CS 263: Counting and Sampling

Nima Anari

Stanford University

slides for

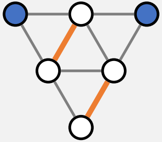
High-Dimensional Expanders

Review

Monomer-dimer system

Prob of matching $\propto$

$$\prod_{e\in M} \underset{\substack{\uparrow \\ \text{dimer}}}{\lambda_e} \cdot \prod_{v \nsim M} \underset{\substack{\uparrow \\ \text{monomer}}}{z_v}$$



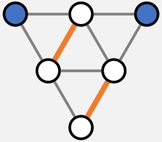
Review

Monomer-dimer system

Prob of matching $\propto$

$$\prod_{e \in M} \lambda_e \cdot \prod_{v \not\in M} z_v$$

dimer monomer



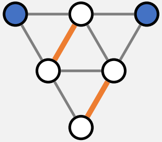
- Assume wlog $z_v = 1$

Review

Monomer-dimer system

Prob of matching $\propto$

$$\prod_{e\in M}\underset{\substack{\uparrow\\\text{dimer}}}{\lambda_e}\cdot\prod_{v\not\sim M}\underset{\substack{\uparrow\\\text{monomer}}}{z_v}$$



- Assume wlog $z_v = 1$
- [Jerrum-Sinclair]: Metropolis mixes in

$$\text{poly}(n, \lambda_{\text{max}})$$

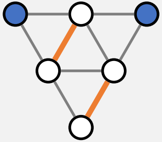
Review

Monomer-dimer system

Prob of matching $\propto$

$$\prod_{e \in M} \underset{\uparrow}{\lambda_e} \cdot \prod_{v \not\sim M} \underset{\uparrow}{z_v}$$

dimer monomer



- Assume wlog $z_v = 1$
- [Jerrum-Sinclair]: Metropolis mixes in

$$\text{poly}(n, \lambda_{\max})$$

- [Jerrum-Sinclair-Vigoda]: on bipartite graphs, can sample in

$$\text{poly}(n, \log \lambda_{\max})$$

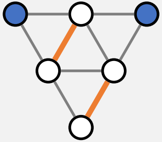
Review

Monomer-dimer system

Prob of matching $\propto$

$$\prod_{e \in M} \underset{\uparrow}{\lambda_e} \cdot \prod_{v \not\sim M} \underset{\uparrow}{z_v}$$

dimer monomer



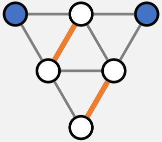
- Assume wlog $z_v = 1$
- [Jerrum-Sinclair]: Metropolis mixes in
$$\text{poly}(n, \lambda_{\text{max}})$$
- [Jerrum-Sinclair-Vigoda]: on bipartite graphs, can sample in
$$\text{poly}(n, \log \lambda_{\text{max}})$$
- Open: non-bipartite graphs

Review

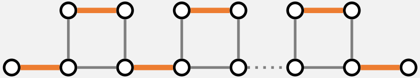
Monomer-dimer system

Prob of matching $\propto$

$$\prod_{e\in M}\underset{\overset{\uparrow}{\text{dimer}}}{\lambda_e}\cdot\prod_{v\not\in M}\underset{\overset{\uparrow}{\text{monomer}}}{z_v}$$



Challenging example: chain of boxes



- Assume wlog $z_v = 1$
- [Jerrum-Sinclair]: Metropolis mixes in
  $$\text{poly}(n, \lambda_{\max})$$
- [Jerrum-Sinclair-Vigoda]: on bipartite graphs, can sample in
  $$\text{poly}(n, \log\lambda_{\max})$$
- Open: non-bipartite graphs

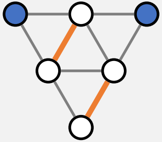
Review

Monomer-dimer system

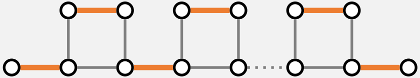
Prob of matching $\propto$

$$\prod_{e\in M}\lambda_e \cdot \prod_{v\not\in M} z_v$$

dimer monomer



Challenging example: chain of boxes



- Assume wlog $z_v = 1$
- [Jerrum-Sinclair]: Metropolis mixes in
  $$\text{poly}(n, \lambda_{\max})$$
- [Jerrum-Sinclair-Vigoda]: on bipartite graphs, can sample in
  $$\text{poly}(n, \log \lambda_{\max})$$
- Open: non-bipartite graphs

- To resolve we reweigh:
  $$\mu(M) \propto \frac{\lambda^M}{\lambda(\Omega_{\text{monomers}(M)})}$$

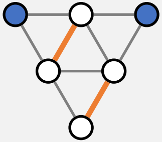
Review

Monomer-dimer system

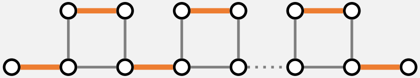
Prob of matching $\propto$

$$\prod_{e\in M}\lambda_e \cdot \prod_{v\not\sim M} z_v$$

dimer monomer



Challenging example: chain of boxes



- Assume wlog $z_v = 1$
- [Jerrum-Sinclair]: Metropolis mixes in
$$\text{poly}(n, \lambda_{\max})$$
- [Jerrum-Sinclair-Vigoda]: on bipartite graphs, can sample in
$$\text{poly}(n, \log \lambda_{\max})$$
- Open: non-bipartite graphs

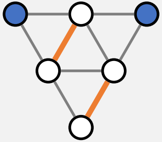
- To resolve we reweigh:
$$\mu(M) \propto \frac{\lambda^M}{\lambda(\Omega_{\text{monomers}(M)})}$$
- Metropolis on perfect and near-perfect $\propto \mu$ mixes fast.

Review

Monomer-dimer system

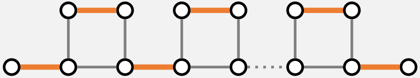
Prob of matching $\propto$

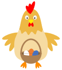
$$\prod_{e\in M}\underset{\text{dimer}}{\underset{\uparrow}{\lambda_e}}\cdot\prod_{v\not\in M}\underset{\text{monomer}}{\underset{\uparrow}{z_v}}$$



- Assume wlog $z_v = 1$
- [Jerrum-Sinclair]: Metropolis mixes in $$\text{poly}(n, \lambda_{\text{max}})$$
- [Jerrum-Sinclair-Vigoda]: on bipartite graphs, can sample in $$\text{poly}(n, \log \lambda_{\text{max}})$$
- Open: non-bipartite graphs

Challenging example: chain of boxes



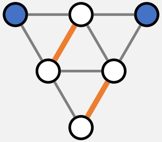
- To resolve we reweigh: $$\mu(M) \propto \frac{\lambda^M}{\lambda(\Omega_{\text{monomers}(M)})}$$
- Metropolis on perfect and near-perfect $\propto \mu$ mixes fast.
- The  problem: how to estimate $\lambda(\Omega_S)$?

Review

Monomer-dimer system

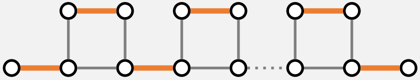
Prob of matching $\propto$

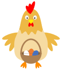
$$\prod_{e\in M}\underset{\substack{\uparrow\\ \text{dimer}}}{\lambda_e}\cdot\prod_{v\not\in M}\underset{\substack{\uparrow\\ \text{monomer}}}{z_v}$$



- Assume wlog $z_v = 1$
- [Jerrum-Sinclair]: Metropolis mixes in
$$\text{poly}(n, \lambda_{\max})$$
- [Jerrum-Sinclair-Vigoda]: on bipartite graphs, can sample in
$$\text{poly}(n, \log\lambda_{\max})$$
- Open: non-bipartite graphs

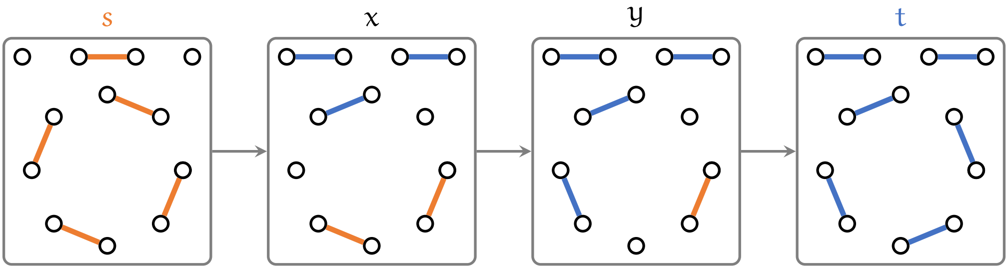
Challenging example: chain of boxes

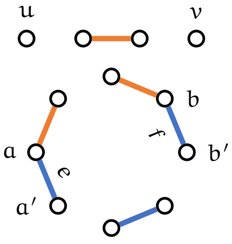


- To resolve we reweigh:
$$\mu(M) \propto \frac{\lambda^{M}}{\lambda(\Omega_{\text{monomers}(M)})}$$
- Metropolis on perfect and near-perfect $\propto \mu$ mixes fast.
- The  problem: how to estimate $\lambda(\Omega_S)$?
- Solution: slowly change λ

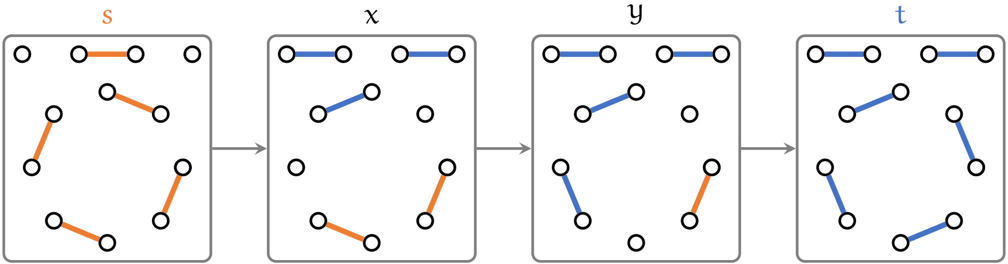
↑

official name: cooling schedule

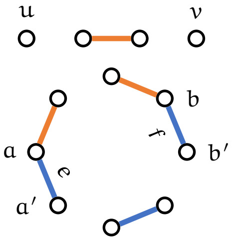




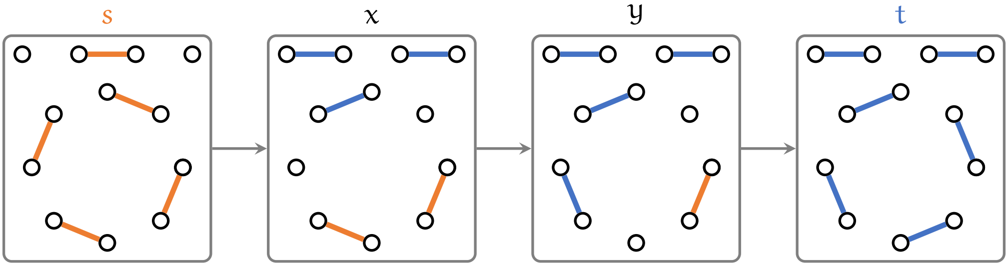
$s \oplus t \oplus x$



$$\mathsf{enc} = s \oplus t \oplus x - e - f$$



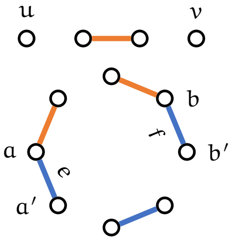
$s \oplus t \oplus x$



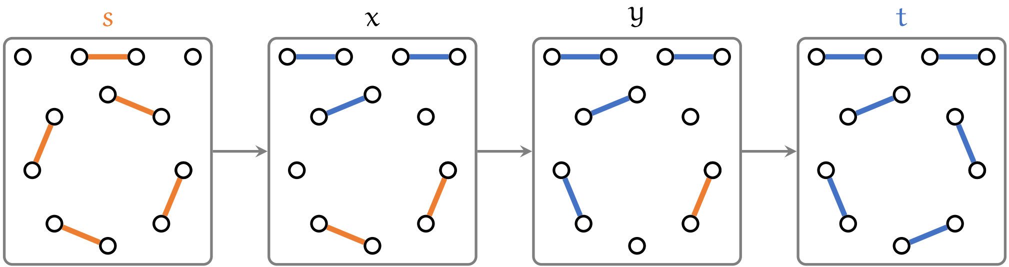
$\text{enc} = s \oplus t \oplus x - e - f$

- Note that $\lambda^{s}\lambda^{t} = \lambda_e\lambda_f\lambda^{x}\lambda^{\text{enc}}$. Let e's endpoints be a, a' and f's endpoints be b, b'. Prove: ← via injective map

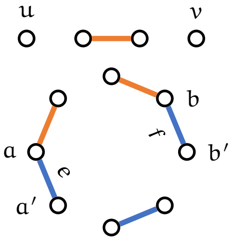
$$\lambda(\underset{\overset{\uparrow}{t}}{\Omega_{\emptyset}})\lambda(\underset{\overset{\uparrow}{s}}{\Omega_{\{u,v\}}}) \geqslant \frac{1}{\text{poly}(n)} \cdot \lambda_e\lambda_f\lambda(\underset{\overset{\uparrow}{x}}{\Omega_{\{a,b\}}})\lambda(\underset{\overset{\uparrow}{\text{enc}}}{\Omega_{\{u,v,a',b'\}}})$$



$s \oplus t \oplus x$



$\mathsf{enc} = s \oplus t \oplus x - e - f$



$s \oplus t \oplus x$

- Note that $\lambda^{s}\lambda^{t} = \lambda_e\lambda_f\lambda^{x}\lambda^{\mathsf{enc}}$. Let e's endpoints be a, a' and f's endpoints be b, b'. Prove: ← via injective map

$$\lambda(\Omega_{\underset{\substack{\uparrow\\ t}}{\emptyset}})\lambda(\Omega_{\underset{\substack{\uparrow\\ s}}{\{u,v\}}}) \geqslant \frac{1}{\mathsf{poly}(n)} \cdot \lambda_e\lambda_f\lambda(\Omega_{\underset{\substack{\uparrow\\ x}}{\{a,b\}}})\lambda(\Omega_{\underset{\substack{\uparrow\\ \mathsf{enc}}}{\{u,v,a',b'\}}})$$

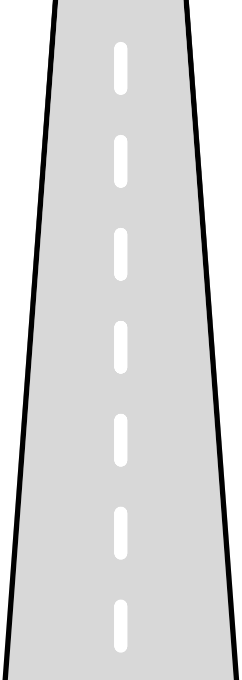
- Thus $\mu(s)\mu(t) \leqslant \mathsf{poly}(n) \cdot \mu(x)\mu(\mathsf{enc})$. Similar ineqs yield $\mu(s)\mu(t) \leqslant \mathsf{poly}(n) \cdot \mu(y)\mu(\mathsf{enc})$. So cong $\leqslant \mathsf{poly}(n)$.

Down-Up Walks

- Matroids
- Down and up kernels
- Simplicial complex

High-Dimensional Expansion

- Local-to-global

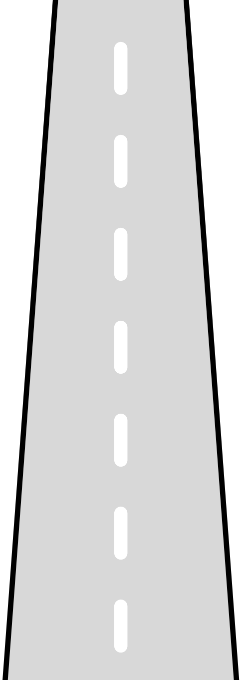


Down-Up Walks

- Matroids
- Down and up kernels
- Simplicial complex

High-Dimensional Expansion

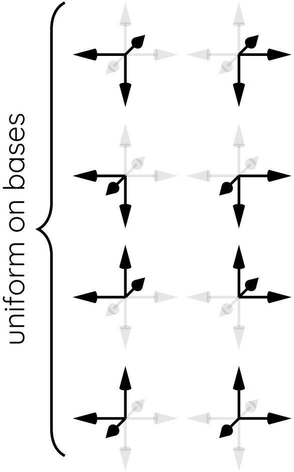
- Local-to-global



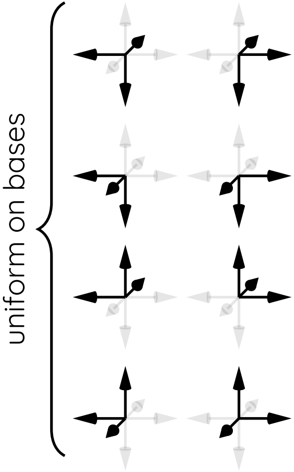
Sample S with $\mathbb{P}[S] \propto \mu(S)$:

$$\underbrace{\mu : \binom{U}{k} \to \mathbb{R}_{\geqslant 0}}_{\text{weighted hypergraph}}$$

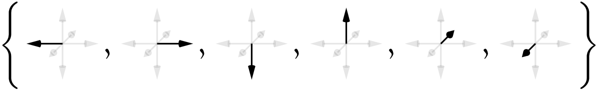
Example: matroids



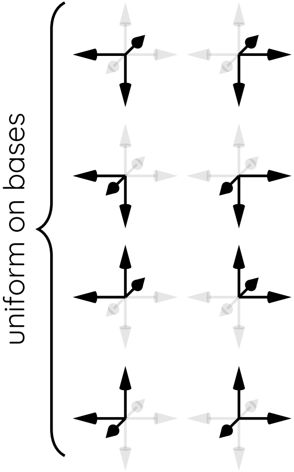
Example: matroids

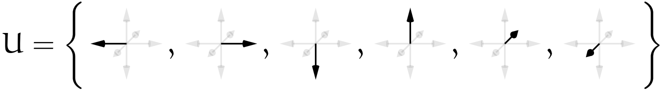


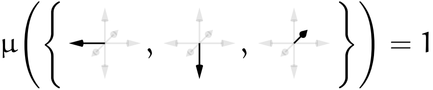
$u =$



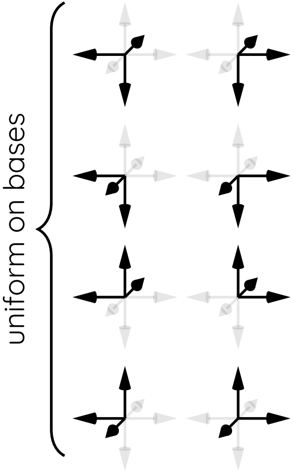
Example: matroids

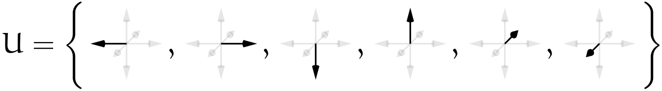


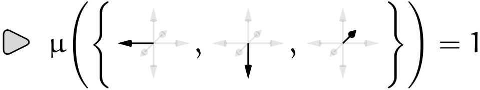


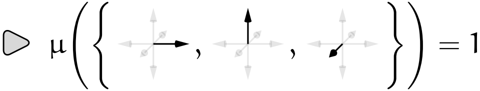
- 

Example: matroids

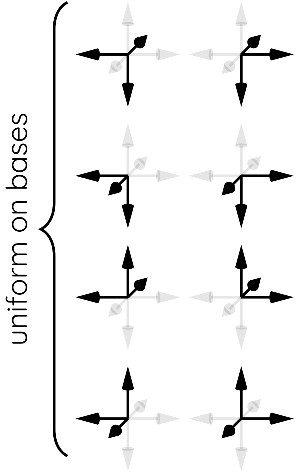


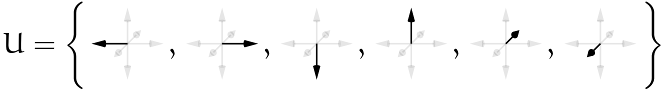


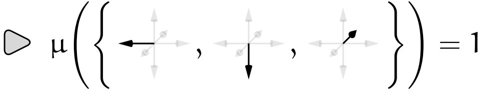
- 

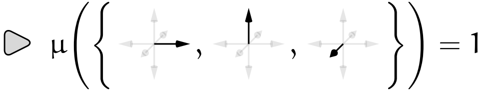
- 

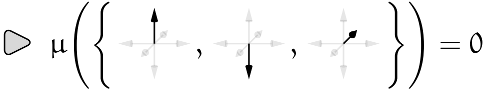
Example: matroids



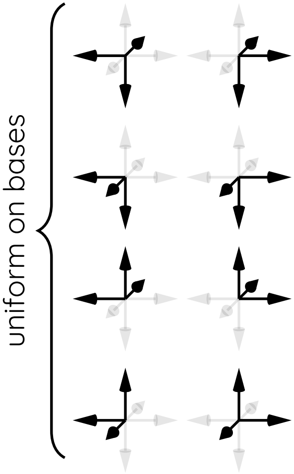


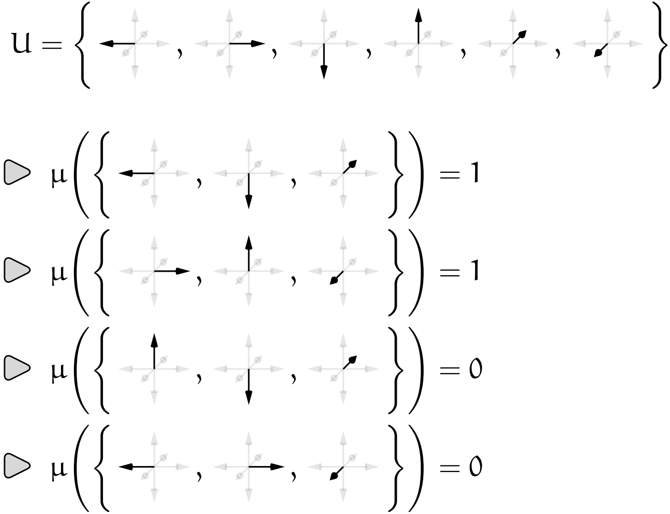






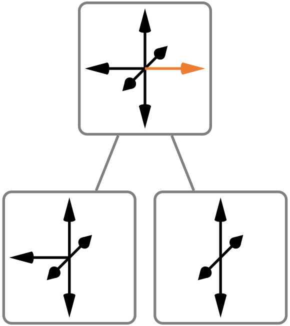
Example: matroids



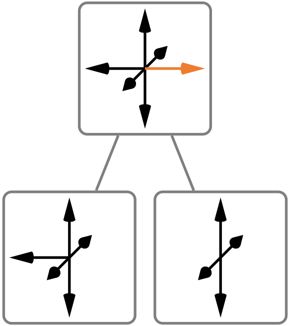


- Note: matroids larger than linear matroids. Results apply to all.

- Note: matroids larger than linear matroids. Results apply to all.
- Matroids are self-reducible:



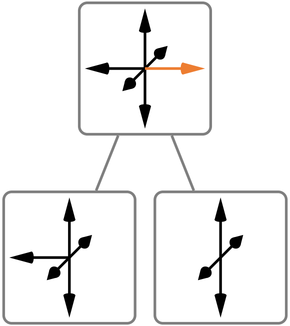
- Note: matroids larger than linear matroids. Results apply to all.
- Matroids are self-reducible:



Example: hypercube

Matroid on $\{\pm\mathbb{1}_1, \ldots, \pm\mathbb{1}_n\}$

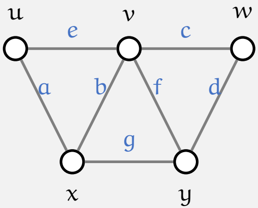
- Note: matroids larger than linear matroids. Results apply to all.
- Matroids are self-reducible:



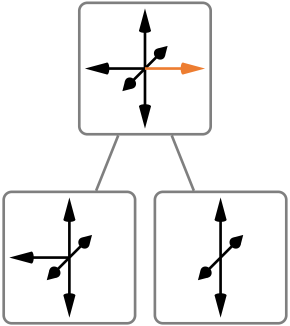
Example: hypercube

Matroid on $\{\pm\mathbb{1}_1, \ldots, \pm\mathbb{1}_n\}$

Example: spanning trees



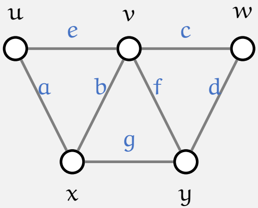
- Note: matroids larger than linear matroids. Results apply to all.
- Matroids are self-reducible:



Example: hypercube

Matroid on $\{\pm\mathbb{1}_1,\dots,\pm\mathbb{1}_n\}$

Example: spanning trees



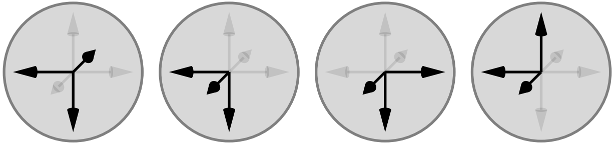
$$\begin{array}{c|ccccccc|} & a & b & c & d & e & f & g \\ u & +1 & 0 & 0 & 0 & +1 & 0 & 0 \\ v & 0 & -1 & +1 & 0 & -1 & -1 & 0 \\ w & 0 & 0 & -1 & +1 & 0 & 0 & 0 \\ x & -1 & +1 & 0 & 0 & 0 & 0 & -1 \\ y & 0 & 0 & 0 & -1 & 0 & +1 & +1 \end{array}$$

vertex-edge adj matrix

Down-up walks

Random walk:

1. Drop element u.a.r.
2. Add element with prob. $\propto \mu$(resulting set).

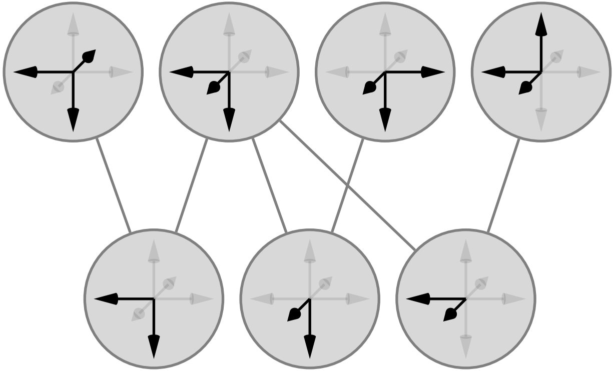


$\mu(S) = \mathbb{1}[S \text{ is basis}]$.

Down-up walks

Random walk:

1. Drop element u.a.r.
2. Add element with prob. $\propto \mu$(resulting set).

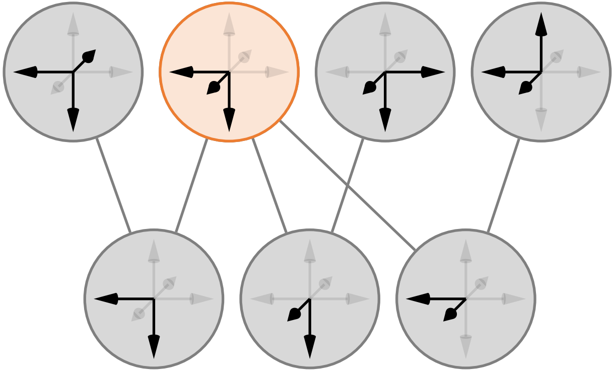


$\mu(S) = \mathbb{1}[S \text{ is basis}]$.

Down-up walks

Random walk:

1. Drop element u.a.r.
2. Add element with prob. $\propto \mu(\text{resulting set})$.

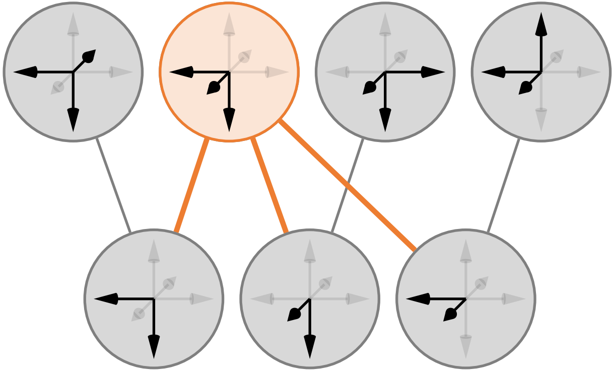


$$\mu(S) = \mathbb{1}[S \text{ is basis}].$$

Down-up walks

Random walk:

1. Drop element u.a.r.
2. Add element with prob. $\propto \mu$(resulting set).

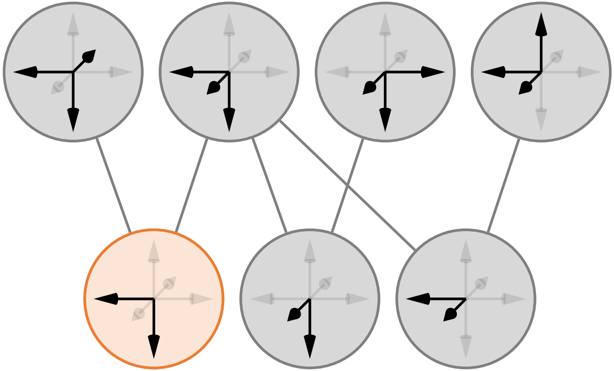


$\mu(S) = \mathbb{1}[S \text{ is basis}]$.

Down-up walks

Random walk:

1. Drop element u.a.r.
2. Add element with prob. $\propto \mu$(resulting set).

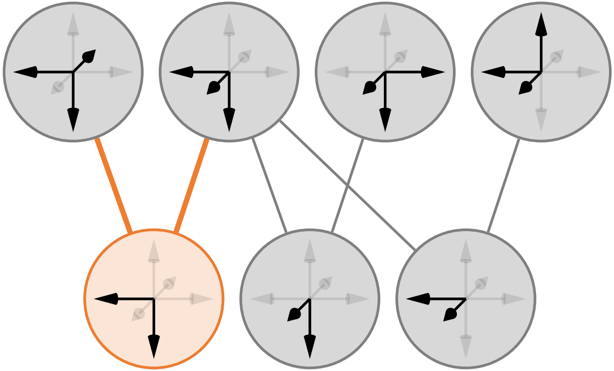


$\mu(S) = \mathbb{1}[S \text{ is basis}]$.

Down-up walks

Random walk:

1. Drop element u.a.r.
2. Add element with prob. $\propto \mu(\text{resulting set})$.

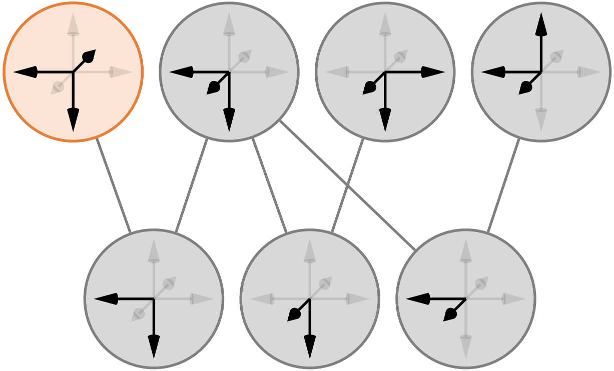


$\mu(S) = \mathbb{1}[S \text{ is basis}]$.

Down-up walks

Random walk:

1. Drop element u.a.r.
2. Add element with prob. $\propto \mu$(resulting set).



$\mu(S) = \mathbb{1}[S \text{ is basis}]$.

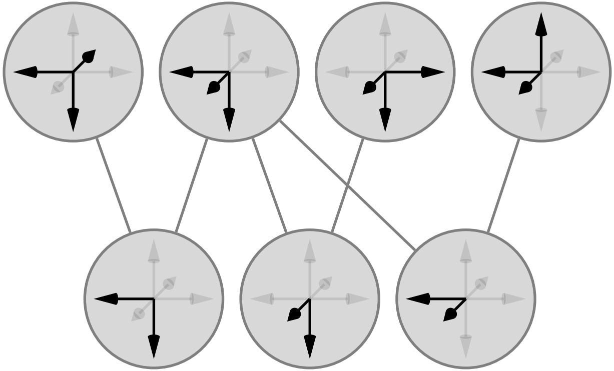
Down-up walks

Random walk:

1. Drop element u.a.r.
2. Add element with prob. $\propto \mu$(resulting set).

[A-Liu-OveisGharan-Vinzant'19]

$t_{\text{mix}} = O(k^2 \log n) \longleftarrow$ Poincaré



$\mu(S) = \mathbb{1}[S$ is basis$]$.

Down-up walks

Random walk:

① Drop element u.a.r.

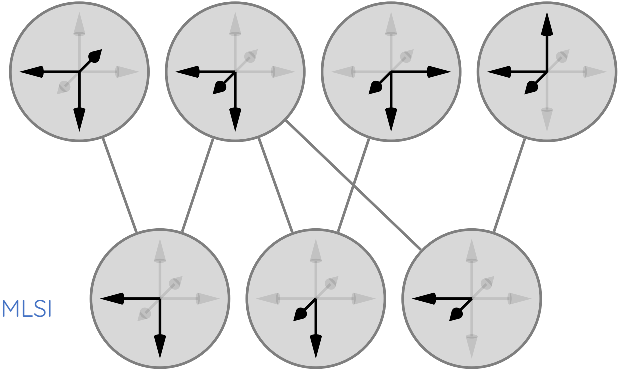
② Add element with prob. $\propto \mu$(resulting set).

[A-Liu-OveisGharan-Vinzant'19]

$t_{\text{mix}} = O(k^2 \log n) \longleftarrow$ Poincaré

[Cryan-Guo-Mousa'20]

$t_{\text{mix}} = O(k(\log k + \log\log n)) \longleftarrow$ MLSI



$\mu(S) = \mathbb{1}[S \text{ is basis}].$

Down-up walks

Random walk:

1. Drop element u.a.r.
2. Add element with prob. $\propto \mu(\text{resulting set})$.

[A-Liu-OveisGharan-Vinzant'19]

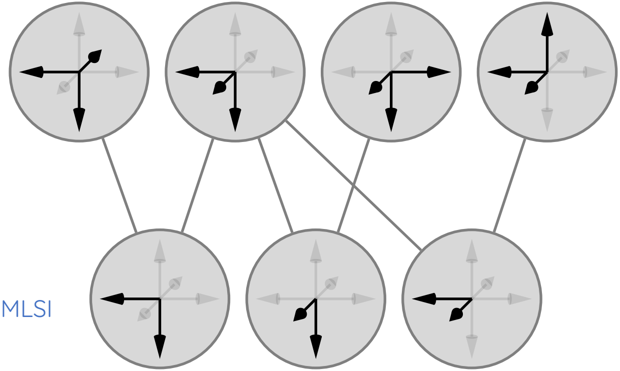
$t_{\text{mix}} = O(k^2 \log n) \longleftarrow$ Poincaré

[Cryan-Guo-Mousa'20]

$t_{\text{mix}} = O(k(\log k + \log\log n)) \longleftarrow$ MLSI

[A-Liu-OveisGharan-Vinzant-Vuong'21]

$t_{\text{mix}} = O(k \log k) \longleftarrow$ exchange



$\mu(S) = \mathbb{1}[S \text{ is basis}]$.

- Down-up walk is of the form

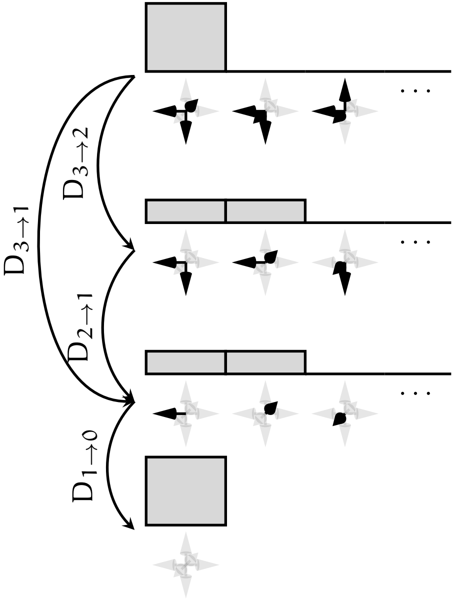
$$\underset{\substack{\nearrow \\ \text{down}}}{N}\,\underset{\substack{\nwarrow \\ \text{up}}}{N^\circ}$$

- Down-up walk is of the form

$$\underset{\text{down}}{N}\underset{\text{up}}{N^{\circ}}$$

Down kernels

$D_{k\to\ell}$ is Markov kernel from $\Omega = \binom{u}{k}$ to $\Omega' = \binom{u}{\ell}$: drop $k-\ell$ u.r. elements.



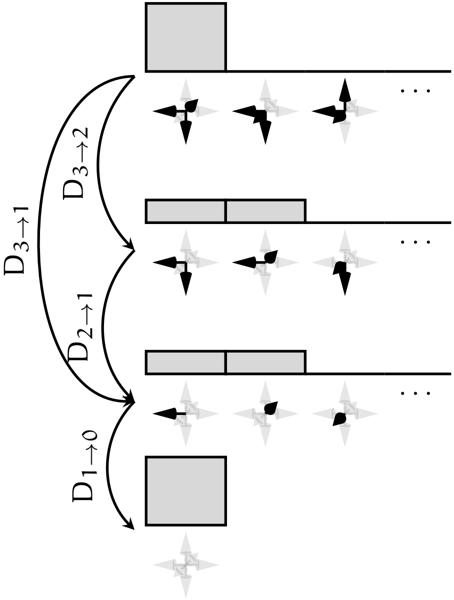
- Down-up walk is of the form

$$\underset{\substack{\nearrow\\ \text{down}}}{N}\underset{\substack{\nwarrow\\ \text{up}}}{N^{\circ}}$$

Down kernels

$D_{k\to\ell}$ is Markov kernel from $\Omega = \binom{u}{k}$ to $\Omega' = \binom{u}{\ell}$: drop $k-\ell$ u.r. elements.

- Distribution on level ℓ: $\mu D_{k\to\ell}$



- Down-up walk is of the form

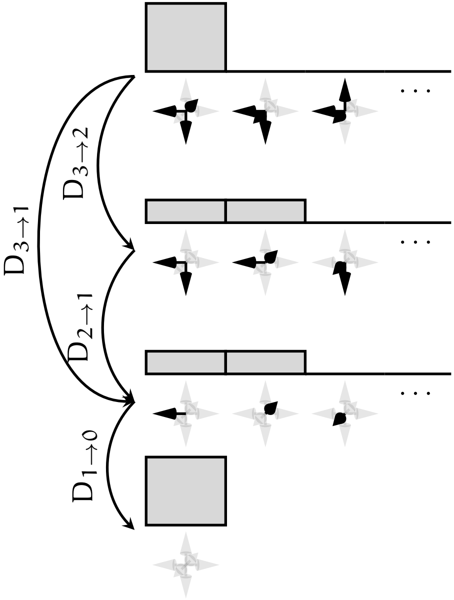
$$\underset{\text{down}}{N}\,\underset{\text{up}}{N^{\circ}}$$

Down kernels

$D_{k\to\ell}$ is Markov kernel from $\Omega = \binom{u}{k}$ to $\Omega' = \binom{u}{\ell}$: drop $k-\ell$ u.r. elements.

- Distribution on level ℓ: $\mu D_{k\to\ell}$
- Time-reversals are up kernels:

$$U_{\ell\to k} = D^{\circ}_{k\to\ell}$$



- Down-up walk is of the form

$$\underset{\text{down}}{N}\,\underset{\text{up}}{N^{\circ}}$$

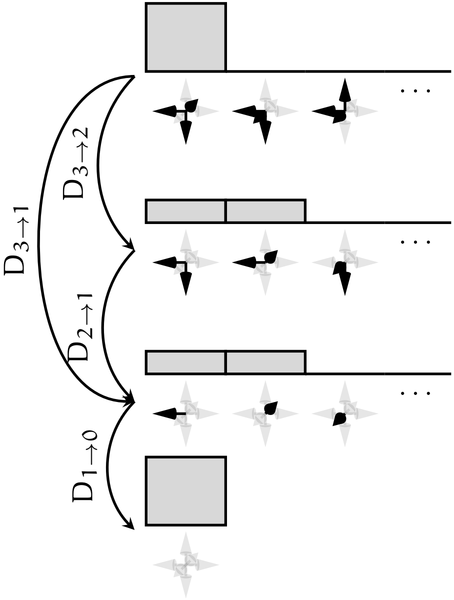
Down kernels

$D_{k\to\ell}$ is Markov kernel from $\Omega = \binom{u}{k}$ to $\Omega' = \binom{u}{\ell}$: drop $k-\ell$ u.r. elements.

- Distribution on level ℓ: $\mu D_{k\to\ell}$
- Time-reversals are up kernels:

$$U_{\ell\to k} = D^{\circ}_{k\to\ell}$$

- $k \leftrightarrow \ell$ down-up walk: $D_{k\to\ell}U_{\ell\to k}$



- Down-up walk is of the form

$$\underset{\text{down}}{N}\underset{\text{up}}{N^{\circ}}$$

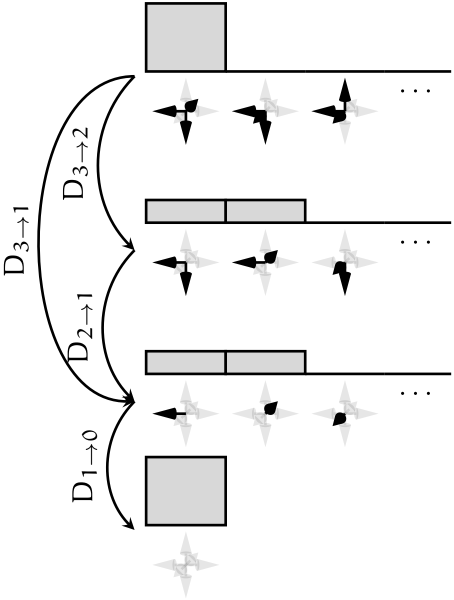
Down kernels

$D_{k\to\ell}$ is Markov kernel from $\Omega = \binom{U}{k}$ to $\Omega' = \binom{U}{\ell}$: drop $k-\ell$ u.r. elements.

- Distribution on level ℓ: $\mu D_{k\to\ell}$
- Time-reversals are up kernels:

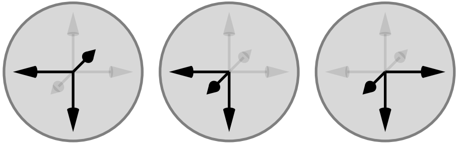
$$U_{\ell\to k} = D^{\circ}_{k\to\ell}$$

- $k \leftrightarrow \ell$ down-up walk: $D_{k\to\ell}U_{\ell\to k}$
- Algorithmically useful: $\ell = k-1$ or more generally $\ell = k - O(1)$



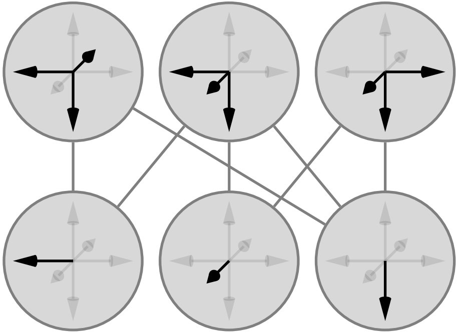
HDX recipe

1. Dist on $\binom{U}{k}$ ← simplicial complex
2. Local: show $\mathcal{D}_\phi$ contraction for $D_{k\to 1}$
3. Global: transfer contraction for $D_{k\to 1}$ to $D_{k\to k-1}$



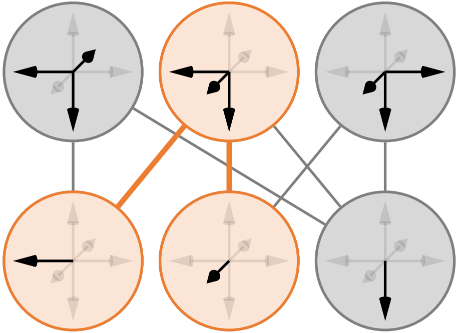
HDX recipe

1. Dist on $\binom{U}{k}$ ← simplicial complex
2. Local: show $\mathcal{D}_\phi$ contraction for $D_{k\to 1}$
3. Global: transfer contraction for $D_{k\to 1}$ to $D_{k\to k-1}$



HDX recipe

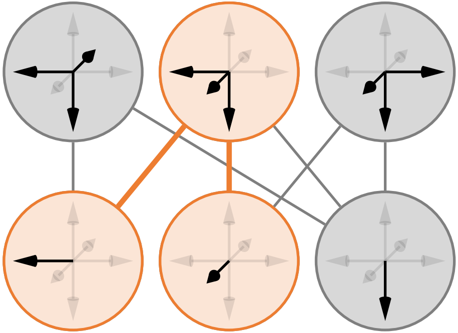
1. Dist on $\binom{U}{k}$ ← simplicial complex
2. Local: show $\mathcal{D}_\phi$ contraction for $D_{k\to 1}$
3. Global: transfer contraction for $D_{k\to 1}$ to $D_{k\to k-1}$



HDX recipe

1. Dist on $\binom{U}{k}$ ← simplicial complex
2. Local: show $\mathcal{D}_\phi$ contraction for $D_{k\to1}$
3. Global: transfer contraction for $D_{k\to1}$ to $D_{k\to k-1}$

▷ How to convert to simplicial complex?

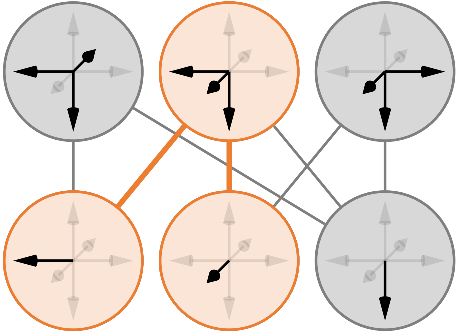


HDX recipe

① Dist on $\binom{U}{k}$ ← simplicial complex

② Local: show $\mathcal{D}_\phi$ contraction for $D_{k \to 1}$

③ Global: transfer contraction for $D_{k \to 1}$ to $D_{k \to k-1}$

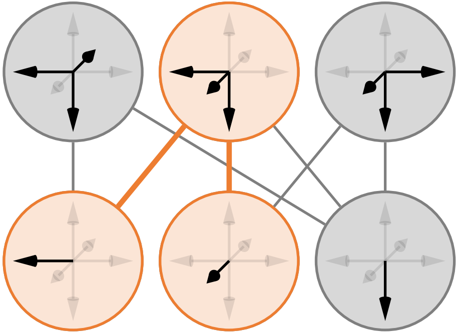


- How to convert to simplicial complex?
- For product spaces $\Omega_1 \times \cdots \times \Omega_n$:

$$\binom{\Omega_1 \sqcup \cdots \sqcup \Omega_n}{n}$$

HDX recipe

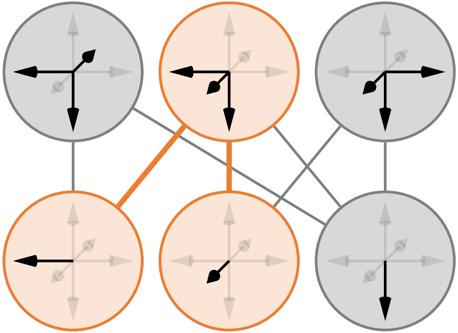
1. Dist on $\binom{U}{k}$ ← simplicial complex
2. Local: show $\mathcal{D}_\phi$ contraction for $D_{k\to 1}$
3. Global: transfer contraction for $D_{k\to 1}$ to $D_{k\to k-1}$



- How to convert to simplicial complex?
- For product spaces $\Omega_1 \times \cdots \times \Omega_n$:
  $$\binom{\Omega_1 \sqcup \cdots \sqcup \Omega_n}{n}$$
- Put zero mass on invalid sets.

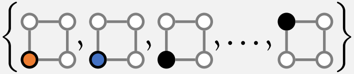
HDX recipe

1. Dist on $\binom{\mathcal{U}}{k}$ ← simplicial complex
2. Local: show $\mathcal{D}_\phi$ contraction for $D_{k\to 1}$
3. Global: transfer contraction for $D_{k\to 1}$ to $D_{k\to k-1}$



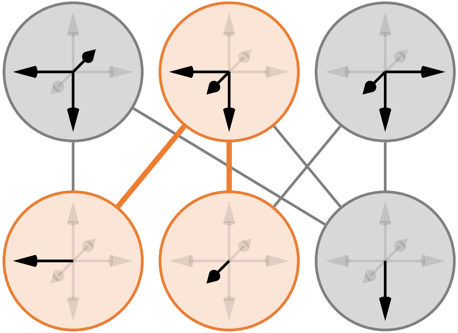
- How to convert to simplicial complex?
- For product spaces $\Omega_1 \times \cdots \times \Omega_n$: $\binom{\Omega_1 \sqcup \cdots \sqcup \Omega_n}{n}$
- Put zero mass on invalid sets.

Example: coloring

$\mathcal{U} =$ 

HDX recipe

1. Dist on $\binom{\mathcal{U}}{k}$ ← simplicial complex
2. Local: show $\mathcal{D}_\phi$ contraction for $D_{k\to 1}$
3. Global: transfer contraction for $D_{k\to 1}$ to $D_{k\to k-1}$

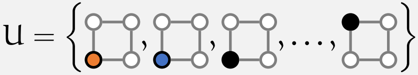


- How to convert to simplicial complex?
- For product spaces $\Omega_1 \times \cdots \times \Omega_n$:

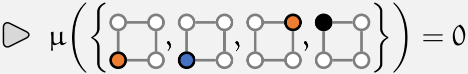
$$\binom{\Omega_1 \sqcup \cdots \sqcup \Omega_n}{n}$$

- Put zero mass on invalid sets.

Example: coloring



$\mathcal{U} = \{ \ , \ , \ , \dots, \ \}$



$\mu(\{ \ , \ , \ , \ \}) = 0$

HDX recipe

① Dist on $\binom{\mathcal{U}}{k}$ ← simplicial complex

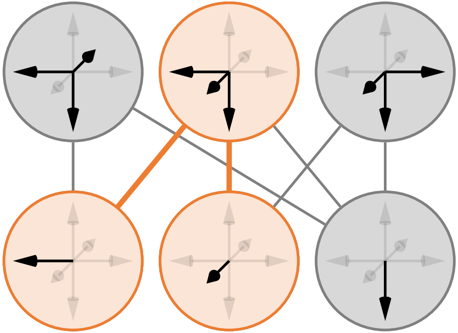
② Local: show $\mathcal{D}_\phi$ contraction for $D_{k\to 1}$

③ Global: transfer contraction for $D_{k\to 1}$ to $D_{k\to k-1}$

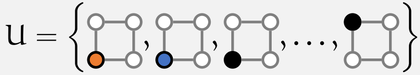
- How to convert to simplicial complex?
- For product spaces $\Omega_1 \times \cdots \times \Omega_n$:

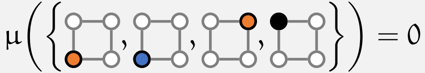
$$\binom{\Omega_1 \sqcup \cdots \sqcup \Omega_n}{n}$$

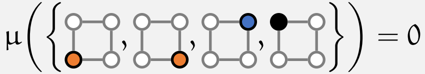
- Put zero mass on invalid sets.



Example: coloring



- 

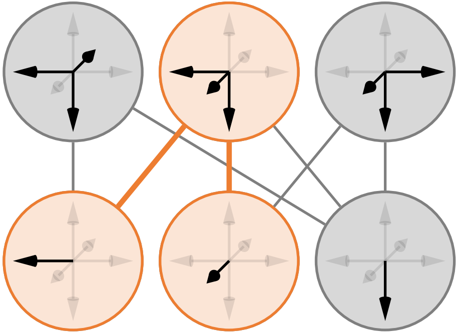
- 

HDX recipe

① Dist on $\binom{U}{k}$ ← simplicial complex

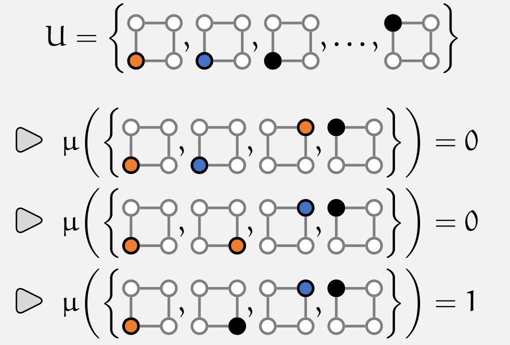
② Local: show $\mathcal{D}_\phi$ contraction for $D_{k\to1}$

③ Global: transfer contraction for $D_{k\to1}$ to $D_{k\to k-1}$



- How to convert to simplicial complex?
- For product spaces $\Omega_1 \times \cdots \times \Omega_n$:
  $$\binom{\Omega_1 \sqcup \cdots \sqcup \Omega_n}{n}$$
- Put zero mass on invalid sets.

Example: coloring

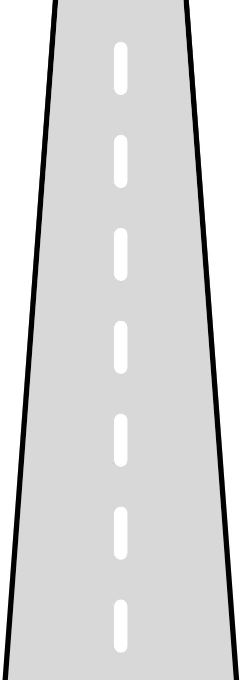


Down-Up Walks

- Matroids
- Down and up kernels
- Simplicial complex

High-Dimensional Expansion

- Local-to-global

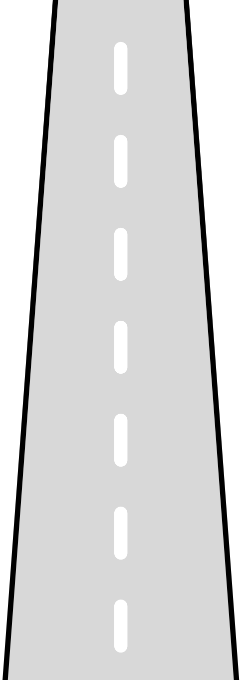


Down-Up Walks

- Matroids
- Down and up kernels
- Simplicial complex

High-Dimensional Expansion

- Local-to-global



- Goal: analyze $\mathcal{D}_\phi$ contraction for the Markov kernels $D_{k\to\ell}$

- Goal: analyze $\mathcal{D}_\phi$ contraction for the Markov kernels $D_{k \to \ell}$
- From data processing, we know

  $$\mathcal{D}_\phi(\nu \parallel \mu) - \mathcal{D}_\phi(\nu D_{k\to\ell} \parallel \mu D_{k\to\ell})$$

  equals

  $$\mathbb{E}_{T \sim \mu D_{k\to\ell}}\left[\mathrm{Ent}^{\phi}_{U_{\ell\to k}(T,\cdot)}\left[\frac{\nu}{\mu}\right]\right]$$

- Goal: analyze $\mathcal{D}_\phi$ contraction for the Markov kernels $D_{k\to\ell}$
- From data processing, we know

  $\mathcal{D}_\phi(\nu \parallel \mu) - \mathcal{D}_\phi(\nu D_{k\to\ell} \parallel \mu D_{k\to\ell})$

  equals

$$\mathbb{E}_{T\sim\mu D_{k\to\ell}}\left[\mathrm{Ent}^{\phi}_{U_{\ell\to k}(T,\cdot)}\left[\frac{\nu}{\mu}\right]\right]$$

Conditionals

For a set $|T| = \ell$ and dist μ on $\binom{U}{k}$, we call $U_{\ell\to k}(T,\cdot)$ the conditional dist:

$$\mathrm{dist}_{S\sim\mu}(S \mid T \subseteq S)$$

- Goal: analyze $\mathcal{D}_\phi$ contraction for the Markov kernels $D_{k\to\ell}$
- From data processing, we know

  $\mathcal{D}_\phi(\nu \parallel \mu) - \mathcal{D}_\phi(\nu D_{k\to\ell} \parallel \mu D_{k\to\ell})$

  equals

$$\mathbb{E}_{T\sim\mu D_{k\to\ell}}\left[\mathrm{Ent}^{\phi}_{U_{\ell\to k}(T,\cdot)}\left[\frac{\nu}{\mu}\right]\right]$$

Conditionals

For a set $|T| = \ell$ and dist μ on $\binom{U}{k}$, we call $U_{\ell\to k}(T,\cdot)$ the conditional dist:

$$\mathrm{dist}_{S\sim\mu}(S \mid T \subseteq S)$$

Links

Removing the known part T, we get links μ_T:

$$\mu_T = \mathrm{dist}_{S\sim\mu}(S - T \mid T \subseteq S)$$

- Goal: analyze $\mathcal{D}_\phi$ contraction for the Markov kernels $D_{k\to\ell}$
- From data processing, we know
$$\mathcal{D}_\phi(\nu \parallel \mu) - \mathcal{D}_\phi(\nu D_{k\to\ell} \parallel \mu D_{k\to\ell})$$
equals
$$\mathbb{E}_{T\sim\mu D_{k\to\ell}}\left[\mathrm{Ent}^{\phi}_{U_{\ell\to k}(T,\cdot)}\left[\frac{\nu}{\mu}\right]\right]$$

Conditionals

For a set $|T| = \ell$ and dist μ on $\binom{U}{k}$, we call $U_{\ell\to k}(T,\cdot)$ the conditional dist:

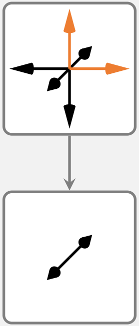
$$\mathrm{dist}_{S\sim\mu}(S \mid T \subseteq S)$$

Links

Removing the known part T, we get links μ_T:

$$\mu_T = \mathrm{dist}_{S\sim\mu}(S - T \mid T \subseteq S)$$

Example: matroids



Conditionals: project all in $U - T$ on $\mathrm{span}(T)^\perp$.

Links: remove T and set vector space to $\mathrm{span}(T)^\perp$.

Links of matroids are matroids. ☺

Local-to-global

Theorem **[Alev-Lau, . . .]**

Suppose for each $|T| = t$, we know $D_{k-t\to 1}$ contracts $\mathcal{D}_\phi(\cdot \parallel \mu_T)$ by $1-\rho_t$. Then

$$\frac{\mathcal{D}_\phi(\nu D_{k\to\ell} \parallel \mu D_{k\to\ell})}{\mathcal{D}_\phi(\nu \parallel \mu)} \leqslant 1-\rho,$$

where

$$\rho = \rho_0\rho_1\cdots\rho_{l-1}.$$

Local-to-global

Theorem **[Alev-Lau, . . .]**

Suppose for each $|T| = t$, we know $D_{k-t\to 1}$ contracts $\mathcal{D}_\phi(\cdot \,\|\, \mu_T)$ by $1-\rho_t$. Then

$$\frac{\mathcal{D}_\phi(\nu D_{k\to\ell} \,\|\, \mu D_{k\to\ell})}{\mathcal{D}_\phi(\nu \,\|\, \mu)} \leqslant 1 - \rho,$$

where

$$\rho = \rho_0 \rho_1 \cdots \rho_{l-1}.$$

- For matroids and $\mathcal{D}_\phi \in \{\chi^2, \mathcal{D}_{KL}\}$:

$$\rho_0 \geqslant 1 - \frac{1}{k}$$

Local-to-global

Theorem [Alev-Lau, . . .]

Suppose for each $|T| = t$, we know $D_{k-t\to 1}$ contracts $\mathcal{D}_\phi(\cdot \parallel \mu_T)$ by $1-\rho_t$. Then

$$\frac{\mathcal{D}_\phi(\nu D_{k\to\ell} \parallel \mu D_{k\to\ell})}{\mathcal{D}_\phi(\nu \parallel \mu)} \leqslant 1 - \rho,$$

where

$$\rho = \rho_0 \rho_1 \cdots \rho_{l-1}.$$

- For matroids and $\mathcal{D}_\phi \in \{\chi^2, \mathcal{D}_{KL}\}$:
$$\rho_0 \geqslant 1 - \frac{1}{k}$$

- This automatically means
$$\rho_t \geqslant 1 - \frac{1}{k-t}.$$

Local-to-global

Theorem [Alev-Lau, . . .]

Suppose for each $|T| = t$, we know $D_{k-t\to 1}$ contracts $\mathcal{D}_\phi(\cdot \parallel \mu_T)$ by $1-\rho_t$. Then

$$\frac{\mathcal{D}_\phi(\nu D_{k\to\ell} \parallel \mu D_{k\to\ell})}{\mathcal{D}_\phi(\nu \parallel \mu)} \leqslant 1-\rho,$$

where

$$\rho = \rho_0\rho_1\cdots\rho_{l-1}.$$

- For matroids and $\mathcal{D}_\phi \in \{\chi^2, \mathcal{D}_{KL}\}$:
$$\rho_0 \geqslant 1-\frac{1}{k}$$

- This automatically means
$$\rho_t \geqslant 1-\frac{1}{k-t}.$$

- By local-to-global, for $D_{k\to\ell}$, $\rho \geqslant$

$$\left(1-\frac{1}{k}\right)\cdots\left(1-\frac{1}{k-\ell+1}\right) = \frac{k-\ell}{k}$$

Local-to-global

Theorem [Alev-Lau, . . .]

Suppose for each $|T| = t$, we know $D_{k-t\to 1}$ contracts $\mathcal{D}_\phi(\cdot \,\|\, \mu_T)$ by $1-\rho_t$. Then

$$\frac{\mathcal{D}_\phi(\nu D_{k\to\ell} \,\|\, \mu D_{k\to\ell})}{\mathcal{D}_\phi(\nu \,\|\, \mu)} \leqslant 1-\rho,$$

where

$$\rho = \rho_0\rho_1 \cdots \rho_{l-1}.$$

- For matroids and $\mathcal{D}_\phi \in \{\chi^2, \mathcal{D}_{KL}\}$:
  $$\rho_0 \geqslant 1 - \tfrac{1}{k}$$
- This automatically means
  $$\rho_t \geqslant 1 - \tfrac{1}{k-t}.$$
- By local-to-global, for $D_{k\to\ell}$, $\rho \geqslant$
  $$\left(1-\frac{1}{k}\right)\cdots\left(1-\frac{1}{k-\ell+1}\right) = \frac{k-\ell}{k}$$
- For $\ell = k-1$ (algorithmically relevant), we get $\rho \geqslant 1/k$. 🙂

Local-to-global

Theorem [Alev-Lau, . . .]

Suppose for each $|T| = t$, we know $D_{k-t\to 1}$ contracts $\mathcal{D}_\phi(\cdot \parallel \mu_T)$ by $1-\rho_t$. Then

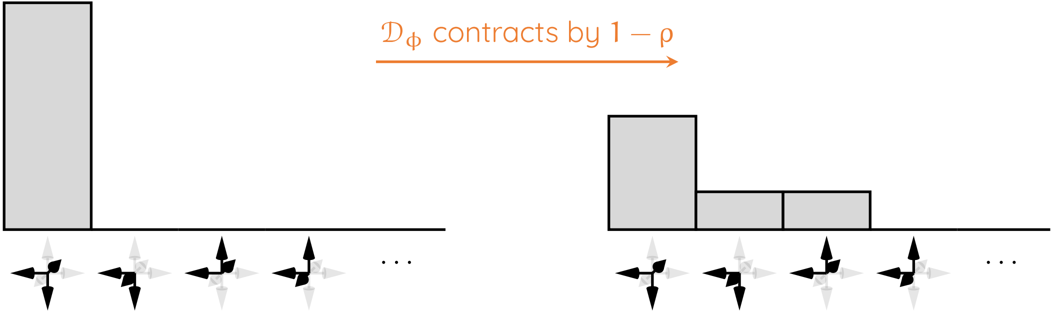
$$\frac{\mathcal{D}_\phi(\nu D_{k\to\ell} \parallel \mu D_{k\to\ell})}{\mathcal{D}_\phi(\nu \parallel \mu)} \leqslant 1-\rho,$$

where

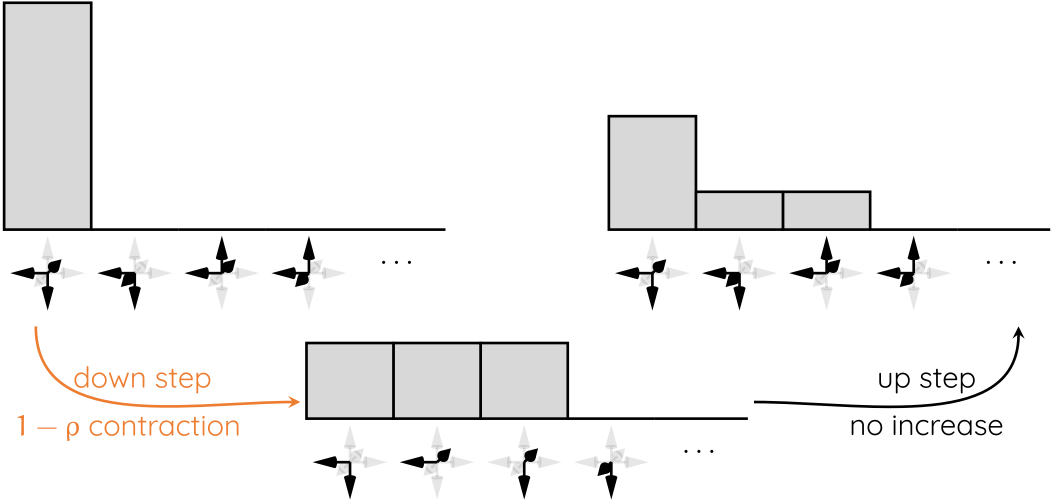
$$\rho = \rho_0\rho_1\cdots\rho_{l-1}.$$

- For matroids and $\mathcal{D}_\phi \in \{\chi^2, \mathcal{D}_{KL}\}$:
$$\rho_0 \geqslant 1 - \tfrac{1}{k}$$

- This automatically means
$$\rho_t \geqslant 1 - \tfrac{1}{k-t}.$$
- By local-to-global, for $D_{k\to\ell}$, $\rho \geqslant$
$$\left(1-\frac{1}{k}\right)\cdots\left(1-\frac{1}{k-\ell+1}\right) = \frac{k-\ell}{k}$$
- For $\ell = k-1$ (algorithmically relevant), we get $\rho \geqslant 1/k$. 🙂
- This trasnfers to down-up walk. By data processing, $U_{k-1\to k}$ cannot increase $\mathcal{D}_\phi$.



$\mathcal{D}_\phi$ contracts by $1 - \rho$



down step

1 − ρ contraction

up step

no increase

Proof:

- Consider $f = \nu/\mu$.

Proof:

- Consider $f = \nu/\mu$.
- We will track quantity:

$$z_T = \mathsf{Ent}^{\phi}_{U_{|T|\to k}(T,\cdot)}[f]$$

Proof:

- Consider $f = \nu/\mu$.
- We will track quantity:

$$z_T = \mathrm{Ent}^{\phi}_{U_{|T|\to k}(T,\cdot)}[f]$$

- Consider process $T_0 \subseteq T_1 \subseteq \ldots$: sample $S \sim \mu$, and u.a.r. permute its elements to be $S = \{e_1, \ldots, e_k\}$, and let $T_i = \{e_1, \ldots, e_i\}$.

Proof:

- Consider $f = \nu/\mu$.
- We will track quantity:

$$z_T = \mathsf{Ent}^{\phi}_{U_{|T|\to k}(T,\cdot)}[f]$$

- Consider process $T_0 \subseteq T_1 \subseteq \dots$: sample $S \sim \mu$, and u.a.r. permute its elements to be $S = \{e_1, \dots, e_k\}$, and let $T_i = \{e_1, \dots, e_i\}$.
- Claim: from assumption

$$\mathbb{E}\left[z_{T_{i+1}} \mid T_i\right] \geqslant \rho_i z_{T_i}.$$

Proof:

- Consider $f = \nu/\mu$.
- We will track quantity:

$$z_T = \mathrm{Ent}^{\phi}_{U_{|T|\to k}(T,\cdot)}[f]$$

- Consider process $T_0 \subseteq T_1 \subseteq \ldots$: sample $S \sim \mu$, and u.a.r. permute its elements to be $S = \{e_1, \ldots, e_k\}$, and let $T_i = \{e_1, \ldots, e_i\}$.
- Claim: from assumption

$$\mathbb{E}\left[z_{T_{i+1}} \mid T_i\right] \geqslant \rho_i z_{T_i}.$$

- Chaining the inequalities we get

$$\mathbb{E}[z_{T_\ell}] \geqslant \rho_0 \cdots \rho_{\ell-1} z_0.$$

Proof:

- Consider $f = \nu/\mu$.
- We will track quantity:

$$z_T = \mathsf{Ent}^{\phi}_{\mathcal{U}_{|T|\to k}(T,\cdot)}[f]$$

- Consider process $T_0 \subseteq T_1 \subseteq \ldots$: sample $S \sim \mu$, and u.a.r. permute its elements to be $S = \{e_1, \ldots, e_k\}$, and let $T_i = \{e_1, \ldots, e_i\}$.
- Claim: from assumption

$$\mathbb{E}\left[z_{T_{i+1}} \mid T_i\right] \geqslant \rho_i z_{T_i}.$$

- Chaining the inequalities we get

$$\mathbb{E}[z_{T_\ell}] \geqslant \rho_0 \cdots \rho_{\ell-1} z_0.$$

- This is the same as

$$\frac{\mathcal{D}_\phi(\nu D_{k\to\ell} \parallel \mu D_{k\to\ell})}{\mathcal{D}_\phi(\nu \parallel \mu)} \leqslant 1 - \rho,$$

where

$$\rho = \rho_0 \cdots \rho_{\ell-1}.$$