Deterministic counting from coupling independence

Heng Guo (University of Edinburgh)

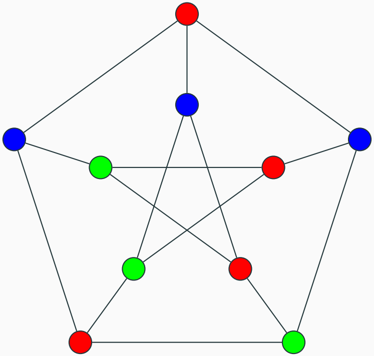
Based on joint work with

Xiaoyu Chen (Nanjing University) Weiming Feng (Hong Kong University)
Xinyuan Zhang (Nanjing University) Zongrui Zou (Nanjing University)

Amsterdam, Mar 28, 2025

Colourings

Graph (proper) colouring



3-colouring of the Petersen graph

Approximating the partition function: total number of proper colourings

(equivalent to sampling near-uniform colourings à la Jerrum, Valiant, and Vazirani, 1986)

- Glauber dynamics: rapid mixing if $q > 2\Delta$ by Jerrum (1995) / Salas and Sokal (1997)
- Flip dynamics:
  - rapid mixing if $q > \frac{11}{6}\Delta$ by Vigoda (2000)
  - improved to $q \geqslant \left(\frac{11}{6} - \varepsilon_0\right)\Delta$ for $\varepsilon_0 \approx 10^{-5}$ by ST Chen and Moitra (2019) / Delcourt, Perarnau, and Postle (2019)
  - further improved to $q \geqslant 1.809\Delta$ for $\Delta \geqslant 125$ by Carlson and Vigoda (2025)
- **NP**-hard if $q < \Delta$ and even by Galanis, Štefankovič, and Vigoda (2015)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite Δ-regular tree (namely a Bethe lattice), by Jonasson (2002).

Randomly colour a graph

Approximating the partition function: total number of proper colourings

(equivalent to sampling near-uniform colourings à la Jerrum, Valiant, and Vazirani, 1986)

- Glauber dynamics: rapid mixing if $q > 2\Delta$ by Jerrum (1995) / Salas and Sokal (1997)
- Flip dynamics:
  - rapid mixing if $q > \frac{11}{6}\Delta$ by Vigoda (2000)
  - improved to $q \geqslant \left(\frac{11}{6} - \varepsilon_0\right)\Delta$ for $\varepsilon_0 \approx 10^{-5}$ by ST Chen and Moitra (2019) / Delcourt, Perarnau, and Postle (2019)
  - further improved to $q \geqslant 1.809\Delta$ for $\Delta \geqslant 125$ by Carlson and Vigoda (2025)
- **NP**-hard if $q < \Delta$ and even by Galanis, Štefankovič, and Vigoda (2015)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite Δ-regular tree (namely a Bethe lattice), by Jonasson (2002).

Randomly colour a graph

Approximating the partition function: total number of proper colourings

(equivalent to sampling near-uniform colourings à la Jerrum, Valiant, and Vazirani, 1986)

- Glauber dynamics: rapid mixing if $q > 2\Delta$ by Jerrum (1995) / Salas and Sokal (1997)
- Flip dynamics:
  - rapid mixing if $q > \frac{11}{6}\Delta$ by Vigoda (2000)
  - improved to $q \geqslant \left(\frac{11}{6} - \varepsilon_0\right)\Delta$ for $\varepsilon_0 \approx 10^{-5}$ by ST Chen and Moitra (2019) / Delcourt, Perarnau, and Postle (2019)
  - further improved to $q \geqslant 1.809\Delta$ for $\Delta \geqslant 125$ by Carlson and Vigoda (2025)
- **NP**-hard if $q < \Delta$ and even by Galanis, Štefankovič, and Vigoda (2015)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite Δ-regular tree (namely a Bethe lattice), by Jonasson (2002).

Approximating the partition function: total number of proper colourings
(equivalent to sampling near-uniform colourings à la Jerrum, Valiant, and Vazirani, 1986)

- Glauber dynamics: rapid mixing if $q > 2\Delta$ by Jerrum (1995) / Salas and Sokal (1997)
- Flip dynamics:
  - rapid mixing if $q > \frac{11}{6}\Delta$ by Vigoda (2000)
  - improved to $q \geqslant \left(\frac{11}{6} - \varepsilon_0\right)\Delta$ for $\varepsilon_0 \approx 10^{-5}$ by ST Chen and Moitra (2019) / Delcourt, Perarnau, and Postle (2019)
  - further improved to $q \geqslant 1.809\Delta$ for $\Delta \geqslant 125$ by Carlson and Vigoda (2025)
- **NP**-hard if $q < \Delta$ and even by Galanis, Štefankovič, and Vigoda (2015)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite Δ-regular tree (namely a Bethe lattice), by Jonasson (2002).

Randomly colour a graph

Approximating the partition function: total number of proper colourings

(equivalent to sampling near-uniform colourings à la Jerrum, Valiant, and Vazirani, 1986)

- Glauber dynamics: rapid mixing if $q > 2\Delta$ by Jerrum (1995) / Salas and Sokal (1997)
- Flip dynamics:
  - rapid mixing if $q > \frac{11}{6}\Delta$ by Vigoda (2000)
  - improved to $q \geqslant \left(\frac{11}{6} - \varepsilon_0\right)\Delta$ for $\varepsilon_0 \approx 10^{-5}$ by ST Chen and Moitra (2019) / Delcourt, Perarnau, and Postle (2019)
  - further improved to $q \geqslant 1.809\Delta$ for $\Delta \geqslant 125$ by Carlson and Vigoda (2025)
- **NP**-hard if $q < \Delta$ and even by Galanis, Štefankovič, and Vigoda (2015)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite Δ-regular tree (namely a Bethe lattice), by Jonasson (2002).

Randomly colour a graph

Approximating the partition function: total number of proper colourings
(equivalent to sampling near-uniform colourings à la Jerrum, Valiant, and Vazirani, 1986)

- Glauber dynamics: rapid mixing if $q > 2\Delta$ by Jerrum (1995) / Salas and Sokal (1997)
- Flip dynamics:
  - rapid mixing if $q > \frac{11}{6}\Delta$ by Vigoda (2000)
  - improved to $q \geqslant \left(\frac{11}{6} - \varepsilon_0\right)\Delta$ for $\varepsilon_0 \approx 10^{-5}$
    by ST Chen and Moitra (2019) / Delcourt, Perarnau, and Postle (2019)
  - further improved to $q \geqslant 1.809\Delta$ for $\Delta \geqslant 125$ by Carlson and Vigoda (2025)
- **NP**-hard if $q < \Delta$ and even by Galanis, Štefankovič, and Vigoda (2015)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite Δ-regular tree (namely a Bethe lattice), by Jonasson (2002).

Approximating the partition function: total number of proper colourings
(equivalent to sampling near-uniform colourings à la Jerrum, Valiant, and Vazirani, 1986)

- Glauber dynamics: rapid mixing if $q > 2\Delta$ by Jerrum (1995) / Salas and Sokal (1997)
- Flip dynamics:
  - rapid mixing if $q > \frac{11}{6}\Delta$ by Vigoda (2000)
  - improved to $q \geqslant \left(\frac{11}{6} - \varepsilon_0\right)\Delta$ for $\varepsilon_0 \approx 10^{-5}$ by ST Chen and Moitra (2019) / Delcourt, Perarnau, and Postle (2019)
  - further improved to $q \geqslant 1.809\Delta$ for $\Delta \geqslant 125$ by Carlson and Vigoda (2025)
- **NP**-hard if $q < \Delta$ and even by Galanis, Štefankovič, and Vigoda (2015)

It is conjectured that there is a threshold and $q_c = \Delta + 1$. This is the uniqueness threshold of Gibbs measures in an infinite Δ-regular tree (namely a Bethe lattice), by Jonasson (2002).

Deterministic approximate counting

It is possible to approximate the partition function without sampling.

Fully Polynomial-Time Approximation Schemes (FPTAS):

- Gamarnik and Katz (2007): $q > 2.844\Delta$
- Lu and Yin (2013): $q > 2.581\Delta$
- JC Liu, Srivastava, and Sinclair (2019): $q \geqslant 2\Delta$
- Bencs, Berrekkal, and Regts (2024): $q \geqslant (2 - \varepsilon_1)\Delta$ for $\varepsilon_1 \approx 0.002$

The first two results use correlation decay, while the other two rely on zero-freeness of polynomials.

Deterministic approximate counting

It is possible to approximate the partition function without sampling.

Fully Polynomial-Time Approximation Schemes (FPTAS):

- Gamarnik and Katz (2007): $q > 2.844\Delta$
- Lu and Yin (2013): $q > 2.581\Delta$
- JC Liu, Srivastava, and Sinclair (2019): $q \geqslant 2\Delta$
- Bencs, Berrekkal, and Regts (2024): $q \geqslant (2 - \varepsilon_1)\Delta$ for $\varepsilon_1 \approx 0.002$

The first two results use correlation decay, while the other two rely on zero-freeness of polynomials.

It is possible to approximate the partition function without sampling.

Fully Polynomial-Time Approximation Schemes (FPTAS):

- Gamarnik and Katz (2007): $q > 2.844\Delta$
- Lu and Yin (2013): $q > 2.581\Delta$
- JC Liu, Srivastava, and Sinclair (2019): $q \geqslant 2\Delta$
- Bencs, Berrekkal, and Regts (2024): $q \geqslant (2 - \varepsilon_1)\Delta$ for $\varepsilon_1 \approx 0.002$

The first two results use correlation decay, while the other two rely on zero-freeness of polynomials.

Our result

Theorem

FPTAS exists when

- $q \geqslant \left(\frac{11}{6} - \varepsilon_0\right)\Delta$ *for* $\varepsilon_0 \approx 10^{-5}$*;*
- $q \geqslant 1.809\Delta$ *for* $\Delta \geqslant 125$*;*
- $q \geqslant \Delta + 3$ *and the girth is sufficiently large (depending only on* Δ *and* q*).*

Basically, our results match their randomised counterparts:

- ST Chen and Moitra (2019) / Delcourt, Perarnau, and Postle (2019)
- Carlson and Vigoda (2025)
- ZC Chen, KK Liu, Mani, Moitra (2023)

CI $\Rightarrow$ FPTAS

Coupling independence

Coupling independence was introduced by XY Chen and XY Zhang (2023), and it implies spectral independence (Anari, KK Liu, Oveis Gharan, 2020).

Let μ be the Gibbs distribution (e.g. the uniform distribution over all proper colourings).

Let μ^σ be the conditional distribution where σ is a partial configuration.

Roughly speaking, CI means that for any two partial configurations σ and τ that differ at one vertex, there is a coupling $\mathcal{C}$ between μ^σ and μ^τ such that the expected difference is at most some constant.

(or, using Wasserstein distance, $\mathcal{W}_1(\mu^\sigma, \mu^\tau) \leqslant C$ for some C)

Coupling independence

Coupling independence was introduced by XY Chen and XY Zhang (2023), and it implies spectral independence (Anari, KK Liu, Oveis Gharan, 2020).

Let μ be the Gibbs distribution (e.g. the uniform distribution over all proper colourings).

Let μ^σ be the conditional distribution where σ is a partial configuration.

Roughly speaking, CI means that for any two partial configurations σ and τ that differ at one vertex, there is a coupling $\mathcal{C}$ between μ^σ and μ^τ such that the expected difference is at most some constant.

(or, using Wasserstein distance, $\mathcal{W}_1(\mu^\sigma, \mu^\tau) \leqslant C$ for some C)

Coupling independence

Coupling independence was introduced by XY Chen and XY Zhang (2023), and it implies spectral independence (Anari, KK Liu, Oveis Gharan, 2020).

Let μ be the Gibbs distribution (e.g. the uniform distribution over all proper colourings).

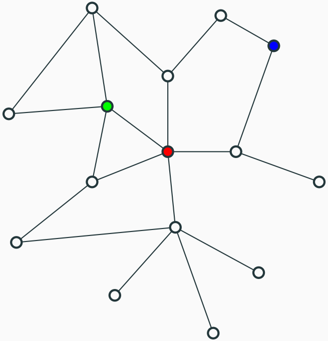
Let μ^σ be the conditional distribution where σ is a partial configuration.

Roughly speaking, CI means that for any two partial configurations σ and τ that differ at one vertex, there is a coupling $\mathcal{C}$ between μ^σ and μ^τ such that the expected difference is at most some constant.

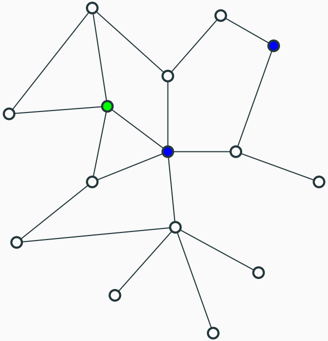
(or, using Wasserstein distance, $\mathcal{W}_1(\mu^\sigma, \mu^\tau) \leqslant C$ for some C)

CI example

σ

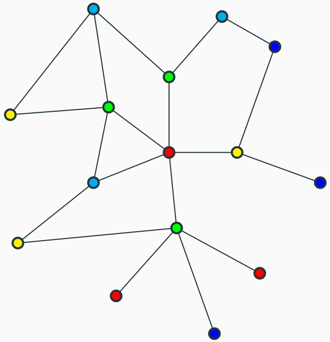


τ

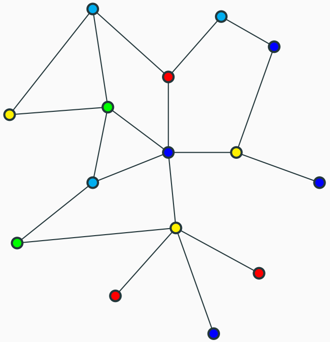


CI EXAMPLE

σ'

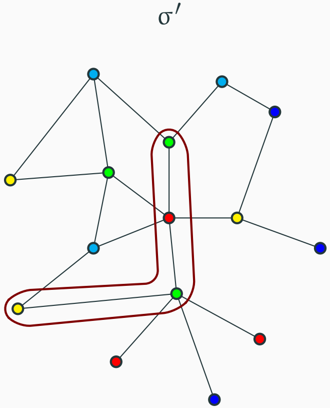


τ'



The pair (σ', τ') is drawn from the coupling. The marginal distribution of σ' (or τ') is μ^σ (or μ^τ).

CI example



The pair (σ', τ') is drawn from the coupling. The marginal distribution of σ' (or τ') is μ^{σ} (or μ^{τ}).

General result

Theorem

FPTAS exists for the partition function of any "permissive" spin systems with coupling independence.

Here, "permissive" means for any partial configuration σ, there is at least one valid configuration extending σ (so that μ^σ is well-defined).

In particular, "permissive" implies constant marginal lower bounds.

Run time is roughly $n^{\Delta^{O(C(\log b^{-1}+\log C+\log\log\Delta))}\log q}$.

Our result unifies most known FPTASes for spin systems.

Recently, ZJ Chen, Wang, CH Zhang, and ZH Zhang (2025) applied this theorem to get FPTAS for edge colourings when $q \geqslant 3\Delta$.

General result

Theorem

FPTAS exists for the partition function of any "permissive" spin systems with coupling independence.

Here, "permissive" means for any partial configuration σ, there is at least one valid configuration extending σ (so that μ^σ is well-defined).

In particular, "permissive" implies constant marginal lower bounds.

Run time is roughly $n^{\Delta^{O(C(\log b^{-1}+\log C+\log\log\Delta))}\log q}$.

Our result unifies most known FPTASes for spin systems.

Recently, ZJ Chen, Wang, CH Zhang, and ZH Zhang (2025) applied this theorem to get FPTAS for edge colourings when $q \geqslant 3\Delta$.

General result

Theorem

FPTAS exists for the partition function of any "permissive" spin systems with coupling independence.

Here, "permissive" means for any partial configuration σ, there is at least one valid configuration extending σ (so that μ^σ is well-defined).

In particular, "permissive" implies constant marginal lower bounds.

Run time is roughly $n^{\Delta^{O(C(\log b^{-1} + \log C + \log\log \Delta))} \log q}$.

Our result unifies most known FPTASes for spin systems.

Recently, ZJ Chen, Wang, CH Zhang, and ZH Zhang (2025) applied this theorem to get FPTAS for edge colourings when $q \geqslant 3\Delta$.

Total influence decay

Definition

We say *total influence decays* with rate δ, if for any $v \in V$ and any two partial configurations σ and τ that differ at v,

$$\sum_{u \in S_\ell(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant \delta(\ell).$$

We show that CI is equivalent to exponential decay of total influence.

$$\text{CI} \Rightarrow \sum_{\ell} \sum_{u \in S_\ell(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant C$$

$$\Rightarrow \exists R_0 \leqslant 2C \text{ s.t. } \sum_{u \in S_{R_0}(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant 1/2$$

Using this, construct a coupling layer by layer, such that the expected difference decay by 1/2 at each layer, and the distance between layers is at most 2C.

Total influence decay

Definition

We say *total influence decays* with rate δ, if for any $v \in V$ and any two partial configurations σ and τ that differ at v,

$$\sum_{u \in S_\ell(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant \delta(\ell).$$

We show that CI is equivalent to exponential decay of total influence.

$$\mathrm{CI} \Rightarrow \sum_\ell \sum_{u \in S_\ell(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant C$$

$$\Rightarrow \exists R_0 \leqslant 2C \text{ s.t.} \sum_{u \in S_{R_0}(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant 1/2$$

Using this, construct a coupling layer by layer, such that the expected difference decay by 1/2 at each layer, and the distance between layers is at most 2C.

Definition

We say *total influence decays* with rate δ, if for any $v \in V$ and any two partial configurations σ and τ that differ at v,

$$\sum_{u \in S_\ell(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant \delta(\ell).$$

We show that CI is equivalent to exponential decay of total influence.

$$\text{CI} \Rightarrow \sum_{\ell} \sum_{u \in S_\ell(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant C$$

$$\Rightarrow \exists R_0 \leqslant 2C \text{ s.t. } \sum_{u \in S_{R_0}(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant 1/2$$

Using this, construct a coupling layer by layer, such that the expected difference decay by 1/2 at each layer, and the distance between layers is at most 2C.

Total influence decay

Definition

We say *total influence decays* with rate δ, if for any $v \in V$ and any two partial configurations σ and τ that differ at v,

$$\sum_{u \in S_\ell(v)} D_{TV}(\mu_u^\sigma, \mu_u^\tau) \leqslant \delta(\ell).$$

We show that CI is equivalent to exponential decay of total influence.

$$\text{CI} \Rightarrow \sum_\ell \sum_{u \in S_\ell(v)} D_{TV}(\mu_u^\sigma, \mu_u^\tau) \leqslant C$$

$$\Rightarrow \exists R_0 \leqslant 2C \text{ s.t. } \sum_{u \in S_{R_0}(v)} D_{TV}(\mu_u^\sigma, \mu_u^\tau) \leqslant 1/2$$

Using this, construct a coupling layer by layer, such that the expected difference decay by 1/2 at each layer, and the distance between layers is at most 2C.

Total influence decay

Definition

We say *total influence decays* with rate δ, if for any $v \in V$ and any two partial configurations σ and τ that differ at v,

$$\sum_{u \in S_\ell(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant \delta(\ell).$$

We show that CI is equivalent to exponential decay of total influence.

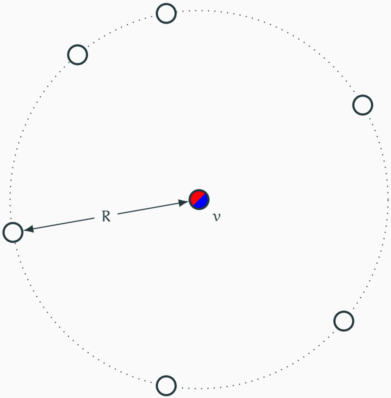
$$\begin{aligned} \mathrm{CI} \quad &\Rightarrow \sum_\ell \sum_{u \in S_\ell(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant C \\ &\Rightarrow \exists R_0 \leqslant 2C \text{ s.t. } \sum_{u \in S_{R_0}(v)} D_{TV}\left(\mu_u^\sigma, \mu_u^\tau\right) \leqslant 1/2 \end{aligned}$$

Using this, construct a coupling layer by layer, such that the expected difference decay by $1/2$ at each layer, and the distance between layers is at most $2C$.

A partial coupling

Inspired by (CLMM'23), we consider the following coupling:

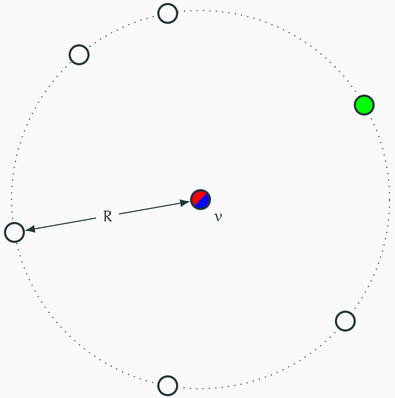
1. Choose u u.a.r. from unfixed vertices in $S_R(v)$;
2. Maximally couple between μ_u^σ and μ_u^τ;
3. If the colours at u are the same, go to 1. Otherwise, output (σ, τ).



A partial coupling

Inspired by (CLMM'23), we consider the following coupling:

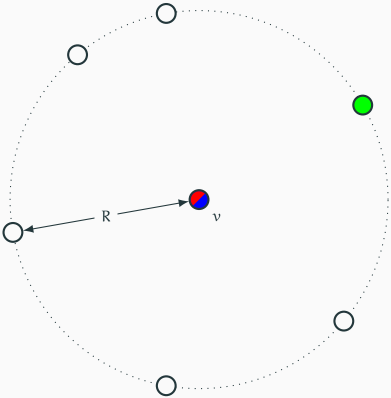
1. Choose u u.a.r. from unfixed vertices in $S_R(v)$;
2. Maximally couple between μ_u^σ and μ_u^τ;
3. If the colours at u are the same, go to 1. Otherwise, output (σ, τ).



A partial coupling

Inspired by (CLMM'23), we consider the following coupling:

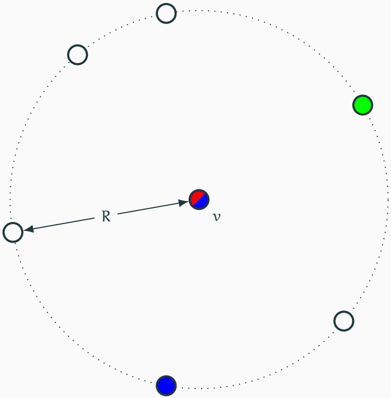
1. Choose u u.a.r. from unfixed vertices in $S_R(v)$;
2. Maximally couple between μ_u^σ and μ_u^τ;
3. If the colours at u are the same, go to 1. Otherwise, output (σ, τ).



A partial coupling

Inspired by (CLMM'23), we consider the following coupling:

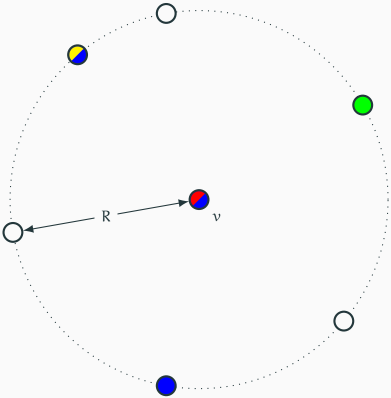
1. Choose u u.a.r. from unfixed vertices in $S_R(v)$;
2. Maximally couple between μ_u^σ and μ_u^τ;
3. If the colours at u are the same, go to 1. Otherwise, output (σ, τ).



A partial coupling

Inspired by (CLMM'23), we consider the following coupling:

1. Choose u u.a.r. from unfixed vertices in $S_R(v)$;
2. Maximally couple between μ_u^σ and μ_u^τ;
3. If the colours at u are the same, go to 1. Otherwise, output (σ, τ).



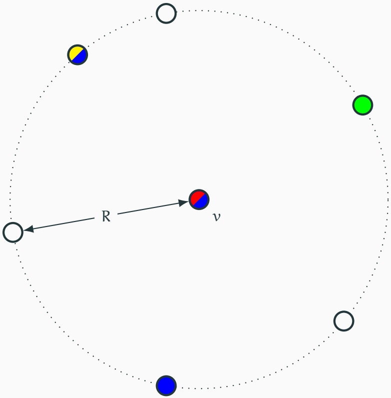
A partial coupling

Inspired by (CLMM'23), we consider the following coupling:

1. Choose u u.a.r. from unfixed vertices in $S_R(v)$;
2. Maximally couple between μ_u^σ and μ_u^τ;
3. If the colours at u are the same, go to 1. Otherwise, output (σ, τ).

The total probability of early exit is at most

$$\delta(R) \sum_{i \leqslant \Delta^R} \frac{1}{i} \leqslant O\left(\delta(R) R \log \Delta\right).$$



Solve coupling with LP

Moitra (2019) introduced a way to solve couplings with linear programming.

A coupling $\mathcal{C}$ between μ^σ and μ^τ must satisfy

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} \mathcal{C}(\sigma', \tau') = \mu^\sigma(\sigma')$$

$$\forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} \mathcal{C}(\sigma', \tau') = \mu^\tau(\tau')$$

Or, equivalently,

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} \mathcal{C}(\sigma', \tau')\mu(\sigma)/\mu(\sigma') = 1$$

$$\forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} \mathcal{C}(\sigma', \tau')\mu(\tau)/\mu(\tau') = 1$$

Thus, consider new variables:

$$x_{\sigma', \tau'} := \frac{\mathcal{C}(\sigma', \tau')\mu(\sigma)}{\mu(\sigma')} \qquad y_{\sigma', \tau'} := \frac{\mathcal{C}(\sigma', \tau')\mu(\tau)}{\mu(\tau')}$$

Solve coupling with LP

Moitra (2019) introduced a way to solve couplings with linear programming.

A coupling $\mathcal{C}$ between μ^σ and μ^τ must satisfy

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} \mathcal{C}(\sigma', \tau') = \mu^\sigma(\sigma')$$
$$\forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} \mathcal{C}(\sigma', \tau') = \mu^\tau(\tau')$$

Or, equivalently,

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} \mathcal{C}(\sigma', \tau')\mu(\sigma)/\mu(\sigma') = 1$$
$$\forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} \mathcal{C}(\sigma', \tau')\mu(\tau)/\mu(\tau') = 1$$

Thus, consider new variables:

$$x_{\sigma',\tau'} := \frac{\mathcal{C}(\sigma', \tau')\mu(\sigma)}{\mu(\sigma')} \qquad y_{\sigma',\tau'} := \frac{\mathcal{C}(\sigma', \tau')\mu(\tau)}{\mu(\tau')}$$

Solve coupling with LP

Moitra (2019) introduced a way to solve couplings with linear programming.

A coupling $\mathcal{C}$ between μ^σ and μ^τ must satisfy

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} \mathcal{C}(\sigma', \tau') = \mu^\sigma(\sigma')$$
$$\forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} \mathcal{C}(\sigma', \tau') = \mu^\tau(\tau')$$

Or, equivalently,

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} \mathcal{C}(\sigma', \tau')\mu(\sigma)/\mu(\sigma') = 1$$
$$\forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} \mathcal{C}(\sigma', \tau')\mu(\tau)/\mu(\tau') = 1$$

Thus, consider new variables:

$$x_{\sigma',\tau'} := \frac{\mathcal{C}(\sigma', \tau')\mu(\sigma)}{\mu(\sigma')} \qquad y_{\sigma',\tau'} := \frac{\mathcal{C}(\sigma', \tau')\mu(\tau)}{\mu(\tau')}$$

Solve coupling with LP (cont.)

We want to solve $r := \frac{\mu(\sigma)}{\mu(\tau)}$. In fact, $\forall \sigma' \in \Omega_\sigma$ and $\tau' \in \Omega_\tau$, $r = \frac{\mu(\sigma)}{\mu(\tau)} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}} \cdot \frac{\mu(\sigma')}{\mu(\tau')} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}}$.

Now, consider the system:

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} x_{\sigma',\tau'} = 1, \qquad \forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} y_{\sigma',\tau'} = 1,$$

$$\forall (\sigma', \tau') \in \Omega_\sigma \times \Omega_\tau, \quad r^- y_{\sigma',\tau'} \leqslant x_{\sigma',\tau'} \leqslant r^+ y_{\sigma',\tau'}.$$

If $r^- \leqslant r \leqslant r^+$, then the true values form a solution.

If there is a set of solutions, by the inequalities $r^- \leqslant \frac{\sum_{\sigma',\tau'} x_{\sigma',\tau'}}{\sum_{\sigma',\tau'} y_{\sigma',\tau'}} \leqslant r^+$.

Use the equalities $\sum_{\sigma',\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} \sum_{\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} 1 = |\Omega_\sigma|$. Thus $r^- \leqslant r \leqslant r^+$.

(Recall that $\mu(\sigma) = \frac{|\Omega_\sigma|}{|\Omega|}$ and $\mu(\tau) = \frac{|\Omega_\tau|}{|\Omega|}$.)

Thus, we can use binary search to approximate r.

Solve coupling with LP (cont.)

We want to solve $r := \frac{\mu(\sigma)}{\mu(\tau)}$. In fact, $\forall \sigma' \in \Omega_\sigma$ and $\tau' \in \Omega_\tau$, $r = \frac{\mu(\sigma)}{\mu(\tau)} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}} \cdot \frac{\mu(\sigma')}{\mu(\tau')} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}}$.

Now, consider the system:

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} x_{\sigma',\tau'} = 1, \qquad \forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} y_{\sigma',\tau'} = 1,$$

$$\forall (\sigma', \tau') \in \Omega_\sigma \times \Omega_\tau, \quad r^- y_{\sigma',\tau'} \leqslant x_{\sigma',\tau'} \leqslant r^+ y_{\sigma',\tau'}.$$

If $r^- \leqslant r \leqslant r^+$, then the true values form a solution.

If there is a set of solutions, by the inequalities $r^- \leqslant \frac{\sum_{\sigma',\tau'} x_{\sigma',\tau'}}{\sum_{\sigma',\tau'} y_{\sigma',\tau'}} \leqslant r^+$.

Use the equalities $\sum_{\sigma',\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} \sum_{\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} 1 = |\Omega_\sigma|$. Thus $r^- \leqslant r \leqslant r^+$.

(Recall that $\mu(\sigma) = \frac{|\Omega_\sigma|}{|\Omega|}$ and $\mu(\tau) = \frac{|\Omega_\tau|}{|\Omega|}$.)

Thus, we can use binary search to approximate r.

Solve coupling with LP (cont.)

We want to solve $r := \frac{\mu(\sigma)}{\mu(\tau)}$. In fact, $\forall \sigma' \in \Omega_\sigma$ and $\tau' \in \Omega_\tau$, $r = \frac{\mu(\sigma)}{\mu(\tau)} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}} \cdot \frac{\mu(\sigma')}{\mu(\tau')} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}}$.

Now, consider the system:

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} x_{\sigma',\tau'} = 1, \qquad \forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} y_{\sigma',\tau'} = 1,$$

$$\forall (\sigma', \tau') \in \Omega_\sigma \times \Omega_\tau, \quad r^- y_{\sigma',\tau'} \leqslant x_{\sigma',\tau'} \leqslant r^+ y_{\sigma',\tau'}.$$

If $r^- \leqslant r \leqslant r^+$, then the true values form a solution.

If there is a set of solutions, by the inequalities $r^- \leqslant \frac{\sum_{\sigma',\tau'} x_{\sigma',\tau'}}{\sum_{\sigma',\tau'} y_{\sigma',\tau'}} \leqslant r^+$.

Use the equalities $\sum_{\sigma',\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} \sum_{\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} 1 = |\Omega_\sigma|$. Thus $r^- \leqslant r \leqslant r^+$.

(Recall that $\mu(\sigma) = \frac{|\Omega_\sigma|}{|\Omega|}$ and $\mu(\tau) = \frac{|\Omega_\tau|}{|\Omega|}$.)

Thus, we can use binary search to approximate r.

Solve coupling with LP (cont.)

We want to solve $r := \frac{\mu(\sigma)}{\mu(\tau)}$. In fact, $\forall \sigma' \in \Omega_\sigma$ and $\tau' \in \Omega_\tau$, $r = \frac{\mu(\sigma)}{\mu(\tau)} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}} \cdot \frac{\mu(\sigma')}{\mu(\tau')} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}}$.

Now, consider the system:

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} x_{\sigma',\tau'} = 1, \qquad \forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} y_{\sigma',\tau'} = 1,$$

$$\forall (\sigma', \tau') \in \Omega_\sigma \times \Omega_\tau, \quad r^- y_{\sigma',\tau'} \leqslant x_{\sigma',\tau'} \leqslant r^+ y_{\sigma',\tau'}.$$

If $r^- \leqslant r \leqslant r^+$, then the true values form a solution.

If there is a set of solutions, by the inequalities $r^- \leqslant \frac{\sum_{\sigma',\tau'} x_{\sigma',\tau'}}{\sum_{\sigma',\tau'} y_{\sigma',\tau'}} \leqslant r^+$.

Use the equalities $\sum_{\sigma',\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} \sum_{\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} 1 = |\Omega_\sigma|$. Thus $r^- \leqslant r \leqslant r^+$.

(Recall that $\mu(\sigma) = \frac{|\Omega_\sigma|}{|\Omega|}$ and $\mu(\tau) = \frac{|\Omega_\tau|}{|\Omega|}$.)

Thus, we can use binary search to approximate r.

Solve coupling with LP (cont.)

We want to solve $r := \frac{\mu(\sigma)}{\mu(\tau)}$. In fact, $\forall \sigma' \in \Omega_\sigma$ and $\tau' \in \Omega_\tau$, $r = \frac{\mu(\sigma)}{\mu(\tau)} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}} \cdot \frac{\mu(\sigma')}{\mu(\tau')} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}}$.

Now, consider the system:

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} x_{\sigma',\tau'} = 1, \qquad \forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} y_{\sigma',\tau'} = 1,$$

$$\forall (\sigma', \tau') \in \Omega_\sigma \times \Omega_\tau, \quad r^- y_{\sigma',\tau'} \leqslant x_{\sigma',\tau'} \leqslant r^+ y_{\sigma',\tau'}.$$

If $r^- \leqslant r \leqslant r^+$, then the true values form a solution.

If there is a set of solutions, by the inequalities $r^- \leqslant \frac{\sum_{\sigma',\tau'} x_{\sigma',\tau'}}{\sum_{\sigma',\tau'} y_{\sigma',\tau'}} \leqslant r^+$.

Use the equalities $\sum_{\sigma',\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} \sum_{\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} 1 = |\Omega_\sigma|$. Thus $r^- \leqslant r \leqslant r^+$.

(Recall that $\mu(\sigma) = \frac{|\Omega_\sigma|}{|\Omega|}$ and $\mu(\tau) = \frac{|\Omega_\tau|}{|\Omega|}$.)

Thus, we can use binary search to approximate r.

Solve coupling with LP (cont.)

We want to solve $r := \frac{\mu(\sigma)}{\mu(\tau)}$. In fact, $\forall \sigma' \in \Omega_\sigma$ and $\tau' \in \Omega_\tau$, $r = \frac{\mu(\sigma)}{\mu(\tau)} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}} \cdot \frac{\mu(\sigma')}{\mu(\tau')} = \frac{x_{\sigma',\tau'}}{y_{\sigma',\tau'}}$.

Now, consider the system:

$$\forall \sigma' \in \Omega_\sigma, \quad \sum_{\tau' \in \Omega_\tau} x_{\sigma',\tau'} = 1, \qquad \forall \tau' \in \Omega_\tau, \quad \sum_{\sigma' \in \Omega_\sigma} y_{\sigma',\tau'} = 1,$$

$$\forall (\sigma', \tau') \in \Omega_\sigma \times \Omega_\tau, \quad r^- y_{\sigma',\tau'} \leqslant x_{\sigma',\tau'} \leqslant r^+ y_{\sigma',\tau'}.$$

If $r^- \leqslant r \leqslant r^+$, then the true values form a solution.

If there is a set of solutions, by the inequalities $r^- \leqslant \frac{\sum_{\sigma',\tau'} x_{\sigma',\tau'}}{\sum_{\sigma',\tau'} y_{\sigma',\tau'}} \leqslant r^+$.

Use the equalities $\sum_{\sigma',\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} \sum_{\tau'} x_{\sigma',\tau'} = \sum_{\sigma'} 1 = |\Omega_\sigma|$. Thus $r^- \leqslant r \leqslant r^+$.

(Recall that $\mu(\sigma) = \frac{|\Omega_\sigma|}{|\Omega|}$ and $\mu(\tau) = \frac{|\Omega_\tau|}{|\Omega|}$.)

Thus, we can use binary search to approximate r.

Solve coupling with LP (cont.)

Only one small issue: the system has an exponential size.

Moitra (2019) gives a succinct way to write the system:

- instead of listing all final outcomes, couple vertices one by one;
- each intermediate state has its own variable;
- write linear constraints to "guess" the transition probability between intermediate states.

The system is still too big. For an intermediate (σ_0, τ_0), $r = \frac{x_{\sigma_0,\tau_0}}{y_{\sigma_0,\tau_0}} \cdot \frac{\mu(\sigma_0)}{\mu(\tau_0)}$.

Key observation: if σ_0 and τ_0 share the same boundary, $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ is easy to compute.

Thus we prioritise this. When the probability of failure is $\exp(-\Omega(\ell))$ for ℓ steps, we can truncate the process to have a polynomial sized LP.

Solve coupling with LP (cont.)

Only one small issue: the system has an exponential size.

Moitra (2019) gives a succinct way to write the system:

- instead of listing all final outcomes, couple vertices one by one;
- each intermediate state has its own variable;
- write linear constraints to "guess" the transition probability between intermediate states.

The system is still too big. For an intermediate (σ_0, τ_0), $r = \frac{x_{\sigma_0,\tau_0}}{y_{\sigma_0,\tau_0}} \cdot \frac{\mu(\sigma_0)}{\mu(\tau_0)}$.

Key observation: if σ_0 and τ_0 share the same boundary, $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ is easy to compute.

Thus we prioritise this. When the probability of failure is $\exp(-\Omega(\ell))$ for ℓ steps, we can truncate the process to have a polynomial sized LP.

Solve coupling with LP (cont.)

Only one small issue: the system has an exponential size.

Moitra (2019) gives a succinct way to write the system:

- instead of listing all final outcomes, couple vertices one by one;
- each intermediate state has its own variable;
- write linear constraints to "guess" the transition probability between intermediate states.

The system is still too big. For an intermediate (σ_0, τ_0), $r = \frac{x_{\sigma_0,\tau_0}}{y_{\sigma_0,\tau_0}} \cdot \frac{\mu(\sigma_0)}{\mu(\tau_0)}$.

Key observation: if σ_0 and τ_0 share the same boundary, $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ is easy to compute.

Thus we prioritise this. When the probability of failure is $\exp(-\Omega(\ell))$ for ℓ steps, we can truncate the process to have a polynomial sized LP.

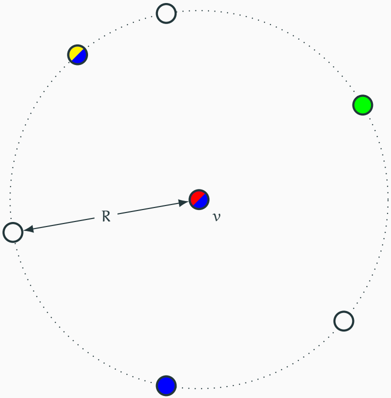
Solve coupling with LP (cont.)

We handle $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ differently.

Solve coupling with LP (cont.)

We handle $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ differently.

Recall that our partial coupling may exit early with (σ_0, τ_0) that differ on two vertices.



Solve coupling with LP (cont.)

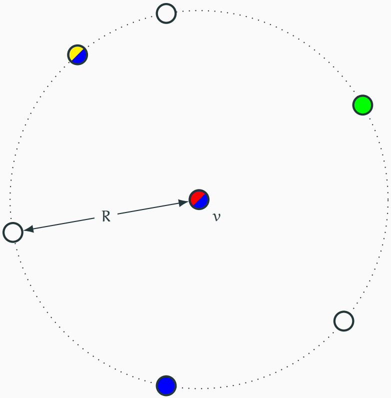
We handle $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ differently.

Recall that our partial coupling may exit early with (σ_0, τ_0) that differ on two vertices.

Thus, there is a ρ such that

$$\frac{\mu(\sigma_0)}{\mu(\tau_0)} = \frac{\mu(\sigma_0)}{\mu(\rho)} \cdot \frac{\mu(\rho)}{\mu(\tau_0)},$$

where both pairs (σ_0, ρ) and (ρ, τ_0) differ on one vertex.



Solve coupling with LP (cont.)

We handle $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ differently.

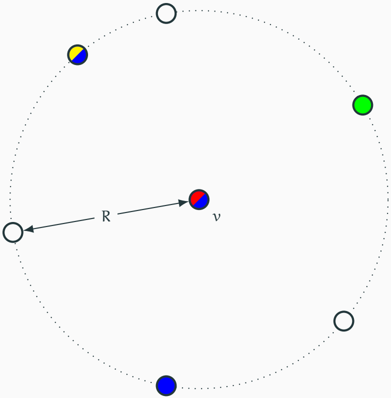
Recall that our partial coupling may exit early with (σ_0, τ_0) that differ on two vertices.

Thus, there is a ρ such that

$$\frac{\mu(\sigma_0)}{\mu(\tau_0)} = \frac{\mu(\sigma_0)}{\mu(\rho)} \cdot \frac{\mu(\rho)}{\mu(\tau_0)},$$

where both pairs (σ_0, ρ) and (ρ, τ_0) differ on one vertex.

We recursively solve $\frac{\mu(\sigma_0)}{\mu(\rho)}$ and $\frac{\mu(\rho)}{\mu(\tau_0)}$.



A recursive marginal estimator

Our LP is similar to Moitra's.

Because the radius R is a constant, the number of intermediate states (and consequently the number of constraints) is a constant for each LP.

The main question is how errors accumulate.

Say each recursive call has ε error. As $\frac{\mu(\sigma_0)}{\mu(\tau_0)} = \frac{\mu(\sigma_0)}{\mu(\sigma_1)} \cdot \frac{\mu(\sigma_1)}{\mu(\tau_0)}$, the error to $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ is 2ε.

This error is only relevant when the coupling exists early, which happens w.p. $O(\delta(R) R \log \Delta)$.

We need a new kind of constraints to bound these early exits, where we use the marginal bounds.

Eventually, we choose a constant R such that $\delta(R)$ absorbs $R \log \Delta$ and marginal bounds, so that the new error $\widehat{\varepsilon} \leqslant \varepsilon/2$.

A recursive marginal estimator

Our LP is similar to Moitra's.

Because the radius R is a constant, the number of intermediate states (and consequently the number of constraints) is a constant for each LP.

The main question is how errors accumulate.

Say each recursive call has ε error. As $\frac{\mu(\sigma_0)}{\mu(\tau_0)} = \frac{\mu(\sigma_0)}{\mu(\sigma_1)} \cdot \frac{\mu(\sigma_1)}{\mu(\tau_0)}$, the error to $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ is 2ε.

This error is only relevant when the coupling exists early, which happens w.p. $O\left(\delta(R) R \log \Delta\right)$.

We need a new kind of constraints to bound these early exits, where we use the marginal bounds.

Eventually, we choose a constant R such that $\delta(R)$ absorbs $R \log \Delta$ and marginal bounds, so that the new error $\widehat{\varepsilon} \leqslant \varepsilon/2$.

A recursive marginal estimator

Our LP is similar to Moitra's.

Because the radius R is a constant, the number of intermediate states (and consequently the number of constraints) is a constant for each LP.

The main question is how errors accumulate.

Say each recursive call has ε error. As $\frac{\mu(\sigma_0)}{\mu(\tau_0)} = \frac{\mu(\sigma_0)}{\mu(\sigma_1)} \cdot \frac{\mu(\sigma_1)}{\mu(\tau_0)}$, the error to $\frac{\mu(\sigma_0)}{\mu(\tau_0)}$ is 2ε.

This error is only relevant when the coupling exists early, which happens w.p. $O(\delta(R) R \log \Delta)$.

We need a new kind of constraints to bound these early exits, where we use the marginal bounds.

Eventually, we choose a constant R such that $\delta(R)$ absorbs $R \log \Delta$ and marginal bounds, so that the new error $\widehat{\varepsilon} \leqslant \varepsilon/2$.

How to get CI?

To use our result, one needs to establish CI.

KK Liu (2021) and Blanca, Caputo, ZC Chen, Parisi, Štefankovič, and Vigoda (2022) have shown that contractive Markov chain coupling implies SI. They actually implicitly did this through CI.

Alternatively, CI can be established by constructing the coupling directly. For example, ZJ Chen, Wang, CH Zhang, and ZH Zhang (2025) did so for edge-colourings.

Contrastive coupling implies CI

Suppose P, Q are two chains on Ω with stationary μ^σ and μ^τ, and

- for any $X \in \Omega$, $\mathcal{W}_d(P(X), Q(X)) \leqslant C$; (Typically $C = \Delta$ for Glauber dynamics)
- for any $X, Y \in \Omega$, $\mathcal{W}_d(Q(X), Q(Y)) \leqslant (1-\delta)d(X,Y)$. (Contrastive)

Let $X \sim \mu^\sigma$ and $Y \sim \mu^\tau$.

Thus, $P(X) \sim \mu^\sigma$ and $Q(Y) \sim \mu^\tau$. We have

$$\begin{aligned}\mathcal{W}_d(X,Y) &= \mathcal{W}_d(P(X), Q(Y)) \\ &\leqslant \mathcal{W}_d(P(X), Q(X)) + \mathcal{W}_d(Q(X), Q(Y)) \\ &\leqslant C + (1-\delta)\mathcal{W}_d(X,Y),\end{aligned}$$

which implies that $\mathcal{W}_d(X,Y) \leqslant \frac{C}{\delta}$.

Contrastive coupling implies CI

Suppose P, Q are two chains on Ω with stationary μ^σ and μ^τ, and

- for any $X \in \Omega$, $\mathcal{W}_d(P(X), Q(X)) \leqslant C$; (Typically $C = \Delta$ for Glauber dynamics)
- for any $X, Y \in \Omega$, $\mathcal{W}_d(Q(X), Q(Y)) \leqslant (1-\delta)d(X, Y)$. (Contrastive)

Let $X \sim \mu^\sigma$ and $Y \sim \mu^\tau$.

Thus, $P(X) \sim \mu^\sigma$ and $Q(Y) \sim \mu^\tau$. We have

$$\begin{aligned}\mathcal{W}_d(X, Y) &= \mathcal{W}_d(P(X), Q(Y)) \\ &\leqslant \mathcal{W}_d(P(X), Q(X)) + \mathcal{W}_d(Q(X), Q(Y)) \\ &\leqslant C + (1-\delta)\mathcal{W}_d(X, Y),\end{aligned}$$

which implies that $\mathcal{W}_d(X, Y) \leqslant \frac{C}{\delta}$.

Concluding remarks

Conclusions

- CI $\Rightarrow$ FPTAS

  (at least in bounded degree graphs)

- Contrastive Markov chain coupling $\Rightarrow$ FPTAS

  (at least in bounded degree graphs)

- Dobrushin-Shlosman condition $\Rightarrow$ FPTAS

  (at least in bounded degree graphs)

Open problems

- Does SI imply FPTAS as well?
  - In particular, for edge-colourings, SI holds for $q \geqslant (2 + o(1))\Delta$ (WZZ'24), but CI is only known to hold for $q \geqslant 3\Delta$ (CWZZ'25).
  - Another example is the Ising model with $\|J\|_{op} \leqslant 1$, where SI is known (Eldan, Koehler, and Zeitouni, 2021) but CI is not.
- Our method is somewhat similar to correlation decay.
  - Can it be useful in proving better algorithmic bounds for colourings?
  - Can it help with bounding zeros?

Thank you!

arXiv:2410.23225