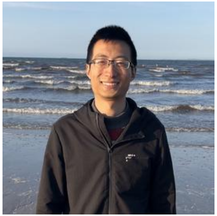
Optimal Mixing for Randomly Sampling Edge Colorings on Trees Down to the Max Degree

Xiaoyu Chen

Nanjing University

based on joint work with

Charlie Carlson



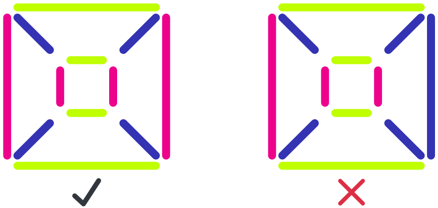
Weiming Feng

Eric Vigoda

Proper edge-coloring

- Given a graph $G = (V, E)$ with max degree Δ
- let $[q] = \{1, \cdots, q\}$ be a set of colors
- We say $\sigma : E \to [q]$ is a (proper) q-edge-coloring if

$$\forall e, f \in E, \text{ s.t. } e \cap f \neq \varnothing, \quad \sigma(e) \neq \sigma(f)$$

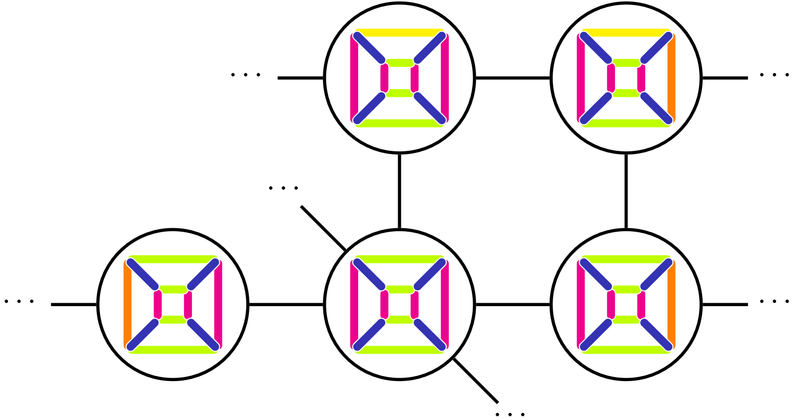


- When $q \geqslant \Delta + 1$, q-edge-coloring can be found in poly time (Vizing's thm)
- Decide whether a given graph has Δ-edge-coloring is NP-hard [Hol81]
- We are interested in **sampling** a uniformly random q-edge-coloring

Glauber dynamics

The Glauber dynamics $(X_t)_{t \geqslant 1}$ updates its current state X_t by

1. picking an edge $e \in E$ uniformly at random;
2. updating $X_t(e)$ to a uniformly random color c in $[q]$ such that c does not appear in the neighborhood of e.



Glauber dynamics

The Glauber dynamics $(X_t)_{t \geqslant 1}$ updates its current state X_t by

1. picking an edge $e \in E$ uniformly at random;
2. updating $X_t(e)$ to a uniformly random color c in $[q]$ such that c does not appear in the neighborhood of e.

Let G be a general graph with m edges and max degree Δ

- When $q \geqslant 2\Delta$, the Glauber dynamics is ergodic:
  $\mathrm{Law}(X_t) \xrightarrow{p} \mu, \quad \mu$ is the uniform distribution of proper q-edge-colorings
- When $q < 2\Delta$, the Glauber dynamics is reducible (disconnect) [HJNP19]

Glauber dynamics

The Glauber dynamics $(X_t)_{t \geqslant 1}$ updates its current state X_t by

1. picking an edge $e \in E$ uniformly at random;
2. updating $X_t(e)$ to a uniformly random color c in $[q]$ such that c does not appear in the neighborhood of e.

Let G be a general graph with m edges and max degree Δ

- When $q \geqslant 2\Delta$, the Glauber dynamics is ergodic:
  $\mathrm{Law}(X_t) \xrightarrow{p} \mu$, μ is the uniform distribution of proper q-edge-colorings
- When $q < 2\Delta$, the Glauber dynamics is reducible (disconnect) [HJNP19]

- We use the T_{mix} to denote the first time t such that $d_{TV}(X_t, \mu) \leqslant 1/100$
- Let $1 = \lambda_1 > \lambda_2 \geqslant \cdots \geqslant 0$ be eigenvalues of Glauber dynamics. Let $T_{\mathrm{rel}} := \frac{1}{1-\lambda_2}$ be the relaxation time of Glauber dynamics ($T_{\mathrm{rel}} \approx T_{\mathrm{mix}}$)

Glauber dynamics

The Glauber dynamics $(X_t)_{t \geqslant 1}$ updates its current state X_t by

1. picking an edge $e \in E$ uniformly at random;
2. updating $X_t(e)$ to a uniformly random color c in $[q]$ such that c does not appear in the neighborhood of e.

Let G be a general graph with m edges and max degree Δ

- When $q \geqslant 2\Delta$, the Glauber dynamics is ergodic:
  $\mathrm{Law}(X_t) \xrightarrow{p} \mu$, μ is the uniform distribution of proper q-edge-colorings
- When $q < 2\Delta$, the Glauber dynamics is reducible (disconnect) [HJNP19]

- We use the T_{mix} to denote the first time t such that $d_{TV}(X_t, \mu) \leqslant 1/100$
- Let $1 = \lambda_1 > \lambda_2 \geqslant \cdots \geqslant 0$ be eigenvalues of Glauber dynamics. Let $T_{\mathrm{rel}} := \frac{1}{1-\lambda_2}$ be the relaxation time of Glauber dynamics ($T_{\mathrm{rel}} \approx T_{\mathrm{mix}}$)

- When $q \geqslant (2 + o(1))\Delta$, it is known that
  $T_{\mathrm{rel}} = O(m^{10/9})$ and $T_{\mathrm{mix}} = O(m \log m)$ (when $\Delta = O(1)$) [WZZ24]

Glauber dynamics

The Glauber dynamics $(X_t)_{t \geq 1}$ updates its current state X_t by

1. picking an edge $e \in E$ uniformly at random;
2. updating $X_t(e)$ to a uniformly random color c in $[q]$ such that c does not appear in the neighborhood of e.

Let G be a **tree** with n vertices (#edge $\approx n$) and max degree Δ

- When $q \geq \Delta + 1$, the Glauber dynamics is ergodic: $\mathrm{Law}(X_t) \xrightarrow{p} \mu$
- When $q \leq \Delta$, the state space of Glauber dynamics becomes disconnected

Glauber dynamics

The Glauber dynamics $(X_t)_{t \geqslant 1}$ updates its current state X_t by

1. picking an edge $e \in E$ uniformly at random;
2. updating $X_t(e)$ to a uniformly random color c in $[q]$ such that c does not appear in the neighborhood of e.

Let G be a **tree** with n vertices (#edge $\approx n$) and max degree Δ

- When $q \geqslant \Delta + 1$, the Glauber dynamics is ergodic: $\mathrm{Law}(X_t) \xrightarrow{p} \mu$
- When $q \leqslant \Delta$, the state space of Glauber dynamics becomes disconnected
- When $q \geqslant \Delta + 1$, [DHP20] proved $T_{\text{mix}} = O(n^C)$ for $C = 60$
- When $q = \Delta + 1$, $\Delta = 2$, [DGJ06] showed that $T_{\mathrm{mix}} = \Theta(n^3 \log n)$

Glauber dynamics

The Glauber dynamics $(X_t)_{t \geqslant 1}$ updates its current state X_t by

1. picking an edge $e \in E$ uniformly at random;
2. updating $X_t(e)$ to a uniformly random color c in $[q]$ such that c does not appear in the neighborhood of e.

Let G be a **tree** with n vertices (#edge $\approx n$) and max degree Δ

- When $q \geqslant \Delta + 1$, the Glauber dynamics is ergodic: $\mathrm{Law}(X_t) \xrightarrow{p} \mu$
- When $q \leqslant \Delta$, the state space of Glauber dynamics becomes disconnected
- When $q \geqslant \Delta + 1$, [DHP20] proved $T_{\text{mix}} = O(n^C)$ for $C = 60$
- When $q = \Delta + 1$, $\Delta = 2$, [DGJ06] showed that $T_{\mathrm{mix}} = \Theta(n^3 \log n)$

Question: what is the exact mixing time for Glauber dynamics on trees?

Results on trees

Let G be a **tree** with n vertices (#edge $\approx n$) and max degree Δ

- When $q \geqslant \Delta + 1$, [DHP20] proved $T_{\mathrm{mix}} = O(n^C)$ for $C = 60$
- When $q = \Delta + 1$, $\Delta = 2$, [DGJ06] showed that $T_{\mathrm{mix}} = \Theta(n^3 \log n)$

our result: general trees with $\Delta = O(1)$

- When $q \geqslant \Delta + 2$, we show $T_{\mathrm{rel}} = O(n)$ for Glauber dynamics
- When $q \geqslant \Delta + 1$, we show $T_{\mathrm{rel}} = O(n)$ for neighboring edge dynamics

allowing update 2 neighboring edges instead of 1 edge

our result: mixing time for Glauber dynamics on trees

- $T_{\mathrm{mix}} = O(T_{\mathrm{rel}}) \times D \log n$, where D is the diameter of the tree
- $T_{\mathrm{mix}} = O(n \log^2 n)$ for Δ-regular complete trees when $q \geqslant \Delta + 2$

Results on Δ-regular complete trees

- When $q \geqslant \Delta + 1$, [DHP20] proved $T_{mix} = O(n^{2+o_\Delta(1)})$ for Glauber dynamics

our result: relaxation time for Glauber dynamics

- When $q = \Delta + 1$, $T_{rel} = \Delta n^{1+O(1/\log \Delta)}$

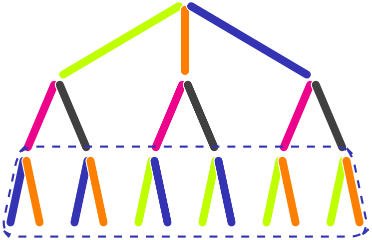
our result: lower bound on T_{rel}

- When $q = \Delta + C$, we have $T_{rel} = \Omega(\Delta n)$
- $T_{rel} = \omega(n)$ when $\Delta = \omega(1)$; justify $\Delta = O(1)$ assumption in upper bound

- Recall: when $q = \Delta + 1$, $\Delta = 2$, [DGJ06] showed that $T_{mix} = \Theta(n^3 \log n)$
- Open problem: when $q = \Delta + 1$ and $\Delta \geqslant 3$, how to close the $n^{O(1/\log \Delta)}$ gap between the lower bound and the upper bound?

Technical barriers for this problem

- Arbitrary pinning is not allowed: the state space will be disconnected
  SI based approach fails; and the comparison approach in [MSW04] fails



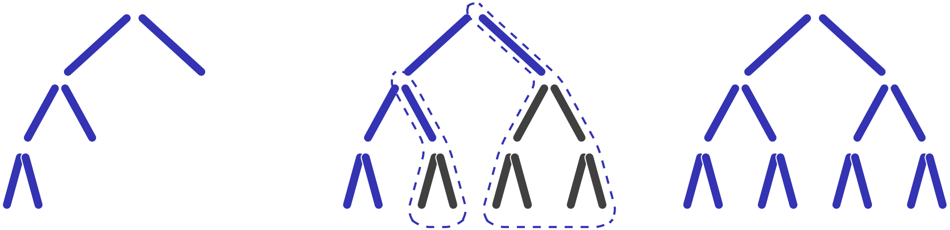
$\Delta = 3, q = \Delta + 2 = 5$

- Similar challenges have arised in sampling vertex coloring on trees [SZ17]
  This approach is hard to apply when sibling variables are not independent

In the rest part of this talk, we will focus on the proof of the following result

When $q = \Delta + 2$, then $T_{\mathrm{rel}} = O_{\Delta,q}(n)$ for Glauber dynamics on trees

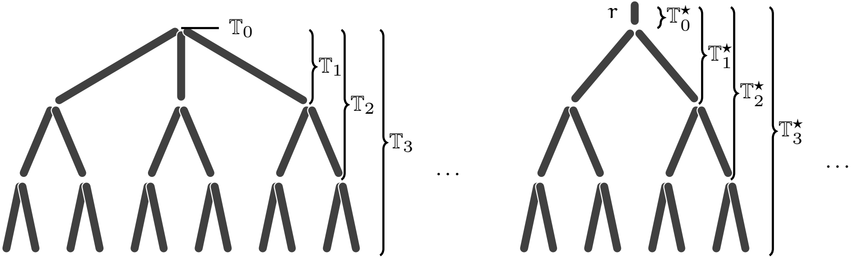
The comparison argument in [DHP20]:



Glauber dynamics: pick an ●-edge and update $\overset{T_{\mathrm{rel}}}{\approx}$ Block dynamics: pick an ●-edge and update the associated block $\overset{T_{\mathrm{rel}}}{\leqslant}$ Glauber dynamics: pick an ●-edge and update

$T_{\mathrm{rel}} = O(n)$ on Δ-regular complete trees $\Longrightarrow$ $T_{\mathrm{rel}} = O(n)$ on arbitrary trees

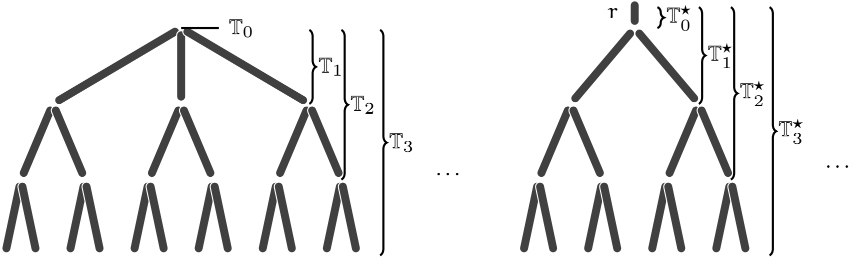
We assume Δ, q be constants and consider Δ-regular complete trees:



μ_k: uniform dist. on $\mathbb{T}_k$; $\mu_k^\star$: uniform dist. on $\mathbb{T}_k^\star$ (r only have $q - \Delta + 1$ colors)

- Let $X \sim \mu_k$, for all $f \in \Omega(\mu_k) \to \mathbb{R}$, let $F := f(X)$ be a random real number

We assume Δ, q be constants and consider Δ-regular complete trees:



μ_k: uniform dist. on $\mathbb{T}_k$; $\mu_k^\star$: uniform dist. on $\mathbb{T}_k^\star$ (r only have $q - \Delta + 1$ colors)

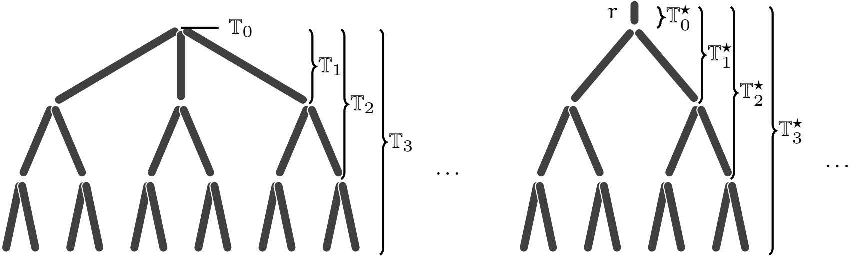
- Let $X \sim \mu_k$, for all $f \in \Omega(\mu_k) \to \mathbb{R}$, let $F := f(X)$ be a random real number
- $T_{\text{rel}} = O(n) \iff$ the approx. tensorization of variance (AT of Var):

$$\mathbf{Var}[F] \leqslant \sum_{i=1}^{k} C_i \sum_{e \in L_r(i)} \mathbb{E}[\mathbf{Var}[F \mid X_{\sim e}]] \qquad (X_{\sim e} = X(\mathbb{T}_k - e))$$

such that $C_i = O(1)$ for all i

$\langle f, f \rangle$ $\qquad$ $\langle f, (I - Q)f \rangle$ $\qquad$ (Rayleigh quotient)

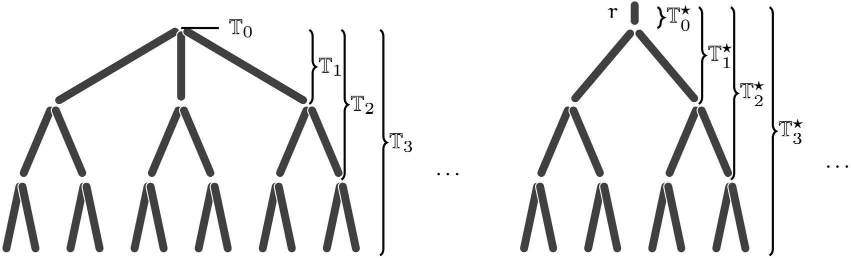
We assume Δ, q be constants and consider Δ-regular complete trees:



μ_k: uniform dist. on $\mathbb{T}_k$; $\mu_k^\star$: uniform dist. on $\mathbb{T}_k^\star$ (r only have $q - \Delta + 1$ colors)

- Let $Y \sim \mu_\ell^\star$, for $f \in \Omega(\mu_\ell^\star) \to \mathbb{R}$, let $F := f(Y)$ be a random real number

We assume Δ, q be constants and consider Δ-regular complete trees:



μ_k: uniform dist. on $\mathbb{T}_k$; $\mu_k^\star$: uniform dist. on $\mathbb{T}_k^\star$ (r only have $q - \Delta + 1$ colors)

- Let $Y \sim \mu_\ell^\star$, for $f \in \Omega(\mu_\ell^\star) \to \mathbb{R}$, let $F := f(Y)$ be a random real number

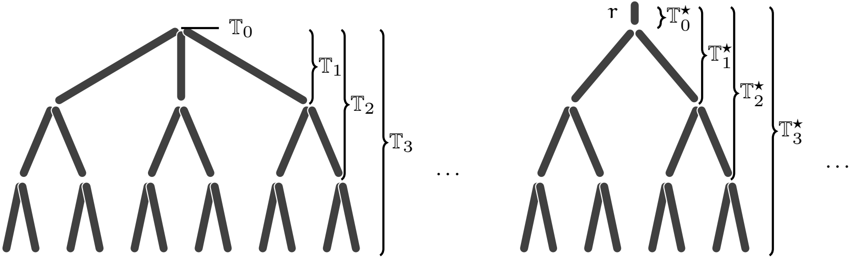
[DHP20]: AT of Var on small tree $\Rightarrow$ AT of Var on large trees (via induction)

- AT of Var for $\mu_k \Longleftarrow$ AT of Var for $\mu_\ell^\star$:

$$\mathbf{Var}[F] \leqslant \sum_{i=0}^{\ell} \alpha_i \sum_{e \in L_r(i)} \mathbb{E}[\mathbf{Var}[F \mid Y_{\sim e}]]$$

such that: $\ell = O(1)$ && $\alpha_\ell < 1$ && $\alpha_j = O(1)$, for $j < \ell$

We assume Δ, q be constants and consider Δ-regular complete trees:



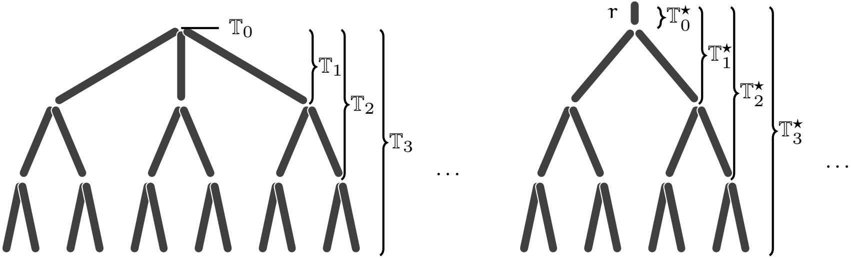
μ_k: uniform dist. on $\mathbb{T}_k$; $\mu_k^\star$: uniform dist. on $\mathbb{T}_k^\star$ (r only have $q - \Delta + 1$ colors)

- Let $Y \sim \mu_\ell^\star$, for $f \in \Omega(\mu_\ell^\star) \to \mathbb{R}$, let $F := f(Y)$ be a random real number

the $\alpha_\ell < 1$ requirement is problematic:

- [DHP20] showed a barrier that when $\ell = 1$, $\alpha_\ell > 1$
- we have considered $\ell = 2$, but doesn't work out
- for larger $\ell = O(1)$, this "small to large" argument doesn't really benefit us

We assume Δ, q be constants and consider Δ-regular complete trees:



μ_k: uniform dist. on $\mathbb{T}_k$; $\mu_k^\star$: uniform dist. on $\mathbb{T}_k^\star$ (r only have $q - \Delta + 1$ colors)

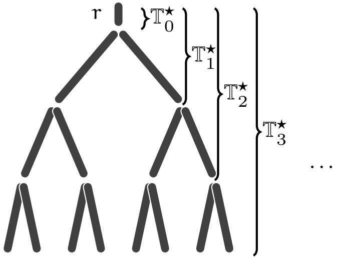
- Let $Y \sim \mu_\ell^\star$, for $f \in \Omega(\mu_\ell^\star) \to \mathbb{R}$, let $F := f(Y)$ be a random real number

theorem (prove by a refined induction)

- AT of Var for $\mu_k \Longleftarrow$ approx. **root**-tensorization of variance (A**R**T of Var):

$$\mathbf{Var}\left[\mathbb{E}\left[F \mid Y(r)\right]\right] \leqslant \sum_{i=0}^{\ell} \alpha_i \sum_{e \in L_r(i)} \mathbb{E}\left[\mathbf{Var}\left[F \mid Y_{\sim e}\right]\right]$$

such that: $\ell = O(1)$ && $\alpha_\ell < 1$ && $\alpha_j = O(1)$, for $j < \ell$



$\mu_k^\star$: uniform dist. on proper q-edge-colorings of $\mathbb{T}_k^\star$

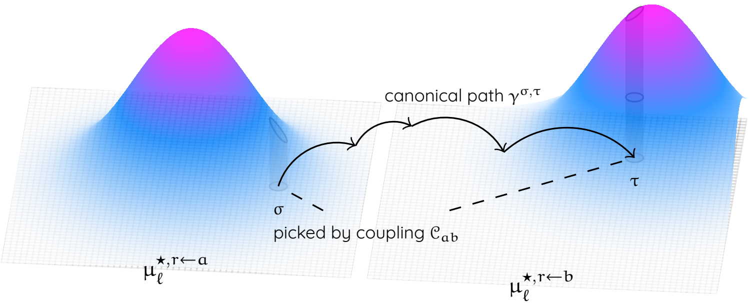
- Let $Y \sim \mu_\ell^\star$, for $f \in \Omega(\mu_\ell^\star) \to \mathbb{R}$ let $F := f(Y)$

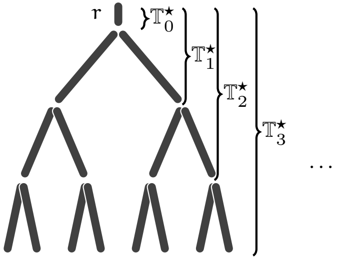
ART of Var:

$$\textbf{Var}\left[\mathbb{E}\left[F \mid Y(r)\right]\right] \leqslant \sum_{i=0}^{\ell} \alpha_i \sum_{e \in L_r(i)} \mathbb{E}\left[\textbf{Var}\left[F \mid Y_{\sim e}\right]\right]$$

s.t. $\ell = O(1)$ && $\alpha_\ell < 1$ && $\alpha_j = O(1)$, for $j < \ell$

Intuition: move the mass from $\mu_\ell^{\star, r\leftarrow a}$ to $\mu_\ell^{\star, r\leftarrow b}$ (via GD) with small congestion





$\mu_k^\star$: uniform dist. on proper q-edge-colorings of $\mathbb{T}_k^\star$

- Let $Y \sim \mu_\ell^\star$, for $f \in \Omega(\mu_\ell^\star) \to \mathbb{R}$ let $F := f(Y)$

ART of Var:

$$\textbf{Var}\left[\mathbb{E}\left[F \mid Y(r)\right]\right] \leqslant \sum_{i=0}^{\ell} \alpha_i \sum_{e \in L_r(i)} \mathbb{E}\left[\textbf{Var}\left[F \mid Y_{\sim e}\right]\right]$$

s.t. $\ell = O(1)$ && $\alpha_\ell < 1$ && $\alpha_j = O(1)$, for $j < \ell$

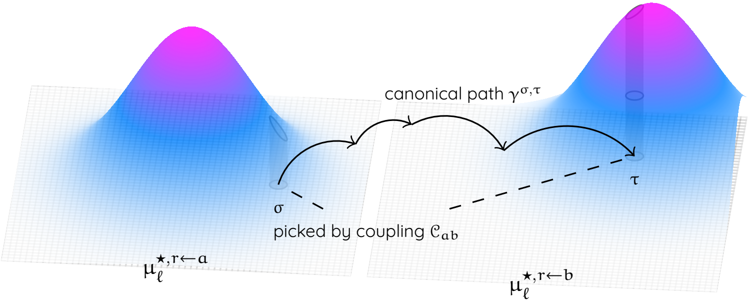
We define the congestion ξ_i for the i-th level as

$$\xi_{(s \mapsto t)} := \frac{\left(\textbf{Pr}_{(\sigma,\tau)\sim\mathcal{C}_{ab}}\left[(s \mapsto t) \in \gamma^{\sigma,\tau}\right]\right)^2}{\mu(s)Q(s,t)} \quad \text{and} \quad \xi_i := \sum_{(s\mapsto t):(s\oplus t)\in L_r(i)} \xi_{(s\mapsto t)}$$

Lemma

It holds that $\alpha_i \leqslant (\ell+1)\xi_i$ for all $i \leqslant \ell$

- $\ell, \Delta, q = O(1)$ so that $\alpha_i = O(1), \forall i \leqslant \ell$
- We only left to show $\alpha_\ell < 1$



$$2\mathbf{Var}\left[\mathbb{E}\left[F \mid Y(r)\right]\right] = \sum_{a,b\in[q-\Delta+1]} \mu^\star_{\ell,r}(a)\mu^\star_{\ell,r}(b)\left(\mathbb{E}\left[F \mid Y(r)=a\right] - \mathbb{E}\left[F \mid Y(r)=b\right]\right)^2$$

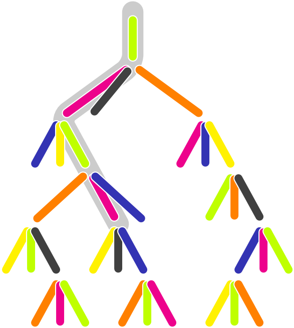
$$= \sum_{a,b\in[q-\Delta+1]} \mu^\star_{\ell,r}(a)\mu^\star_{\ell,r}(b)\left(\mathop{\mathbb{E}}_{(\sigma,\tau)\sim\mathcal{C}_{ab}}\left[f(\sigma)-f(\tau)\right]\right)^2$$

$$= \sum_{a,b\in[q-\Delta+1]} \mu^\star_{\ell,r}(a)\mu^\star_{\ell,r}(b)\left(\mathop{\mathbb{E}}_{(\sigma,\tau)\sim\mathcal{C}_{ab}}\left[\sum_{(s\mapsto t)\in\gamma^{\sigma,\tau}}(f(t)-f(s))\right]\right)^2$$

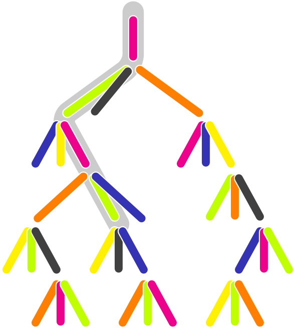
We can finish the proof by applying Cauchy's inequality and rearrange

The coupling

Let $a = $ ●, $b = $ ●; the coupling is defined by flipping the ●-●-alternating path

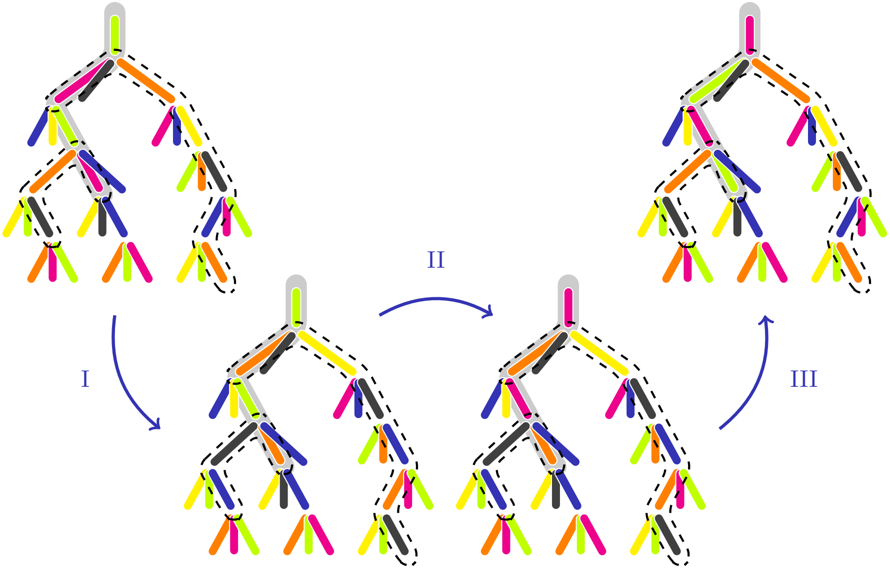


$\sigma \sim \mu_\ell^{\star, r \leftarrow ●}$

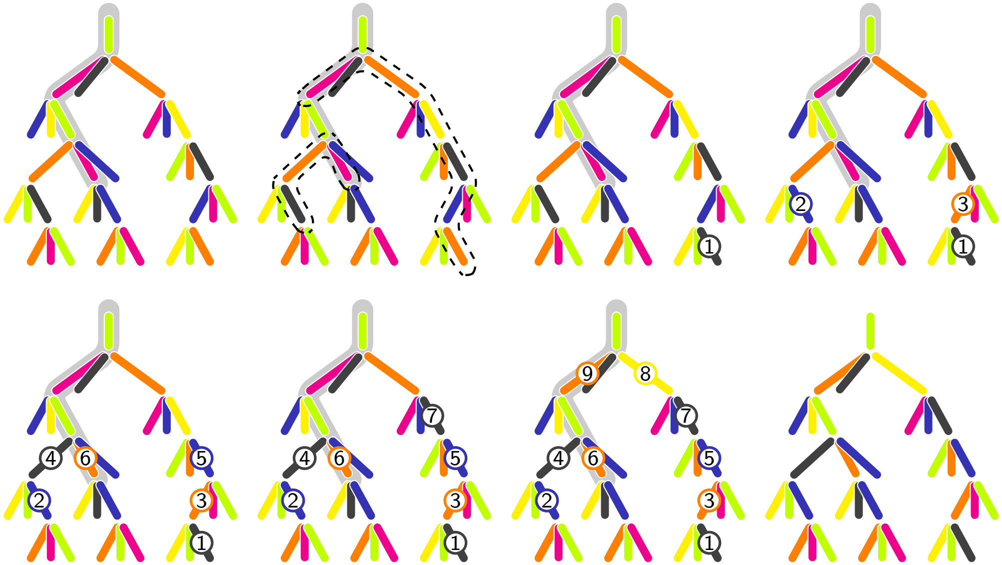


$\tau \sim \mu_\ell^{\star, r \leftarrow ●}$

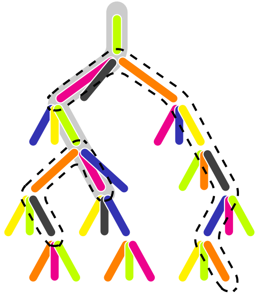
The canonical path (global plan)



The canonical path (Stage-I)



The congestion analysis (over simplified)



- Consider the probability space of $s \sim \mu_\ell^\star$
- For each ●-edge e on the alternating path, let $P(s) := \#\{\text{dashed paths that touch the leaves}\}$

Proof sketch for $\alpha_\ell \leqslant (\ell+1)\xi_\ell < 1$

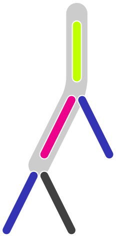
1. We can bound the congestion by

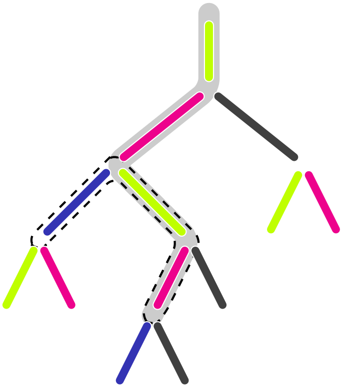
$$\xi_\ell \approx \mathbb{E}\left[\sum_{t:s\oplus t\in L_r(\ell)} \#\{(\sigma,\tau) \mid (s\mapsto t)\in\gamma^{\sigma,\tau}\}\right] = \mathbb{E}\left[P\cdot\Delta^{O(P)}\right]$$

2. There is an exp. decay so that the paths won't be very long
3. W.h.p., there is only very few of them can touch the leaves (P is small)

Why such analysis fails when $q = \Delta + 1$

It should fail, otherwise, we contradict to the $\Theta(n^3 \log n)$ result in [DGJ06] 🤔





This problem can be fixed by allowing neiboring-edge updates 🙂

Thank you

arXiv:2407.04576

Summary: general trees with $\Delta = O(1)$

- When $q \geqslant \Delta + 2$, we show $T_{\mathsf{rel}} = O(n)$ for Glauber dynamics
- When $q \geqslant \Delta + 1$, we show $T_{\mathsf{rel}} = O(n)$ for neighboring edge dynamics

---

- For GD: $T_{\mathsf{mix}} = O(T_{\mathsf{rel}}) \times D \log n$, where D is the diameter of the tree

Open problems

- When $\Delta \geqslant 3$, $q = \Delta + 1$, for GD, show $T_{\mathsf{rel}} = O(n)$ or $\omega(n)$ on Δ-reg. complete trees
  - $T_{\mathsf{rel}} = \Delta n^{1+o_\Delta(1)}$ and $\Omega(\Delta n)$ is already known (this work)
  - when $\Delta = 2$ [DGJ06] shows $T_{\mathsf{rel}} = \Theta(n^3 \log n)$ (special, depth $\neq O(\log n)$)
- How to combine the idea of coupling and canonical path in more general setting?