha) = \min_{x\in V} \sum_{\Lambda\ni x} \alpha_\Lambda$. Some remarks:
1) if μ is a product measure then BF holds with $C = 1$ (follows from Shearer inequality for Shannon entropy).

Entropy factorization

We say that μ satisfies the Block Factorization (BF) of entropy with constant C if for all $f : \Omega \mapsto \mathbb{R}_+$, and for all weights α,

$$\gamma(\alpha)\,\mathrm{Ent} f \;\leqslant\; C \sum_{\Lambda \subset V} \alpha_\Lambda\, \mu[\mathrm{Ent}_\Lambda f],$$

where $\gamma(\alpha) = \min_{x\in V} \sum_{\Lambda \ni x} \alpha_\Lambda$. Some remarks:
1) if μ is a product measure then BF holds with $C = 1$ (follows from Shearer inequality for Shannon entropy).
2) if $\alpha_\Lambda = \frac{1}{|V|}\mathbf{1}(|\Lambda| = 1)$: Approximate Tensorization (AT)

$$\mathrm{Ent} f \;\leqslant\; C \sum_{x \in V} \mu[\mathrm{Ent}_x f].$$

Equivalent to log-Sobolev inequality for Glauber dynamics.
If $G \subset \mathbb{Z}^d$: known under Strong Spatial Mixing (SSM) assumption from Stroock-Zegarlinski '92; Martinelli, Olivieri '94; Cesi '01. For generic graph G AT is known for small enough $|\beta|$: C,Menz,Tetali '14; Marton '14; Bauerschmidt, Bodineau '19, or under negative dependence assumptions: Cryan,Guo,Mousa '19; Hermon,Salez '19.

Consequence of BF for mixing times

Lemma

If BF holds with constant C, then for all weights α,

$$\mathrm{Ent}(P_\alpha f) \leqslant (1-\delta)\mathrm{Ent}(f)\,, \qquad \delta = \gamma(\alpha)/C.$$

In particular, $T_{\mathrm{mix}}(P) = O(\gamma(\alpha)^{-1}\log n)$.

Consequence of BF for mixing times

Lemma

If BF holds with constant C, then for all weights α,

$$\mathrm{Ent}(P_\alpha f) \leqslant (1-\delta)\mathrm{Ent}(f), \qquad \delta = \gamma(\alpha)/C.$$

In particular, $T_{\mathrm{mix}}(P) = O(\gamma(\alpha)^{-1}\log n)$.

Indeed, BF means that

$$\textstyle\sum_\Lambda \alpha_\Lambda \mu[\mathrm{Ent}_\Lambda(f)] \geqslant \frac{\gamma(\alpha)}{C}\,\mathrm{Ent}(f).$$

By convexity of $\mathrm{Ent}(\cdot)$:

$$\begin{aligned}\mathrm{Ent}(P_\alpha f) &\leqslant \textstyle\sum_\Lambda \alpha_\Lambda\, \mu[\mathrm{Ent}(\mu_\Lambda(f))]\\ &= \mathrm{Ent}(f) - \textstyle\sum_\Lambda \alpha_\Lambda \mu[\mathrm{Ent}_\Lambda(f)] \leqslant (1-\delta)\mathrm{Ent}(f).\end{aligned}$$

Note: the mixing time bound is tight up to $O(\log n)$ since the spectral gap always satisfies $\lambda(P_\alpha) \geqslant \gamma(\alpha)$. Often optimal mixing

.

Entropy factorization for $G \subset \mathbb{Z}^d$

Theorem (CP20)

For $G \subset \mathbb{Z}^d$, under SSM , the general BF holds with a constant C independent of n and the boundary conditions.

Entropy factorization for $G \subset \mathbb{Z}^d$

Theorem (CP20)

For $G \subset \mathbb{Z}^d$, under SSM , the general BF holds with a constant C independent of n and the boundary conditions.

1. Reduce to proving a bipartite factorization into even/odd vertices $\alpha_E = \alpha_O = 1/2$.
2. Use suitable recursive strategy to prove it for even/odd case (main difficulty: lack of a simple additive structure).

Entropy factorization for $G \subset \mathbb{Z}^d$

Theorem (CP20)

For $G \subset \mathbb{Z}^d$ *, under SSM , the general BF holds with a constant* C *independent of* n *and the boundary conditions.*

1. Reduce to proving a bipartite factorization into even/odd vertices $\alpha_E = \alpha_O = 1/2$.
2. Use suitable recursive strategy to prove it for even/odd case (main difficulty: lack of a simple additive structure).

Theorem (BCPSV20)

For $G \subset \mathbb{Z}^d$*, under SSM, the Swendsen-Wang dynamics for ferromagnetic Ising/Potts models has* $T_{\mathrm{mix}}(P_{\mathrm{SW}}) = \Theta(\log n)$

Entropy factorization for $G \subset \mathbb{Z}^d$

Theorem (CP20)

For $G \subset \mathbb{Z}^d$, under SSM , the general BF holds with a constant C independent of n and the boundary conditions.

1. Reduce to proving a bipartite factorization into even/odd vertices $\alpha_E = \alpha_O = 1/2$.
2. Use suitable recursive strategy to prove it for even/odd case (main difficulty: lack of a simple additive structure).

Theorem (BCPSV20)

For $G \subset \mathbb{Z}^d$, under SSM, the Swendsen-Wang dynamics for ferromagnetic Ising/Potts models has $T_{\mathrm{mix}}(P_{\mathrm{SW}}) = \Theta(\log n)$

1. Reduce to spin/edge factorization for Edwards-Sokal coupling ν:

$$\mathrm{Ent}_\nu(F) \leqslant C\left[\nu\left(\mathrm{Ent}_\nu(F|\mathsf{spin}) + \mathrm{Ent}_\nu(F|\mathsf{edge})\right)\right].$$

2. Lift the even/odd factorization to spin/edge factorization
3. Lower bound $T_{\mathrm{mix}}(P_{\mathrm{SW}})$ by disagreement percolation estimates.

General graphs: Spectral independence (SI)

[ALO20] introduced SI and used it to prove a poly(n) bound for the Glauber dynamics of the hard-core gas in the uniqueness regime.

$$J(x, a; y, b) = \mu(\sigma_y = b | \sigma_x = a) - \mu(\sigma_y = b) \quad \text{for } x \neq y.$$

J is a $\mathcal{X} \times \mathcal{X}$ matrix, $\mathcal{X} = V \times [q]$. By reversibility J has real eigenvalues $\lambda_i(J)$.

Definition

μ is η-spectrally independent if $\lambda_{\max}(J) \leqslant \eta$. (Note: $\eta \geqslant 0$).

General graphs: Spectral independence (SI)

[ALO20] introduced SI and used it to prove a poly(n) bound for the Glauber dynamics of the hard-core gas in the uniqueness regime.

$$J(x, a; y, b) = \mu(\sigma_y = b | \sigma_x = a) - \mu(\sigma_y = b) \quad \text{for } x \neq y.$$

J is a $\mathcal{X} \times \mathcal{X}$ matrix, $\mathcal{X} = V \times [q]$. By reversibility J has real eigenvalues $\lambda_i(J)$.

Definition

μ is η-spectrally independent if $\lambda_{\max}(J) \leqslant \eta$. (Note: $\eta \geqslant 0$).

Theorem (ALO20)

If μ is η-SI for some $\eta = O(1)$ then the Glauber dynamics has $T_{\text{mix}} = \text{poly}(n)$.

Main idea: η-SI with $\eta = O(1)$ enables a powerful recursive scheme to prove spectral gap for the Glauber dynamics . The approach is very general and was developed in the more general setting of simplicial complexes and matroids .

Extensions of the SI approach

Theorem (CLV20)

If μ is η-SI for some $\eta = O(1)$ then the Glauber dynamics has $T_{\mathrm{mix}} = \Theta(n \log n)$.

Extensions of the SI approach

Theorem (CLV20)

If μ is η-SI for some $\eta = O(1)$ then the Glauber dynamics has $T_{\mathrm{mix}} = \Theta(n \log n)$.

Two main improvements: [CLV20] use entropy instead of variance and also use a finer recursion based on entropy factorizations to conclude a tight bound on entropy tensorization.

Extensions of the SI approach

Theorem (CLV20)

If μ is η-SI for some $\eta = O(1)$ then the Glauber dynamics has $T_{\mathrm{mix}} = \Theta(n \log n)$.

Two main improvements: [CLV20] use entropy instead of variance and also use a finer recursion based on entropy factorizations to conclude a tight bound on entropy tensorization.

Theorem (BCCPSV21)

If μ is η-SI for some $\eta = O(1)$ then the general BF holds with constant $C = O(1)$. Therefore, all α-weighted heat bath dynamics have optimal $T_{\mathrm{mix}} = O(\gamma(\alpha)^{-1} \log n)$. Moreover, for ferromagnetic Ising/Potts, the SW dynamics has $T_{\mathrm{mix}} = O(\log n)$.

Extensions of the SI approach

Theorem (CLV20)

If μ is η-SI for some $\eta = O(1)$ then the Glauber dynamics has $T_{\mathrm{mix}} = \Theta(n \log n)$ *.*

Two main improvements: [CLV20] use entropy instead of variance and also use a finer recursion based on entropy factorizations to conclude a tight bound on entropy tensorization.

Theorem (BCCPSV21)

If μ is η-SI for some $\eta = O(1)$ then the general BF holds with constant $C = O(1)$. Therefore, all α-weighted heat bath dynamics have optimal $T_{\mathrm{mix}} = O(\gamma(\alpha)^{-1} \log n)$ *. Moreover, for ferromagnetic Ising/Potts, the SW dynamics has* $T_{\mathrm{mix}} = O(\log n)$ *.*

We extend [ALO20,CLV20] and prove a multi-partite factorization

$$\mathrm{Ent}(f) \leqslant C \sum_{i=1}^{k} \mu \left[\mathrm{Ent}_{V_i}(f)\right]$$

where V_i are independent sets with $V = \cup_{i=1}^{k} V_i$, and $k \leqslant \Delta + 1$.

Some remarks on SI approach

Strength :

- It applies to more general combinatorial structures than spin systems (simplicial complexes, matroid bases)
- It allows us to prove tight bounds in some cases up to the tree uniqueness threshold . For instance, for ferro-Ising, our results on arbitrary block dynamics and SW dynamics hold for all $\beta < \beta_c(\Delta) = \log(\frac{\Delta}{\Delta-2})$. Previously known only for Glauber dynamics from Mossel,Sly'13.
- SI is very flexible : it covers all standard spatial mixing notions such as Dobrushin-uniqueness condition or SSM, and can be seen to hold as soon as μ admits some form of positive curvature , that is the existence of a contractive coupling . See below for more precise statements

Restrictions:

- our results require bounded degree $\Delta = O(1)$.
- they do not apply to unbounded spins.

Contractive coupling implies Spectral Independence

Hamming distance: $d_{\mathrm{H}}(\sigma, \sigma') = \sum_{x \in V} \mathbf{1}(\sigma_x \neq \sigma'_x)$.

W-1 distance: $W_1(\mu, \nu) = \inf\{\mathbb{E}_\pi[d_{\mathrm{H}}(\sigma, \sigma')], \ \pi \in \mathcal{C}(\mu, \nu)\}$.

A Markov chain P has (Ollivier-Ricci) curvature $\rho \in (0, 1)$ if

$$W_1(P(\sigma, \cdot), P(\sigma', \cdot)) \leqslant (1 - \rho) d_{\mathrm{H}}(\sigma, \sigma'), \qquad \forall \sigma, \sigma' \in \Omega$$

In other words, if there exists a $(1 - \rho)$-contractive coupling.

Contractive coupling implies Spectral Independence

Hamming distance: $d_{\mathrm{H}}(\sigma, \sigma') = \sum_{x \in V} \mathbf{1}(\sigma_x \neq \sigma'_x)$.

W-1 distance: $W_1(\mu, \nu) = \inf\{\mathbb{E}_\pi[d_{\mathrm{H}}(\sigma, \sigma')],\ \pi \in \mathcal{C}(\mu, \nu)\}$.

A Markov chain P has (Ollivier-Ricci) curvature $\rho \in (0, 1)$ if

$$W_1(P(\sigma, \cdot), P(\sigma', \cdot)) \leqslant (1 - \rho) d_{\mathrm{H}}(\sigma, \sigma'), \qquad \forall \sigma, \sigma' \in \Omega$$

In other words, if there exists a $(1 - \rho)$-contractive coupling.

Theorem (BCCPSV21)

If the Glauber dynamics has curvature $\rho \in (0, 1)$, then $\lambda_{\max}(J) \leqslant \frac{2}{\rho n}$. In particular, if $\rho \geqslant \varepsilon / n$ then μ is η-spectrally independent with $\eta = 2/\varepsilon$.

Contractive coupling implies Spectral Independence

Hamming distance: $d_{\mathrm{H}}(\sigma,\sigma') = \sum_{x\in V} \mathbf{1}(\sigma_x \neq \sigma'_x)$.

W-1 distance: $W_1(\mu,\nu) = \inf\{\mathbb{E}_\pi[d_{\mathrm{H}}(\sigma,\sigma')]\,,\ \pi\in\mathcal{C}(\mu,\nu)\}$.

A Markov chain P has (Ollivier-Ricci) curvature $\rho\in(0,1)$ if

$$W_1(P(\sigma,\cdot),P(\sigma',\cdot)) \leqslant (1-\rho)d_{\mathrm{H}}(\sigma,\sigma')\,, \qquad \forall\sigma,\sigma'\in\Omega$$

In other words, if there exists a $(1-\rho)$-contractive coupling.

Theorem (BCCPSV21)

If the Glauber dynamics has curvature $\rho\in(0,1)$, then $\lambda_{\max}(J) \leqslant \frac{2}{\rho n}$. In particular, if $\rho \geqslant \varepsilon/n$ then μ is η-spectrally independent with $\eta = 2/\varepsilon$.

The theorem can be considerably extended by allowing other distances and much more general Markov chains (see below).
But even in the above setting this is quite a strong result:
If Glauber has a contractive coupling then previous theorems show that all heat bath dynamics as well as SW dynamics have optimal entropy decay and optimal mixing. [$\Rightarrow$ Peres-Tetali conjecture ?]

Main ideas

Use $\lambda_{\max}(J) \leqslant \max_{(x,a)\in\mathcal{X}} S(x,a)$,
$S(x,a) = \sum_{(y,b)\in\mathcal{X}} |\mu(\sigma_y = b|\sigma_x = a) - \mu(\sigma_y = b)|$, and

$$S(x,a) = \nu[f] - \mu[f],$$

where $\nu = \mu(\cdot|\sigma_x = a)$, $f(\sigma) = \sum_{(y,b)} \mathrm{sgn}(J(x,a;y,b))\mathbf{1}(\sigma_y = b)$.

Main ideas

Use $\lambda_{\max}(J) \leqslant \max_{(x,a)\in\mathcal{X}} S(x,a)$,
$S(x,a) = \sum_{(y,b)\in\mathcal{X}} |\mu(\sigma_y = b|\sigma_x = a) - \mu(\sigma_y = b)|$, and

$$S(x,a) = \nu[f] - \mu[f],$$

where $\nu = \mu(\cdot|\sigma_x = a)$, $f(\sigma) = \sum_{(y,b)} \mathrm{sgn}(J(x,a;y,b))\mathbf{1}(\sigma_y = b)$.

Since f is 2-Lipschitz: $S(x,a) \leqslant 2W_1(\mu,\nu)$.

Lemma (BCCPSV21)

(Ω, d) finite metric space, μ,ν distr. on Ω, and P, Q two MCs with stationary distr. μ,ν resp. If (P,d) has curvature $\rho > 0$, then

$$W_{1,d}(\mu,\nu) \leq \frac{1}{\rho}\,\nu\left[W_{1,d}(P(\sigma,\cdot), Q(\sigma,\cdot))\right].$$

Main ideas

Use $\lambda_{\max}(J) \leqslant \max_{(x,a)\in\mathcal{X}} S(x,a)$,
$S(x,a) = \sum_{(y,b)\in\mathcal{X}} |\mu(\sigma_y = b|\sigma_x = a) - \mu(\sigma_y = b)|$, and

$$S(x,a) = \nu[f] - \mu[f],$$

where $\nu = \mu(\cdot|\sigma_x = a)$, $f(\sigma) = \sum_{(y,b)} \mathrm{sgn}(J(x,a;y,b))\mathbf{1}(\sigma_y = b)$.

Since f is 2-Lipschitz: $S(x,a) \leqslant 2W_1(\mu,\nu)$.

Lemma (BCCPSV21)

(Ω, d) finite metric space, μ,ν distr. on Ω, and P, Q two MCs with stationary distr. μ,ν resp. If (P,d) has curvature $\rho > 0$, then

$$W_{1,d}(\mu,\nu) \leq \frac{1}{\rho}\,\nu\left[W_{1,d}(P(\sigma,\cdot), Q(\sigma,\cdot))\right].$$

In our case: $W_1(P(\sigma,\cdot), Q(\sigma,\cdot)) \leq \frac{1}{n}$, and therefore $S(x,a) \leqslant \frac{2}{\rho n}$.

Main ideas

Use $\lambda_{\max}(J) \leqslant \max_{(x,a)\in\mathcal{X}} S(x,a)$,
$S(x,a) = \sum_{(y,b)\in\mathcal{X}} |\mu(\sigma_y = b|\sigma_x = a) - \mu(\sigma_y = b)|$, and

$$S(x,a) = \nu[f] - \mu[f],$$

where $\nu = \mu(\cdot|\sigma_x = a)$, $f(\sigma) = \sum_{(y,b)} \mathrm{sgn}(J(x,a;y,b))\mathbf{1}(\sigma_y = b)$.

Since f is 2-Lipschitz: $S(x,a) \leqslant 2W_1(\mu,\nu)$.

Lemma (BCCPSV21)

(Ω, d) finite metric space, μ,ν distr. on Ω, and P, Q two MCs with stationary distr. μ,ν resp. If (P,d) has curvature $\rho > 0$, then

$$W_{1,d}(\mu,\nu) \leq \frac{1}{\rho}\,\nu\left[W_{1,d}(P(\sigma,\cdot), Q(\sigma,\cdot))\right].$$

In our case: $W_1(P(\sigma,\cdot), Q(\sigma,\cdot)) \leq \frac{1}{n}$, and therefore $S(x,a) \leqslant \frac{2}{\rho n}$.
Proof uses Poisson eq. $(1-P)h = f - \mu[f]$,
$\nu[f] - \mu[f] = \nu[(Q-P)h]$,
$(Q-P)h(\sigma) \leqslant \mathrm{L(h)} W_{1,d}(P(\sigma,\cdot), Q(\sigma,\cdot))$, $\mathrm{L(h)} \leqslant \mathrm{L(f)}/\rho$.

Extenstions

Definition

A collection $\mathcal{P} = \{P_\tau,\ \tau \in \Omega\}$ of MCs associated with μ is Φ-local if for any two adjacent pinnings τ, τ' and $\tau' = \tau \cup (x, a)$,

$$W_1(P_\tau(\sigma, \cdot), P_{\tau'}(\sigma, \cdot)) \le \Phi.$$

$\mathcal{P}$ is arbitrary provided P_τ has stat. distr. μ^τ (pinned Gibbs meas.).

Extenstions

Definition

A collection $\mathcal{P} = \{P_\tau,\ \tau \in \Omega\}$ of MCs associated with μ is Φ-local if for any two adjacent pinnings τ, τ' and $\tau' = \tau \cup (x, a)$,

$$W_1(P_\tau(\sigma, \cdot), P_{\tau'}(\sigma, \cdot)) \leq \Phi.$$

$\mathcal{P}$ is arbitrary provided P_τ has stat. distr. μ^τ (pinned Gibbs meas.).

Theorem

If $\mathcal{P}$ is Φ-local and $(\mathcal{P}, d_{\mathrm{H}})$ has curvature $\rho > 0$, then μ is η-spectrally independent with $\eta = \frac{2\Phi}{\rho}$.

Extenstions

Definition

A collection $\mathcal{P} = \{P_\tau,\ \tau \in \Omega\}$ of MCs associated with μ is Φ-local if for any two adjacent pinnings τ, τ' and $\tau' = \tau \cup (x, a)$,

$$W_1(P_\tau(\sigma, \cdot), P_{\tau'}(\sigma, \cdot)) \leq \Phi.$$

$\mathcal{P}$ is arbitrary provided P_τ has stat. distr. μ^τ (pinned Gibbs meas.).

Theorem

If $\mathcal{P}$ is Φ-local and $(\mathcal{P}, d_{\mathrm{H}})$ has curvature $\rho > 0$, then μ is η-spectrally independent with $\eta = \frac{2\Phi}{\rho}$.

Proof: very similar to previous theorem. Moreover, it extends to non-Hamming distance $d \asymp d_{\mathrm{H}}$. This is very useful in applications.

Applications

1. For general spin systems Dobrushin uniqueness implies spectral independence. (Extending results of Hayes '06, Dyer,Goldberg,Jerrum '09 who proved that DU and related conditions imply curvature bounds).

Applications

1. For general spin systems Dobrushin uniqueness implies spectral independence. (Extending results of Hayes '06, Dyer,Goldberg,Jerrum '09 who proved that DU and related conditions imply curvature bounds).

2. Flip dynamics for q-colorings (Vigoda '00, Chen, Delcourt, Moitra, Perarnau, Postle '19) is contractive w.r.t. some $d \asymp d_{\mathrm{H}}$ as soon as $q \geqslant (\frac{11}{6} - \varepsilon_0)\Delta$.

Applications

1. For general spin systems Dobrushin uniqueness implies spectral independence. (Extending results of Hayes '06, Dyer,Goldberg,Jerrum '09 who proved that DU and related conditions imply curvature bounds).

2. Flip dynamics for q-colorings (Vigoda '00, Chen, Delcourt, Moitra, Perarnau, Postle '19) is contractive w.r.t. some $d \asymp d_{\mathrm{H}}$ as soon as $q \geqslant (\frac{11}{6} - \varepsilon_0)\Delta$.

3. Hard-core gas has a contractive coupling up to tree uniqueness threshold (ALO20, CLV20).

Applications

1. For general spin systems Dobrushin uniqueness implies spectral independence. (Extending results of Hayes '06, Dyer,Goldberg,Jerrum '09 who proved that DU and related conditions imply curvature bounds).

2. Flip dynamics for q-colorings (Vigoda '00, Chen, Delcourt, Moitra, Perarnau, Postle '19) is contractive w.r.t. some $d \asymp d_{\mathrm{H}}$ as soon as $q \geqslant (\frac{11}{6} - \varepsilon_0)\Delta$.

3. Hard-core gas has a contractive coupling up to tree uniqueness threshold (ALO20, CLV20).

4. Ferromagnetic Ising model has a contractive coupling up to tree uniqueness threshold (Chen,Liu,Vigoda '20). This extends results of Mossel,Sly'13.

Applications

1. For general spin systems Dobrushin uniqueness implies spectral independence. (Extending results of Hayes '06, Dyer,Goldberg,Jerrum '09 who proved that DU and related conditions imply curvature bounds).

2. Flip dynamics for q-colorings (Vigoda '00, Chen, Delcourt, Moitra, Perarnau, Postle '19) is contractive w.r.t. some $d \asymp d_{\mathrm{H}}$ as soon as $q \geqslant (\frac{11}{6} - \varepsilon_0)\Delta$.

3. Hard-core gas has a contractive coupling up to tree uniqueness threshold (ALO20, CLV20).

4. Ferromagnetic Ising model has a contractive coupling up to tree uniqueness threshold (Chen,Liu,Vigoda '20). This extends results of Mossel,Sly'13.

5. Ferromagnetic Potts model has contractive coupling for $\beta < \beta_1$ (Bordewich, Greenhill, Patel '16 use heat bath block dynamics with bounded block size) where $\beta_1 \approx$ tree uniqueness as $q \to \infty$.