Counting Hypergraph Colourings in the Local Lemma Regime

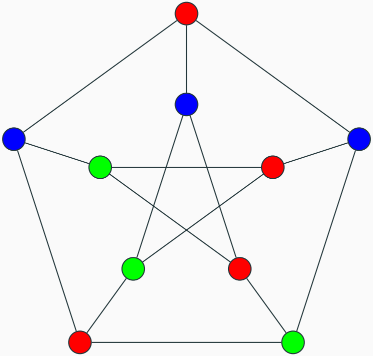
Heng Guo (University of Edinburgh)

Joint with Chao Liao (SJTU), Pinyan Lu (SHUFE), and Chihao Zhang (CUHK)

Chinese Academy of Science, Dec 26 2017

Colourings

Graph (proper) colouring

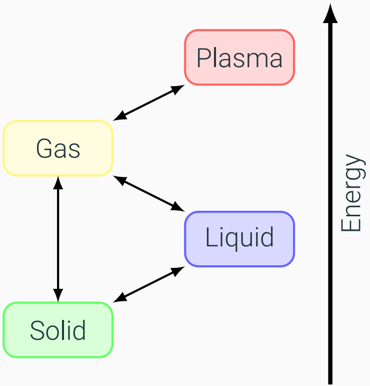


3-colouring of the Petersen graph

Phase transitions

Phase transitions:
as some parameter changes, macroscopic behaviours of the whole system change drastically.

E.g. ice $\to$ water $\to$ water vapor



Computational phase transitions

The complexity of determining whether a graph is q-colourable:

- $q = 1, 2$: easy;
- $q \geqslant 3$: **NP**-hard.

What about graphs with maximum degree Δ?

- $q \geqslant \Delta + 1$: always colourable;
- $q \geqslant \Delta - \sqrt{\Delta} + 3$: polynomial-time (Molloy, Reed '01 '14);
- $q \leqslant \Delta - \sqrt{\Delta} + c$ $(c = 1.5 + O(\Delta^{-0.5}))$: **NP**-hard (Embden-Weinert, Hougardy, and Kreuter '98).

Computational phase transitions

The complexity of determining whether a graph is q-colourable:

- $q = 1, 2$: easy;
- $q \geqslant 3$: **NP**-hard.

What about graphs with maximum degree Δ?

- $q \geqslant \Delta + 1$: always colourable;
- $q \geqslant \Delta - \sqrt{\Delta} + 3$: polynomial-time (Molloy, Reed '01 '14);
- $q \leqslant \Delta - \sqrt{\Delta} + c$ $(c = 1.5 + O(\Delta^{-0.5}))$: **NP**-hard (Embden-Weinert, Hougardy, and Kreuter '98).

Properly colour a planar graph

Things can be more complicated!

On a planar graph, determining whether a graph is q-colourable:

- $q = 2$: easy;
- $q = 3$: **NP**-hard (Dailey '80);
- $q = 4$: quadratic time (Four colour theorem) by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geqslant 5$: linear time (much simpler proof) (RSST '96).

Properly colour a planar graph

Things can be more complicated!

On a planar graph, determining whether a graph is q-colourable:

- $q = 2$: easy;
- $q = 3$: **NP**-hard (Dailey '80);
- $q = 4$: quadratic time (Four colour theorem) by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geqslant 5$: linear time (much simpler proof) (RSST '96).

Properly colour a planar graph

Things can be more complicated!

On a planar graph, determining whether a graph is q-colourable:

- $q = 2$: easy;
- $q = 3$: **NP**-hard (Dailey '80);
- $q = 4$: quadratic time (Four colour theorem)
  by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geqslant 5$: linear time (much simpler proof) (RSST '96).

Properly colour a planar graph

Things can be more complicated!

On a planar graph, determining whether a graph is q-colourable:

- $q = 2$: easy;
- $q = 3$: **NP**-hard (Dailey '80);
- $q = 4$: quadratic time (Four colour theorem)
  by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geqslant 5$: linear time (much simpler proof) (RSST '96).

Properly colour a planar graph

Things can be more complicated!

On a planar graph, determining whether a graph is q-colourable:

- $q = 2$: easy;
- $q = 3$: **NP**-hard (Dailey '80);
- $q = 4$: quadratic time (Four colour theorem)
  by Robertson, Sanders, Seymour, and Thomas (1996);
- $q \geqslant 5$: linear time (much simpler proof) (RSST '96).

Randomly colour a graph

What about generate a random proper colouring?

- $q > 2\Delta$: rapid mixing of Glauber dynamics by Jerrum (1995);
- $q > \frac{11}{6}\Delta$: rapid mixing of WSK dynamics by Vigoda (2000);
- $q < \Delta$ (even q) : **NP**-hard 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

What about generate a random proper colouring?

- $q > 2\Delta$: rapid mixing of Glauber dynamics by Jerrum (1995);
- $q > \frac{11}{6}\Delta$: rapid mixing of WSK dynamics by Vigoda (2000);
- $q < \Delta$ (even q) : **NP**-hard 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

What about generate a random proper colouring?

- $q > 2\Delta$: rapid mixing of Glauber dynamics by Jerrum (1995);
- $q > \frac{11}{6}\Delta$: rapid mixing of WSK dynamics by Vigoda (2000);
- $q < \Delta$: **NP**-hard by Galanis, Štefankovič, and Vigoda (2015); (even q)

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

What about generate a random proper colouring?

- $q > 2\Delta$: rapid mixing of Glauber dynamics by Jerrum (1995);
- $q > \frac{11}{6}\Delta$: rapid mixing of WSK dynamics by Vigoda (2000);
- $q < \Delta$: **NP**-hard by Galanis, Štefankovič, and Vigoda (2015); (even q)

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

What about generate a random proper colouring?

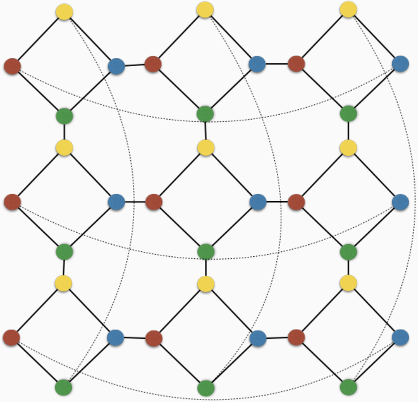
- $q > 2\Delta$: rapid mixing of Glauber dynamics by Jerrum (1995);
- $q > \frac{11}{6}\Delta$: rapid mixing of WSK dynamics by Vigoda (2000);
- $q < \Delta$ (even q) : **NP**-hard by Galanis, Štefankovič, and Vigoda (2015);

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).

FROZEN

Sometimes you just cannot let it go.

$$q = \Delta + 1 = 4$$

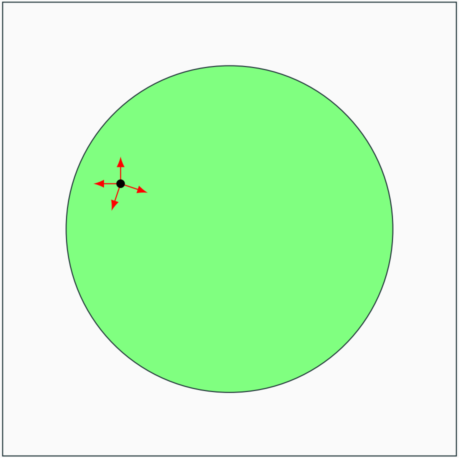


credit: Chihao Zhang

Disconnected state space

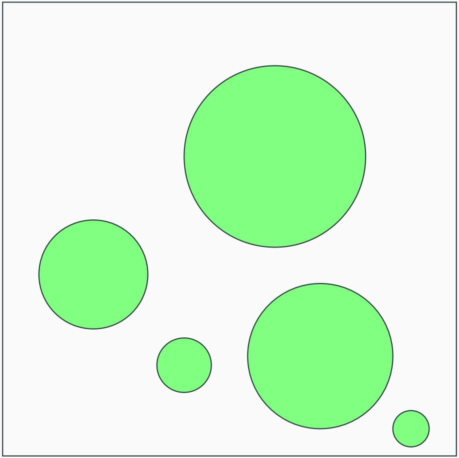
Markov chain is a random walk in the solution space.

(The solution space has to be connected!)



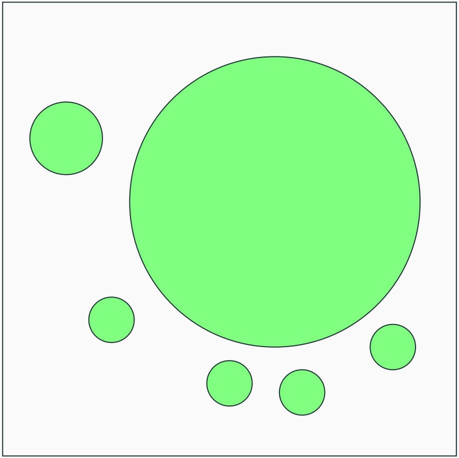
Disconnected state space

A disconnected state space is really bad.



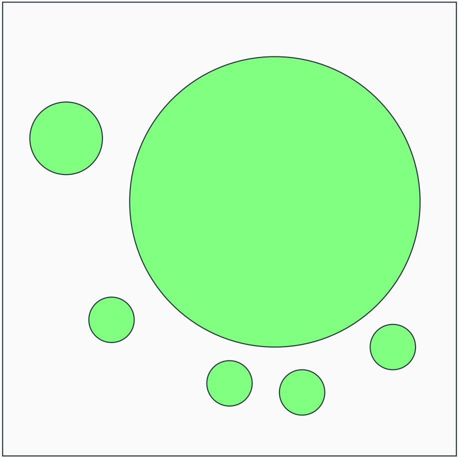
Disconnected state space

There is still some hope if a giant component dominates.

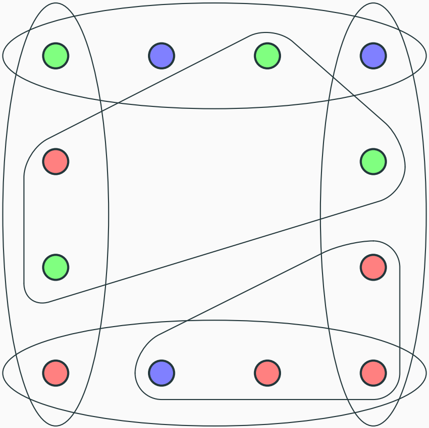


Disconnected state space

There is still some hope if a giant component dominates.
Approximately counting 4-colourings in cubic graphs
by Lu, Yang, Zhang, Zhu (2017). (Not via Markov chains!)



What about hypergraphs?



A proper hypergraph colouring is one where no edge is monochromatic.

Previous results

For k-uniform hypergraphs, Bordewich, Dyer, and Karpinski (2006) show that Glauber dynamics is rapidly mixing if

$$k \geqslant 4 \text{ and } q > \Delta \qquad \text{or} \qquad k = 3 \text{ and } q > 1.5\Delta.$$

However, Lovász local lemma implies that there exists a proper colouring if $q > 2e\Delta^{1/(k-1)}$.

Frieze and Melsted (2011) showed that if $q \ll \Delta$, then there exists a colouring so that no move is possible ("frozen").

Frieze and Anastos (2017) showed that Glauber dynamics still converges rapidly if the hypergraph is simple and $q > \max\{C_k \log n, 500k^3\Delta^{1/(k-1)}\}$.

(Simple: every two hyperedges intersect with at most one vertex.)

Previous results

For k-uniform hypergraphs, Bordewich, Dyer, and Karpinski (2006) show that Glauber dynamics is rapidly mixing if

$$k \geqslant 4 \text{ and } q > \Delta \qquad \text{or} \qquad k = 3 \text{ and } q > 1.5\Delta.$$

However, Lovász local lemma implies that there exists a proper colouring if $q > 2e\Delta^{1/(k-1)}$.

Frieze and Melsted (2011) showed that if $q \ll \Delta$, then there exists a colouring so that no move is possible ("frozen").

Frieze and Anastos (2017) showed that Glauber dynamics still converges rapidly if the hypergraph is simple and $q > \max\{C_k \log n, 500k^3\Delta^{1/(k-1)}\}$.

(Simple: every two hyperedges intersect with at most one vertex.)

Previous results

For k-uniform hypergraphs, Bordewich, Dyer, and Karpinski (2006) show that Glauber dynamics is rapidly mixing if

$$k \geqslant 4 \text{ and } q > \Delta \qquad \text{or} \qquad k = 3 \text{ and } q > 1.5\Delta.$$

However, Lovász local lemma implies that there exists a proper colouring if $q > 2e\Delta^{1/(k-1)}$.

Frieze and Melsted (2011) showed that if $q \ll \Delta$, then there exists a colouring so that no move is possible ("frozen").

Frieze and Anastos (2017) showed that Glauber dynamics still converges rapidly if the hypergraph is simple and $q > \max\{C_k \log n, 500k^3\Delta^{1/(k-1)}\}$.

(Simple: every two hyperedges intersect with at most one vertex.)

Previous results

For k-uniform hypergraphs, Bordewich, Dyer, and Karpinski (2006) show that Glauber dynamics is rapidly mixing if

$$k \geqslant 4 \text{ and } q > \Delta \qquad \text{or} \qquad k = 3 \text{ and } q > 1.5\Delta.$$

However, Lovász local lemma implies that there exists a proper colouring if $q > 2e\Delta^{1/(k-1)}$.

Frieze and Melsted (2011) showed that if $q \ll \Delta$, then there exists a colouring so that no move is possible (“frozen”).

Frieze and Anastos (2017) showed that Glauber dynamics still converges rapidly if the hypergraph is simple and $q > \max\{C_k \log n, 500k^3\Delta^{1/(k-1)}\}$.

(Simple: every two hyperedges intersect with at most one vertex.)

Our results

Theorem

For integers $\Delta \geqslant 2$, $k \geqslant 28$, and $q > 315\Delta^{\frac{14}{k-14}}$, there is an FPTAS for q-colourings in k-uniform hypergraphs with maximum degree Δ.

Theorem

For integers $\Delta \geqslant 2$, $k \geqslant 28$, and $q > 798\Delta^{\frac{16}{k-16/3}}$, there is also an almost-uniform polynomial-time sampler.

Our approach is based on a result by Moitra (2017), whose original approach in this setting would require $k > C \log \Delta$ in addition.

Our results

Theorem

For integers $\Delta \geqslant 2$, $k \geqslant 28$, *and* $q > 315\Delta^{\frac{14}{k-14}}$, *there is an FPTAS for* q*-colourings in* k*-uniform hypergraphs with maximum degree* Δ.

Theorem

For integers $\Delta \geqslant 2$, $k \geqslant 28$, *and* $q > 798\Delta^{\frac{16}{k-16/3}}$, *there is also an almost-uniform polynomial-time sampler.*

Our approach is based on a result by Moitra (2017), whose original approach in this setting would require $k > C \log \Delta$ in addition.

Our results

Theorem

For integers $\Delta \geqslant 2$, $k \geqslant 28$, and $q > 315\Delta^{\frac{14}{k-14}}$*, there is an FPTAS for* q*-colourings in* k*-uniform hypergraphs with maximum degree* Δ*.*

Theorem

For integers $\Delta \geqslant 2$, $k \geqslant 28$, and $q > 798\Delta^{\frac{16}{k-16/3}}$*, there is also an almost-uniform polynomial-time sampler.*

Our approach is based on a result by Moitra (2017), whose original approach in this setting would require $k > C \log \Delta$ in addition.

Tensor perspective

Each vertex is a EQUALITY tensor, q-dimensional, order $\leqslant \Delta$.

Each hyperedge is a COLOURING tensor, q-dimensional, order k.

Thus, the hypergraph can be viewed as a bipartite tensor network. The total number of proper colourings is the contraction of the whole tensor network.

Tensor perspective

Each vertex is a EQUALITY tensor, q-dimensional, order $\leqslant \Delta$.

Each hyperedge is a COLOURING tensor, q-dimensional, order k.

Thus, the hypergraph can be viewed as a bipartite tensor network. The total number of proper colourings is the contraction of the whole tensor network.

Moitra's approach

(with our modifications)

Lovász local lemma

Let $H = (V, \mathcal{E})$ be the hypergraph, and $\Gamma(e)$ be neighbourhood of $e \in \mathcal{E}$.

Theorem (Lovász '77)

If there exists an assignment $x : \mathcal{E} \to (0, 1)$ such that for every $e \in \mathcal{E}$ we have

$$\Pr(e \textit{ is monochromatic}) \leqslant x(e) \prod_{e' \in \Gamma(e)} (1 - x(e')), \tag{1}$$

then a proper colouring exists.

Normally we set $x(e) = \frac{1}{k\Delta}$ and the condition becomes $q > (ek\Delta)^{\frac{1}{k-1}}$.

Let $\mu_{\mathcal{C}}(\cdot)$ be the Gibbs (uniform) distribution on all proper colourings.

Theorem (Haeupler, Saha, and Srinivasan '11)

If (1) holds for every $e \in \mathcal{E}$, then for any event B, it holds that

$$\mu_{\mathcal{C}}(B) \leqslant \Pr(B) \prod_{e \in \Gamma(B)} (1 - x(e))^{-1}.$$

Lovász local lemma

Let $H = (V, \mathcal{E})$ be the hypergraph, and $\Gamma(e)$ be neighbourhood of $e \in \mathcal{E}$.

Theorem (Lovász '77)

If there exists an assignment $x : \mathcal{E} \to (0, 1)$ such that for every $e \in \mathcal{E}$ we have

$$\Pr(e \text{ is monochromatic}) \leqslant x(e) \prod_{e' \in \Gamma(e)} (1 - x(e')), \quad (1)$$

then a proper colouring exists.

Normally we set $x(e) = \frac{1}{k\Delta}$ and the condition becomes $q > (ek\Delta)^{\frac{1}{k-1}}$.

Let $\mu_{\mathcal{C}}(\cdot)$ be the Gibbs (uniform) distribution on all proper colourings.

Theorem (Haeupler, Saha, and Srinivasan '11)

If (1) holds for every $e \in \mathcal{E}$, then for any event B, it holds that

$$\mu_{\mathcal{C}}(B) \leqslant \Pr(B) \prod_{e \in \Gamma(B)} (1 - x(e))^{-1}.$$

Lovász local lemma

Let $H = (V, \mathcal{E})$ be the hypergraph, and $\Gamma(e)$ be neighbourhood of $e \in \mathcal{E}$.

Theorem (Lovász '77)

If there exists an assignment $x : \mathcal{E} \to (0, 1)$ such that for every $e \in \mathcal{E}$ we have

$$\Pr(e \textit{ is monochromatic}) \leqslant x(e) \prod_{e' \in \Gamma(e)} (1 - x(e')), \tag{1}$$

then a proper colouring exists.

Normally we set $x(e) = \frac{1}{k\Delta}$ and the condition becomes $q > (ek\Delta)^{\frac{1}{k-1}}$.

Let $\mu_{\mathcal{C}}(\cdot)$ be the Gibbs (uniform) distribution on all proper colourings.

Theorem (Haeupler, Saha, and Srinivasan '11)

If (1) holds for every $e \in \mathcal{E}$, then for any event B, it holds that

$$\mu_{\mathcal{C}}(B) \leqslant \Pr(B) \prod_{e \in \Gamma(B)} (1 - x(e))^{-1}.$$

Locally uniformity

Lemma

If $t \geqslant k$ *and* $q \geqslant (et\Delta)^{\frac{1}{k-1}}$*, then for any* $v \in V$ *and any colour* $c \in [q]$*,*

$$\frac{1}{q}\left(1 - \frac{1}{t}\right) \leqslant \Pr_{\sigma\sim\mu_{\mathcal{C}}}(\sigma(v) = c) \leqslant \frac{1}{q}\left(1 + \frac{4}{t}\right).$$

We use this lemma with $t \approx \Delta^C$ at various places with various C. Recall that our assumption is of the form $q \geqslant C'\Delta^{\frac{C''}{k}}$.

Under $\mu_{\mathcal{C}}$, all vertices are very close to uniform.

The use of this lemma is when some vertices are already coloured, namely $\mu_{\mathcal{C}}$ conditioned on a partial colouring. Replace k with the minimum number of uncoloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is $\frac{\varepsilon}{n}$-approximation of the marginals.

Locally uniformity

Lemma

If $t \geqslant k$ and $q \geqslant (et\Delta)^{\frac{1}{k-1}}$, then for any $v \in V$ and any colour $c \in [q]$,

$$\frac{1}{q}\left(1 - \frac{1}{t}\right) \leqslant \Pr_{\sigma \sim \mu_{\mathcal{C}}}(\sigma(v) = c) \leqslant \frac{1}{q}\left(1 + \frac{4}{t}\right).$$

We use this lemma with $t \approx \Delta^C$ at various places with various C. Recall that our assumption is of the form $q \geqslant C'\Delta^{\frac{C''}{k}}$.

Under $\mu_{\mathcal{C}}$, all vertices are very close to uniform.

The use of this lemma is when some vertices are already coloured, namely $\mu_{\mathcal{C}}$ conditioned on a partial colouring. Replace k with the minimum number of uncoloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is $\frac{\varepsilon}{n}$-approximation of the marginals.

Locally uniformity

Lemma

If $t \geqslant k$ *and* $q \geqslant (et\Delta)^{\frac{1}{k-1}}$*, then for any* $v \in V$ *and any colour* $c \in [q]$*,*

$$\frac{1}{q}\left(1-\frac{1}{t}\right) \leqslant \Pr_{\sigma\sim\mu_{\mathcal{C}}}(\sigma(v)=c) \leqslant \frac{1}{q}\left(1+\frac{4}{t}\right).$$

We use this lemma with $t \approx \Delta^C$ at various places with various C. Recall that our assumption is of the form $q \geqslant C'\Delta^{\frac{C''}{k}}$.

Under $\mu_{\mathcal{C}}$, all vertices are very close to uniform.

The use of this lemma is when some vertices are already coloured, namely $\mu_{\mathcal{C}}$ conditioned on a partial colouring. Replace k with the minimum number of uncoloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is $\frac{\varepsilon}{n}$-approximation of the marginals.

Locally uniformity

Lemma

If $t \geqslant k$ *and* $q \geqslant (et\Delta)^{\frac{1}{k-1}}$*, then for any* $v \in V$ *and any colour* $c \in [q]$*,*

$$\frac{1}{q}\left(1 - \frac{1}{t}\right) \leqslant \Pr_{\sigma \sim \mu_{\mathcal{C}}}(\sigma(v) = c) \leqslant \frac{1}{q}\left(1 + \frac{4}{t}\right).$$

We use this lemma with $t \approx \Delta^C$ at various places with various C. Recall that our assumption is of the form $q \geqslant C'\Delta^{\frac{C''}{k}}$.

Under $\mu_{\mathcal{C}}$, all vertices are ~~very close to~~ uniform.

The use of this lemma is when some vertices are already coloured, namely $\mu_{\mathcal{C}}$ conditioned on a partial colouring. Replace k with the minimum number of uncoloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is $\frac{\varepsilon}{n}$-approximation of the marginals.

Locally uniformity

Lemma

If $t \geqslant k$ *and* $q \geqslant (et\Delta)^{\frac{1}{k-1}}$*, then for any* $v \in V$ *and any colour* $c \in [q]$*,*

$$\frac{1}{q}\left(1-\frac{1}{t}\right) \leqslant \Pr_{\sigma\sim\mu_{\mathcal{C}}}(\sigma(v)=c) \leqslant \frac{1}{q}\left(1+\frac{4}{t}\right).$$

We use this lemma with $t \approx \Delta^{C}$ at various places with various C. Recall that our assumption is of the form $q \geqslant C'\Delta^{\frac{C''}{k}}$.

Under $\mu_{\mathcal{C}}$, all vertices are ~~very close to~~ uniform.

The use of this lemma is when some vertices are already coloured, namely $\mu_{\mathcal{C}}$ conditioned on a partial colouring. Replace k with the minimum number of uncoloured vertices among all remaining hyperedges.

A good start, but not enough. The goal is $\frac{\varepsilon}{n}$-approximation of the marginals.

Coupling

Let $\mathcal{C}_i$ be the set of colourings where v is coloured c_i, and μ_i be uniform over $\mathcal{C}_i$. We want to couple $\mathcal{C}_1$ and $\mathcal{C}_2$.

Start: $V_1 = \{v\}$, $V_{col} = \{v\}$. Maintain $V_2 = V \setminus V_1$.

Body:

1. For any hyperedge e intersecting both V_1 and V_2, let u be its first vertex. Couple u maximally assuming we know its marginal probabilities.
2. Remove all hyperedges that are satisfied in both copies.
3. If an edge has k_1 vertices coloured, put other vertices in V_1 (failed) and remove it.

Stop: all hyperedges intersecting both V_1 and $V_2 \setminus V_{col}$ are satisfied.

If $q > C\Delta^{\frac{3}{k-k_1}}$, then the coupling stops in $O(\log n)$ steps with probability $1 - O\left(\frac{1}{n^c}\right)$.

Coupling

Let $\mathcal{C}_i$ be the set of colourings where v is coloured c_i, and μ_i be uniform over $\mathcal{C}_i$. We want to couple $\mathcal{C}_1$ and $\mathcal{C}_2$.

Start: $V_1 = \{v\}$, $V_{col} = \{v\}$. Maintain $V_2 = V \setminus V_1$.

Body:

1. For any hyperedge e intersecting both V_1 and V_2, let u be its first vertex. Couple u maximally assuming we know its marginal probabilities.
2. Remove all hyperedges that are satisfied in both copies.
3. If an edge has k_1 vertices coloured, put other vertices in V_1 (failed) and remove it.

Stop: all hyperedges intersecting both V_1 and $V_2 \setminus V_{col}$ are satisfied.

If $q > C\Delta^{\frac{3}{k-k_1}}$, then the coupling stops in $O(\log n)$ steps with probability $1 - O\left(\frac{1}{n^c}\right)$.

Coupling

Let $\mathcal{C}_i$ be the set of colourings where v is coloured c_i, and μ_i be uniform over $\mathcal{C}_i$. We want to couple $\mathcal{C}_1$ and $\mathcal{C}_2$.

Start: $V_1 = \{v\}$, $V_{col} = \{v\}$. Maintain $V_2 = V \setminus V_1$.

Body:

1. For any hyperedge e intersecting both V_1 and V_2, let u be its first vertex. Couple u maximally assuming we know its marginal probabilities.
2. Remove all hyperedges that are satisfied in both copies.
3. If an edge has k_1 vertices coloured, put other vertices in V_1 (failed) and remove it.

Stop: all hyperedges intersecting both V_1 and $V_2 \setminus V_{col}$ are satisfied.

If $q > C\Delta^{\frac{3}{k-k_1}}$, then the coupling stops in $O(\log n)$ steps with probability $1 - O\left(\frac{1}{n^c}\right)$.

Coupling tree

Coupling tree $\mathcal{T}$: each node is a pair of partial colourings (x, y).

The children of (x, y) are all q^2 ways to extend them to the next vertex.

All information can be recovered by simulating the coupling from the start.

Moitra (2017) marks what vertices to couple in advance, whereas our coupling is adaptive.

Coupling tree

Coupling tree $\mathcal{T}$: each node is a pair of partial colourings (x, y).

The children of (x, y) are all q^2 ways to extend them to the next vertex.

All information can be recovered by simulating the coupling from the start.

Moitra (2017) marks what vertices to couple in advance, whereas our coupling is adaptive.

Coupling tree

Coupling tree $\mathcal{T}$: each node is a pair of partial colourings (x, y).

The children of (x, y) are all q^2 ways to extend them to the next vertex.

All information can be recovered by simulating the coupling from the start.

Moitra (2017) marks what vertices to couple in advance, whereas our coupling is adaptive.

A SAMPLER

Let $q_i = \frac{|\mathcal{C}_i|}{|\mathcal{C}|}$ for $i = 1, 2$.

A sampler $\mathbb{S}$:

1. Sample (X, Y) using the coupling;
2. Uniformly output a colouring in $\mathcal{C}_X$ with probability $\frac{q_1}{q_1+q_2}$. Otherwise uniformly output a colouring in $\mathcal{C}_Y$.

The output of $\mathbb{S}$ is uniform over $\mathcal{C}_1 \cup \mathcal{C}_2$.

Linear program

We cannot really run the coupling. Instead, we set up a linear program. The variables are to mimic:

$$p^x_{x,y} = q_1 \cdot \frac{|\mathcal{C}_1 \cup \mathcal{C}_2|}{|\mathcal{C}_x|} \cdot \mu_{cp}(x,y);$$

$$p^y_{x,y} = q_2 \cdot \frac{|\mathcal{C}_1 \cup \mathcal{C}_2|}{|\mathcal{C}_y|} \cdot \mu_{cp}(x,y);$$

$$0 \leqslant p^x_{x,y}, p^y_{x,y} \leqslant 1.$$

The meaning of $p^x_{x,y}$ is, for any $\sigma \in \mathcal{C}_x$,

"conditioned on $\mathbb{S}$ outputting σ, the coupling reaches (x,y)".

This definition is independent from σ.

Linear program

We cannot really run the coupling. Instead, we set up a linear program. The variables are to mimic:

$$p^{x}_{x,y} = q_1 \cdot \frac{|\mathcal{C}_1 \cup \mathcal{C}_2|}{|\mathcal{C}_x|} \cdot \mu_{cp}(x,y);$$

$$p^{y}_{x,y} = q_2 \cdot \frac{|\mathcal{C}_1 \cup \mathcal{C}_2|}{|\mathcal{C}_y|} \cdot \mu_{cp}(x,y);$$

$$0 \leqslant p^{x}_{x,y},\, p^{y}_{x,y} \leqslant 1.$$

The meaning of $p^{x}_{x,y}$ is, for any $\sigma \in \mathcal{C}_x$,

“conditioned on $\mathbb{S}$ outputing σ, the coupling reaches (x,y)”.

This definition is independent from σ.

Constraints 1

From the definition: $\frac{p^x_{x,y}}{p^y_{x,y}} = \frac{q_1}{q_2} \cdot \frac{|\mathcal{C}_y|}{|\mathcal{C}_x|}$.

If (x, y) is a leaf in $\mathcal{T}$, then we can compute $\frac{|\mathcal{C}_x|}{|\mathcal{C}_y|}$ in time $\exp(|V_1|)$.

Constraints 1: For every leaf (x, y), we have the constraints:

$$\underline{r} \leqslant \frac{p^x_{x,y}}{p^y_{x,y}} \cdot \frac{|\mathcal{C}_x|}{|\mathcal{C}_y|} \leqslant \overline{r}.$$

Here $\underline{r}$ and $\overline{r}$ are our guessed lower and upper bounds for $\frac{q_1}{q_2}$.

Constraints 1

From the definition: $\frac{p^{x}_{x,y}}{p^{y}_{x,y}} = \frac{q_1}{q_2} \cdot \frac{|\mathcal{C}_y|}{|\mathcal{C}_x|}$.

If (x, y) is a leaf in $\mathcal{T}$, then we can compute $\frac{|\mathcal{C}_x|}{|\mathcal{C}_y|}$ in time $\exp(|V_1|)$.

Constraints 1: For every leaf (x, y), we have the constraints:

$$\underline{r} \leqslant \frac{p^{x}_{x,y}}{p^{y}_{x,y}} \cdot \frac{|\mathcal{C}_x|}{|\mathcal{C}_y|} \leqslant \overline{r}.$$

Here $\underline{r}$ and $\overline{r}$ are our guessed lower and upper bounds for $\frac{q_1}{q_2}$.

Constraints 1

From the definition: $\frac{p^x_{x,y}}{p^y_{x,y}} = \frac{q_1}{q_2} \cdot \frac{|\mathcal{C}_y|}{|\mathcal{C}_x|}$.

If (x, y) is a leaf in $\mathcal{T}$, then we can compute $\frac{|\mathcal{C}_x|}{|\mathcal{C}_y|}$ in time $\exp(|V_1|)$.

Constraints 1: For every leaf (x, y), we have the constraints:

$$\underline{r} \leqslant \frac{p^x_{x,y}}{p^y_{x,y}} \cdot \frac{|\mathcal{C}_x|}{|\mathcal{C}_y|} \leqslant \overline{r}.$$

Here $\underline{r}$ and $\overline{r}$ are our guessed lower and upper bounds for $\frac{q_1}{q_2}$.

Constraints 2

Constraints 2: For the root $(x_0, y_0) \in \mathcal{T}$, we have

$$p^{x_0}_{x_0,y_0} = p^{y_0}_{x_0,y_0} = 1.$$

Moreover, for every non-leaf $(x, y) \in \mathcal{T}$, let u be the next vertex to couple. We have:

$$\text{for every } c \in [q], \quad p^{x}_{x,y} = \sum_{c' \in [q]} p^{x^{u \leftarrow c}}_{x^{u \leftarrow c}, y^{u \leftarrow c'}};$$

$$\text{for every } c \in [q], \quad p^{y}_{x,y} = \sum_{c' \in [q]} p^{y^{u \leftarrow c}}_{x^{u \leftarrow c'}, y^{u \leftarrow c}}.$$

Recover the marginals

Due to **Constraints 2**, a simple induction shows that for every $\sigma \in \mathcal{C}_1$,

$$\sum_{(x,y)\in\mathcal{L}(T):\ \sigma\models x} p^x_{x,y} = 1\,.$$

Rewrite $|\mathcal{C}_1|$:

$$\begin{aligned}|\mathcal{C}_1| = \sum_{\sigma\in\mathcal{C}_1} 1 &= \sum_{\sigma\in\mathcal{C}_1} \sum_{(x,y)\in\mathcal{L}(\mathcal{T}):\sigma\models x} p^x_{x,y} \\ &= \sum_{(x,y)\in\mathcal{L}(\mathcal{T})} \sum_{\sigma\models x} p^x_{x,y} \\ &= \sum_{(x,y)\in\mathcal{L}(\mathcal{T})} p^x_{x,y}\,|C_x|\,.\end{aligned}$$

Similar equalities hold on the Y side, implying:

$$\frac{q_1}{q_2} = \frac{|\mathcal{C}_1|}{|\mathcal{C}_2|} = \frac{\sum_{(x,y)\in\mathcal{L}(\mathcal{T})} p^x_{x,y}\,|C_x|}{\sum_{(x,y)\in\mathcal{L}(\mathcal{T})} p^y_{x,y}\,|C_y|}\,.$$

Recover the marginals (cont.)

$$\frac{q_1}{q_2} = \frac{\sum_{(x,y)\in\mathcal{L}(\mathcal{T})} p^x_{x,y}\,|C_x|}{\sum_{(x,y)\in\mathcal{L}(\mathcal{T})} p^y_{x,y}\,|C_y|}$$

Recall **Constraints 1**. For any $(x, y) \in \mathcal{L}(\mathcal{T})$,

$$\underline{r} \leqslant \frac{p^x_{x,y}\,|C_x|}{p^y_{x,y}\,|C_y|} \leqslant \overline{r}.$$

It implies that

$$\underline{r} \leqslant \frac{q_1}{q_2} \leqslant \overline{r}.$$

Constraints 3

Unfortunately, the whole linear program is exponentially large, but the coupling stops at $O(\log n)$ size whp. If we truncate at $O(\log n)$ levels, the error should be small, due to local uniformity.

Constraints 3: For every $c, c' \in [q]$ that $c \neq c'$:

$$p^{x^{u \leftarrow c}}_{x^{u \leftarrow c}, y^{u \leftarrow c'}} \leqslant \frac{5}{t} \cdot p^{x}_{x,y};$$
$$p^{y^{u \leftarrow c'}}_{x^{u \leftarrow c}, y^{u \leftarrow c'}} \leqslant \frac{5}{t} \cdot p^{y}_{x,y}.$$

The quantity t will eventually be set as $C(k\Delta)^6$.

Constraints 3

Unfortunately, the whole linear program is exponentially large, but the coupling stops at $O(\log n)$ size whp. If we truncate at $O(\log n)$ levels, the error should be small, due to local uniformity.

Constraints 3: For every $c, c' \in [q]$ that $c \neq c'$:

$$p^{x^{u \leftarrow c}}_{x^{u \leftarrow c}, y^{u \leftarrow c'}} \leqslant \frac{5}{t} \cdot p^{x}_{x,y};$$
$$p^{y^{u \leftarrow c'}}_{x^{u \leftarrow c}, y^{u \leftarrow c'}} \leqslant \frac{5}{t} \cdot p^{y}_{x,y}.$$

The quantity t will eventually be set as $C(k\Delta)^6$.

Truncation error

Recall that

$$|\mathcal{C}_1| = \sum_{\sigma \in \mathcal{C}_1} \sum_{(x,y) \in \mathcal{L}(\mathcal{T}) : \sigma \models x} p_{x,y}^{x}.$$

The truncation error from a particular $\sigma \in \mathcal{C}_1$ comes from conditioned on outputing σ, the coupling lasts too long.

Such "bad" colouring does exist (all early vertices are monochromatic).

We prove two things:

1. The fraction of "bad" colourings is small;
2. For every "good" colouring, the truncation error is small because of **Constraints 3**.

Truncation error

Recall that

$$|\mathcal{C}_1| = \sum_{\sigma \in \mathcal{C}_1} \sum_{(x,y) \in \mathcal{L}(\mathcal{T}) : \sigma \models x} p^{x}_{x,y}.$$

The truncation error from a particular $\sigma \in \mathcal{C}_1$ comes from conditioned on outputing σ, the coupling lasts too long.

Such “bad” colouring does exist (all early vertices are monochromatic).

We prove two things:

1. The fraction of “bad” colourings is small;
2. For every “good” colouring, the truncation error is small because of **Constraints 3**.

Truncation error

Recall that

$$|\mathcal{C}_1| = \sum_{\sigma \in \mathcal{C}_1} \sum_{(x,y) \in \mathcal{L}(\mathcal{T}) : \sigma \models x} p^{x}_{x,y}.$$

The truncation error from a particular $\sigma \in \mathcal{C}_1$ comes from conditioned on outputing σ, the coupling lasts too long.

Such “bad” colouring does exist (all early vertices are monochromatic).

We prove two things:

1. The fraction of “bad” colourings is small;
2. For every “good” colouring, the truncation error is small because of **Constraints 3**.

Counting and sampling

So far we are calculating the marginal probability, which requires that there are sufficiently many vertices in all hyperedges.

- For approximate counting, we use the local lemma again, to find an assignment so that every hyperedge is satisfied by the first $\frac{k}{14}$ vertices. Then we compute the marginal probability of this assignment by fixing vertices one by one.

- For sampling, we use the marginal to colour vertices, similar to the coupling process. With high probability, the remaining connected components are of size $O(\log n)$.

Counting and sampling

So far we are calculating the marginal probability, which requires that there are sufficiently many vertices in all hyperedges.

- For approximate counting, we use the local lemma again, to find an assignment so that every hyperedge is satisfied by the first $\frac{k}{14}$ vertices. Then we compute the marginal probability of this assignment by fixing vertices one by one.

- For sampling, we use the marginal to colour vertices, similar to the coupling process. With high probability, the remaining connected components are of size $O(\log n)$.

Counting and sampling

So far we are calculating the marginal probability, which requires that there are sufficiently many vertices in all hyperedges.

- For approximate counting, we use the local lemma again, to find an assignment so that every hyperedge is satisfied by the first $\frac{k}{14}$ vertices. Then we compute the marginal probability of this assignment by fixing vertices one by one.

- For sampling, we use the marginal to colour vertices, similar to the coupling process. With high probability, the remaining connected components are of size $O(\log n)$.

Concluding remarks

Open problems

- What is the correct threshold for hypergraph colouring?
  - Is it $q \asymp \Delta^{\frac{2}{k}}$?
- What about **NP**-hardness of sampling hypergraph colourings?
- Does this method work for general LLL?
- What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?

Open problems

- What is the correct threshold for hypergraph colouring?
  - Is it $q \asymp \Delta^{\frac{2}{k}}$?
- What about **NP**-hardness of sampling hypergraph colourings?
- Does this method work for general LLL?
- What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?

Open problems

- What is the correct threshold for hypergraph colouring?
    - Is it $q \asymp \Delta^{\frac{2}{k}}$?
- What about **NP**-hardness of sampling hypergraph colourings?
- Does this method work for general LLL?
- What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?

Open problems

- What is the correct threshold for hypergraph colouring?
  - Is it $q \asymp \Delta^{\frac{2}{k}}$?
- What about **NP**-hardness of sampling hypergraph colourings?
- Does this method work for general LLL?
- What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?

Open problems

- What is the correct threshold for hypergraph colouring?
    - Is it $q \asymp \Delta^{\frac{2}{k}}$?
- What about **NP**-hardness of sampling hypergraph colourings?
- Does this method work for general LLL?
- What is the relationship between this method and traditional ones (Markov chains, spatial mixing, etc.)?

A professor is one who can speak on any subject for precisely fifty minutes.

— Norbert Wiener

Thank you!

arXiv:1711.03396