CS 263: Counting and Sampling

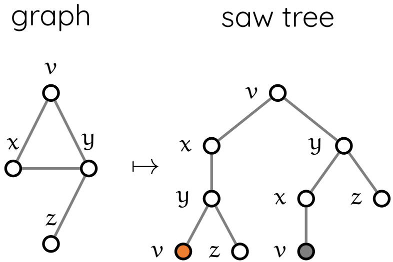
Nima Anari

Stanford University

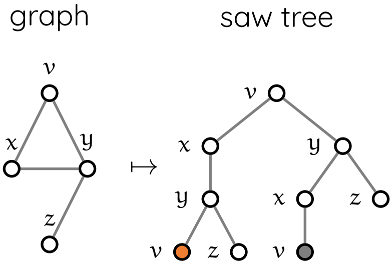
slides for

Universality of HDX

Review



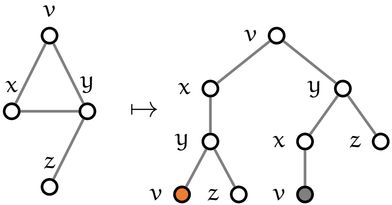
Review



- Root marginals are the same.

Review

graph saw tree



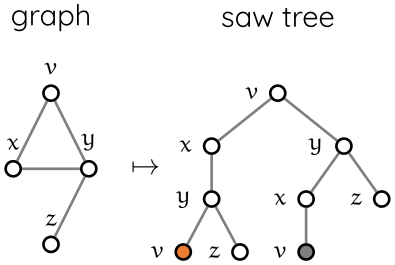
- Root marginals are the same.
- Weak spatial mixing:

$$d_{\mathrm{TV}}(\text{root} \mid \sigma, \text{root} \mid \sigma') \to 0$$

as the following goes to ∞:

$$\min\{d(\text{root}, u) \mid u \in S\}.$$

Review



- Root marginals are the same.
- Weak spatial mixing:

$$d_{TV}(\text{root} \mid \sigma, \text{root} \mid \sigma') \to 0$$

as the following goes to ∞:

$$\min\{d(\text{root}, u) \mid u \in S\}.$$

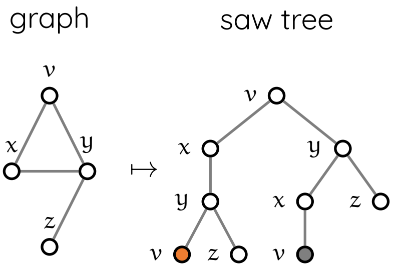
- Strong spatial mixing:

$$d_{TV}(\text{root} \mid \sigma, \text{root} \mid \sigma') \to 0$$

as the following goes to ∞:

$$\min\{d(\text{root}, u) \mid \sigma(u) \neq \sigma'(u)\}.$$

Review



- Root marginals are the same.
- Weak spatial mixing:

$$d_{\mathrm{TV}}(\text{root} \mid \sigma, \text{root} \mid \sigma') \to 0$$

as the following goes to ∞:

$$\min\{d(\text{root}, u) \mid u \in S\}.$$

- Strong spatial mixing:

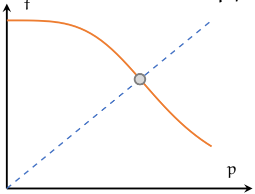
$$d_{\mathrm{TV}}(\text{root} \mid \sigma, \text{root} \mid \sigma') \to 0$$

as the following goes to ∞:

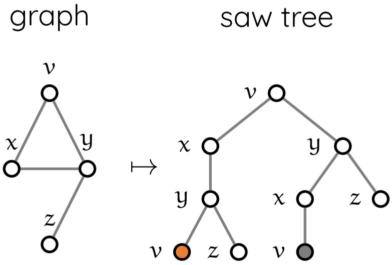
$$\min\{d(\text{root}, u) \mid \sigma(u) \neq \sigma'(u)\}.$$

- [Weitz]'s alg forms truncated saw tree and uses recursion:

$$f(p_1, \dots, p_d) = \frac{1}{1+\lambda p_1 \cdots p_d}$$



Review


graph

saw tree

- Root marginals are the same.
- Weak spatial mixing:

$$d_{TV}(\text{root} \mid \sigma, \text{root} \mid \sigma') \to 0$$

as the following goes to ∞:

$$\min\{d(\text{root}, u) \mid u \in S\}.$$

- Strong spatial mixing:

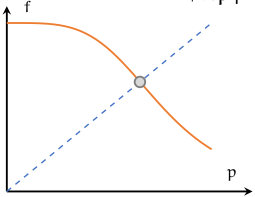
$$d_{TV}(\text{root} \mid \sigma, \text{root} \mid \sigma') \to 0$$

as the following goes to ∞:

$$\min\{d(\text{root}, u) \mid \sigma(u) \neq \sigma'(u)\}.$$

- [Weitz]'s alg forms truncated saw tree and uses recursion:

$$f(p_1, \ldots, p_d) = \frac{1}{1+\lambda p_1 \cdots p_d}$$



- Attractive exactly when $\lambda < \lambda_c(\Delta)$

Remarks

- The tree equivalence works on any 2-spin systems.

Remarks

- The tree equivalence works on any 2-spin systems.
- Correlation decay predicts efficient sampling threshold for anti-ferromagnetic 2-spin systems.

Remarks

- The tree equivalence works on any 2-spin systems.
- Correlation decay predicts efficient sampling threshold for anti-ferromagnetic 2-spin systems.
- In some ferromagnetic systems, efficient sampling [Jerrum-Sinclair] beyond correlation decay.

Remarks

- The tree equivalence works on any 2-spin systems.
- Correlation decay predicts efficient sampling threshold for anti-ferromagnetic 2-spin systems.
- In some ferromagnetic systems, efficient sampling [Jerrum-Sinclair] beyond correlation decay.
- In tree recursion $p \mapsto f(p)$, originally proof of convergence involved multiple iterations of f. The change of variable trick $(\psi \circ f \circ \psi^{-1})$ is due to [Li-Lu-Yin].

Remarks

- The tree equivalence works on any 2-spin systems.
- Correlation decay predicts efficient sampling threshold for anti-ferromagnetic 2-spin systems.
- In some ferromagnetic systems, efficient sampling [Jerrum-Sinclair] beyond correlation decay.
- In tree recursion $p \mapsto f(p)$, originally proof of convergence involved multiple iterations of f. The change of variable trick $(\psi \circ f \circ \psi^{-1})$ is due to [Li-Lu-Yin].
- Open: extend to beyond binary domains. There are some notions (more complicated) of saw tree, but we no longer have things like SSM on trees $\implies$ SSM on graphs.

Remarks

- The tree equivalence works on any 2-spin systems.
- Correlation decay predicts efficient sampling threshold for anti-ferromagnetic 2-spin systems.
- In some ferromagnetic systems, efficient sampling [Jerrum-Sinclair] beyond correlation decay.
- In tree recursion $p \mapsto f(p)$, originally proof of convergence involved multiple iterations of f. The change of variable trick ($\psi \circ f \circ \psi^{-1}$) is due to [Li-Lu-Yin].
- Open: extend to beyond binary domains. There are some notions (more complicated) of saw tree, but we no longer have things like SSM on trees $\implies$ SSM on graphs.
- Conjecture: for q-colorings, we have SSM as soon as $q \geqslant \Delta + 2$.

Remarks

- The tree equivalence works on any 2-spin systems.
- Correlation decay predicts efficient sampling threshold for anti-ferromagnetic 2-spin systems.
- In some ferromagnetic systems, efficient sampling [Jerrum-Sinclair] beyond correlation decay.
- In tree recursion $p \mapsto f(p)$, originally proof of convergence involved multiple iterations of f. The change of variable trick $(\psi \circ f \circ \psi^{-1})$ is due to [Li-Lu-Yin].
- Open: extend to beyond binary domains. There are some notions (more complicated) of saw tree, but we no longer have things like SSM on trees $\implies$ SSM on graphs.
- Conjecture: for q-colorings, we have SSM as soon as $q \geqslant \Delta + 2$.
- SSM for colorings was open even on trees. [Chen-Liu-Mani-Moitra'23] proved it for $q \geqslant \Delta + 3$.

Remarks

- The tree equivalence works on any 2-spin systems.
- Correlation decay predicts efficient sampling threshold for anti-ferromagnetic 2-spin systems.
- In some ferromagnetic systems, efficient sampling [Jerrum-Sinclair] beyond correlation decay.
- In tree recursion $p \mapsto f(p)$, originally proof of convergence involved multiple iterations of f. The change of variable trick $(\psi \circ f \circ \psi^{-1})$ is due to [Li-Lu-Yin].
- Open: extend to beyond binary domains. There are some notions (more complicated) of saw tree, but we no longer have things like SSM on trees $\implies$ SSM on graphs.
- Conjecture: for q-colorings, we have SSM as soon as $q \geqslant \Delta + 2$.
- SSM for colorings was open even on trees. [Chen-Liu-Mani-Moitra'23] proved it for $q \geqslant \Delta + 3$.
- Corollary: large girth graphs.

Remarks

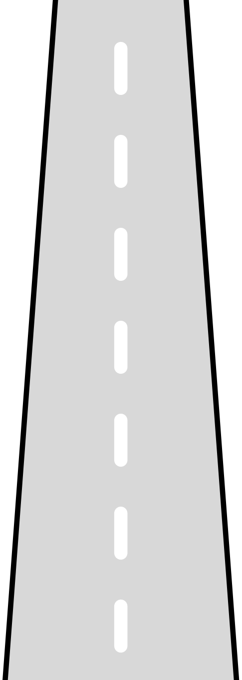
- The tree equivalence works on any 2-spin systems.
- Correlation decay predicts efficient sampling threshold for anti-ferromagnetic 2-spin systems.
- In some ferromagnetic systems, efficient sampling [Jerrum-Sinclair] beyond correlation decay.
- In tree recursion $p \mapsto f(p)$, originally proof of convergence involved multiple iterations of f. The change of variable trick $(\psi \circ f \circ \psi^{-1})$ is due to [Li-Lu-Yin].
- Open: extend to beyond binary domains. There are some notions (more complicated) of saw tree, but we no longer have things like SSM on trees $\implies$ SSM on graphs.
- Conjecture: for q-colorings, we have SSM as soon as $q \geqslant \Delta + 2$.
- SSM for colorings was open even on trees. [Chen-Liu-Mani-Moitra'23] proved it for $q \geqslant \Delta + 3$.
- Corollary: large girth graphs.
- Open: runtime of deterministic algs seem to be $n^{O(\log \Delta)}$, can we remove bad dependency on Δ?

HDX via Correlation Decay

- Influences
- Fast sampling

HDX via Transport

- Universality

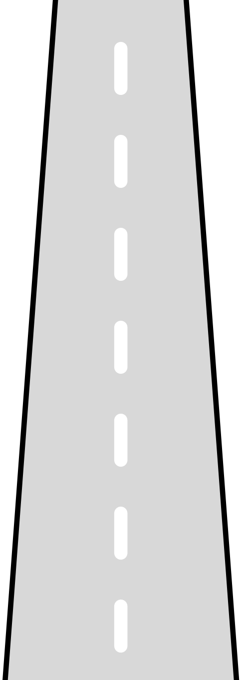


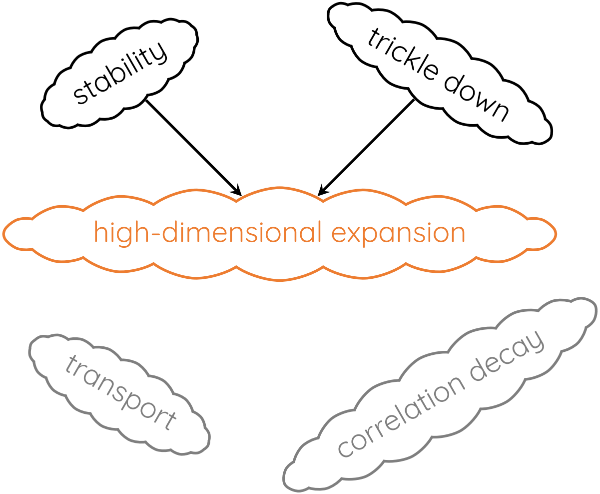
HDX via Correlation Decay

- Influences
- Fast sampling

HDX via Transport

- Universality





stability

trickle down

high-dimensional expansion

transport

correlation decay

Influences

- We now derive spectral independence from SSM.

Influences

- We now derive spectral independence from SSM.

different from Dobrushin's

↓

- Idea: bound influence matrix

$$\mathcal{I}[i \to j] = \mathbb{P}[X_j = 1 \mid X_i = 1] - \mathbb{P}[X_j = 1 \mid X_i = 0]$$

Influences

- We now derive spectral independence from SSM.

different from Dobrushin's
↓

- Idea: bound influence matrix

$$\mathcal{I}[i \to j] = \mathbb{P}[X_j = 1 \mid X_i = 1] - \mathbb{P}[X_j = 1 \mid X_i = 0]$$

- [A-Liu-OveisGharan] showed how to bound ℓ_1 of columns for hardcore. We follow [Chen-Liu-Vigoda]'s approach and bound ℓ_1 of rows.

Influences

- We now derive spectral independence from SSM.

different from Dobrushin's

↓

- Idea: bound influence matrix

$$\mathcal{I}[i \to j] = \mathbb{P}[X_j = 1 \mid X_i = 1] - \mathbb{P}[X_j = 1 \mid X_i = 0]$$

- [A-Liu-OveisGharan] showed how to bound ℓ_1 of columns for hardcore. We follow [Chen-Liu-Vigoda]'s approach and bound ℓ_1 of rows.
- If Ψ is correlation matrix:

$$\ell_1(\text{rows of } \Psi) \leqslant O(1) \cdot \ell_1(\text{rows of } \mathcal{I})$$

Influences

- We now derive spectral independence from SSM.

different from Dobrushin's
↓

- Idea: bound influence matrix

$$\mathcal{J}[i \to j] = \mathbb{P}[X_j = 1 \mid X_i = 1] - \mathbb{P}[X_j = 1 \mid X_i = 0]$$

- [A-Liu-OveisGharan] showed how to bound ℓ_1 of columns for hardcore. We follow [Chen-Liu-Vigoda]'s approach and bound ℓ_1 of rows.
- If Ψ is correlation matrix:

$$\ell_1(\text{rows of } \Psi) \leqslant O(1) \cdot \ell_1(\text{rows of } \mathcal{J})$$

Lemma

If we form saw tree rooted at i, then

$$\mathcal{J}_{\text{graph}}[i \to j] = \sum_{u \text{ copy of } j} \mathcal{J}_{\text{tree}}[i \to u]$$

Influences

- We now derive spectral independence from SSM.

different from Dobrushin's ↓

- Idea: bound influence matrix

$$\mathcal{I}[i \to j] = \mathbb{P}[X_j = 1 \mid X_i = 1] - \mathbb{P}[X_j = 1 \mid X_i = 0]$$

- [A-Liu-OveisGharan] showed how to bound ℓ_1 of columns for hardcore. We follow [Chen-Liu-Vigoda]'s approach and bound ℓ_1 of rows.
- If Ψ is correlation matrix:

$$\ell_1(\text{rows of } \Psi) \leqslant O(1) \cdot \ell_1(\text{rows of } \mathcal{I})$$

Lemma

If we form saw tree rooted at i, then

$$\mathcal{I}_{\text{graph}}[i \to j] = \sum_{u \text{ copy of } j} \mathcal{I}_{\text{tree}}[i \to u]$$

Proof:

Influences

- We now derive spectral independence from SSM.

  different from Dobrushin's

  ↓
- Idea: bound influence matrix

  $$\mathcal{I}[i \to j] = \mathbb{P}[X_j = 1 \mid X_i = 1] - \mathbb{P}[X_j = 1 \mid X_i = 0]$$
- [A-Liu-OveisGharan] showed how to bound ℓ_1 of columns for hardcore. We follow [Chen-Liu-Vigoda]'s approach and bound ℓ_1 of rows.
- If Ψ is correlation matrix:

  $$\ell_1(\text{rows of } \Psi) \leqslant O(1) \cdot \ell_1(\text{rows of } \mathcal{I})$$

Lemma

If we form saw tree rooted at i, then

$$\mathcal{I}_{\text{graph}}[i \to j] = \sum_{u \text{ copy of } j} \mathcal{I}_{\text{tree}}[i \to u]$$

Proof:

- We know $h \cdot g_{\text{graph}} = g_{\text{tree}}$, expanding on z_i we get

  $$h(z_i r + s) = z_i r' + s'$$

  where r, s, r', s' are free from z_i.

Influences

- We now derive spectral independence from SSM.

different from Dobrushin's ↓

- Idea: bound influence matrix

$$\mathcal{J}[i \to j] = \mathbb{P}[X_j = 1 \mid X_i = 1] - \mathbb{P}[X_j = 1 \mid X_i = 0]$$

- [A-Liu-OveisGharan] showed how to bound ℓ_1 of columns for hardcore. We follow [Chen-Liu-Vigoda]'s approach and bound ℓ_1 of rows.
- If Ψ is correlation matrix:

$$\ell_1(\text{rows of } \Psi) \leqslant O(1) \cdot \ell_1(\text{rows of } \mathcal{J})$$

Lemma

If we form saw tree rooted at i, then

$$\mathcal{J}_{\text{graph}}[i \to j] = \sum_{u \text{ copy of } j} \mathcal{J}_{\text{tree}}[i \to u]$$

Proof:

- We know $h \cdot g_{\text{graph}} = g_{\text{tree}}$, expanding on z_i we get

$$h(z_i r + s) = z_i r' + s'$$

where r, s, r', s' are free from z_i.
- The lhs of lemma is $\partial_j \log(r/s)\big|_{z=1}$.

Influences

- We now derive spectral independence from SSM.

different from Dobrushin's

↓

- Idea: bound influence matrix

$$\mathcal{J}[i \to j] = \mathbb{P}[X_j = 1 \mid X_i = 1] - \mathbb{P}[X_j = 1 \mid X_i = 0]$$

- [A-Liu-OveisGharan] showed how to bound ℓ_1 of columns for hardcore. We follow [Chen-Liu-Vigoda]'s approach and bound ℓ_1 of rows.
- If Ψ is correlation matrix:

$$\ell_1(\text{rows of } \Psi) \leqslant O(1) \cdot \ell_1(\text{rows of } \mathcal{J})$$

Lemma

If we form saw tree rooted at i, then

$$\mathcal{J}_{\text{graph}}[i \to j] = \sum_{u \text{ copy of } j} \mathcal{J}_{\text{tree}}[i \to u]$$

Proof:

- We know $h \cdot g_{\text{graph}} = g_{\text{tree}}$, expanding on z_i we get

$$h(z_i r + s) = z_i r' + s'$$

where r, s, r', s' are free from z_i.

- The lhs of lemma is $\partial_j \log(r/s)\big|_{z=1}$.
- The rhs is $\partial_j \log(r'/s')\big|_{z=1}$.

Influences

- We now derive spectral independence from SSM.

different from Dobrushin's

↓

- Idea: bound influence matrix

$$\mathcal{I}[i \to j] = \mathbb{P}[X_j = 1 \mid X_i = 1] - \mathbb{P}[X_j = 1 \mid X_i = 0]$$

- [A-Liu-OveisGharan] showed how to bound ℓ_1 of columns for hardcore. We follow [Chen-Liu-Vigoda]'s approach and bound ℓ_1 of rows.
- If Ψ is correlation matrix:

$$\ell_1(\text{rows of } \Psi) \leqslant O(1) \cdot \ell_1(\text{rows of } \mathcal{I})$$

Lemma

If we form saw tree rooted at i, then

$$\mathcal{I}_{\text{graph}}[i \to j] = \sum_{u \text{ copy of } j} \mathcal{I}_{\text{tree}}[i \to u]$$

Proof:

- We know $h \cdot g_{\text{graph}} = g_{\text{tree}}$, expanding on z_i we get

$$h(z_i r + s) = z_i r' + s'$$

where r, s, r', s' are free from z_i.
- The lhs of lemma is $\partial_j \log(r/s)\big|_{z=1}$.
- The rhs is $\partial_j \log(r'/s')\big|_{z=1}$.
- Equal because $r' = rh$ and $s' = sh$.

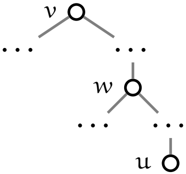
- Remains to show on v-rooted tree:

$$\sum_u |\mathcal{I}[v \to u]| = O(1)$$

- Remains to show on v-rooted tree:

$$\sum_u |\mathcal{I}[v \to u]| = O(1)$$

- Influences multiply on the tree:

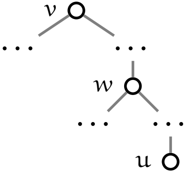


$$\mathcal{I}[v \to u] = \mathcal{I}[v \to w]\,\mathcal{I}[w \to u]$$

- Remains to show on v-rooted tree:

$$\sum_u |\mathcal{I}[v \to u]| = O(1)$$

- Influences multiply on the tree:



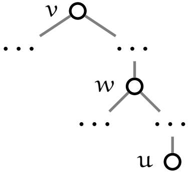
$$\mathcal{I}[v \to u] = \mathcal{I}[v \to w]\, \mathcal{I}[w \to u]$$

- If we track $t_\ell = \sum_{u \in \text{level } \ell} |\mathcal{I}[v \to u]|$, and show for some $\ell = O(1)$, we have $t_\ell \leqslant 1 - \epsilon$, we are done. ☺

- Remains to show on v-rooted tree:

$$\sum_u |\mathcal{I}[v \to u]| = O(1)$$

- Influences multiply on the tree:



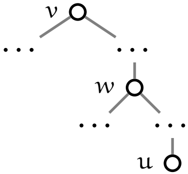
$$\mathcal{I}[v \to u] = \mathcal{I}[v \to w]\,\mathcal{I}[w \to u]$$

- If we track $t_\ell = \sum_{u \in \text{level } \ell} |\mathcal{I}[v \to u]|$, and show for some $\ell = O(1)$, we have $t_\ell \leqslant 1 - \epsilon$, we are done. 🙂
- Let p_u be $\mathbb{P}[X_u = 0]$ in u's subtree.

- Remains to show on v-rooted tree:

$$\sum_u |\mathcal{I}[v \to u]| = O(1)$$

- Influences multiply on the tree:



$$\mathcal{I}[v \to u] = \mathcal{I}[v \to w]\,\mathcal{I}[w \to u]$$

- If we track $t_\ell = \sum_{u \in \text{level } \ell} |\mathcal{I}[v \to u]|$, and show for some $\ell = O(1)$, we have $t_\ell \leqslant 1 - \epsilon$, we are done. 🙂
- Let p_u be $\mathbb{P}[X_u = 0]$ in u's subtree.
- Using $q_u = \log((1 - p_u)/p_u)$, we have recursion

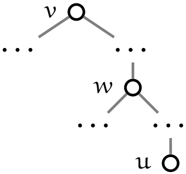
$$q_v = f(q_{u_1}, \ldots, q_{u_k})$$

where $u_1, \ldots, u_k$ are children of v.

- Remains to show on v-rooted tree:

$$\sum_u |\mathcal{J}[v \to u]| = O(1)$$

- Influences multiply on the tree:



$$\mathcal{J}[v \to u] = \mathcal{J}[v \to w]\,\mathcal{J}[w \to u]$$

- If we track $t_\ell = \sum_{u \in \text{level } \ell} |\mathcal{J}[v \to u]|$, and show for some $\ell = O(1)$, we have $t_\ell \leqslant 1 - \epsilon$, we are done. 🙂
- Let p_u be $\mathbb{P}[X_u = 0]$ in u's subtree.
- Using $q_u = \log((1 - p_u)/p_u)$, we have recursion

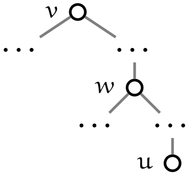
$$q_v = f(q_{u_1}, \ldots, q_{u_k})$$

where $u_1, \ldots, u_k$ are children of v.

- Claim:

$$\mathcal{J}[v \to u_i] = \partial_i f(q_{u_1}, \ldots, q_{u_k})$$

- Remains to show on v-rooted tree:
$$\sum_u |\mathcal{I}[v \to u]| = O(1)$$
- Influences multiply on the tree:



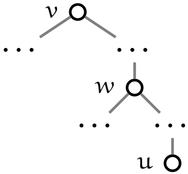
$$\mathcal{I}[v \to u] = \mathcal{I}[v \to w]\, \mathcal{I}[w \to u]$$

- If we track $t_\ell = \sum_{u \in \text{level } \ell} |\mathcal{I}[v \to u]|$, and show for some $\ell = O(1)$, we have $t_\ell \leqslant 1 - \epsilon$, we are done. ☺
- Let p_u be $\mathbb{P}[X_u = 0]$ in u's subtree.
- Using $q_u = \log((1 - p_u)/p_u)$, we have recursion
$$q_v = f(q_{u_1}, \dots, q_{u_k})$$
where $u_1, \dots, u_k$ are children of v.
- Claim:
$$\mathcal{I}[v \to u_i] = \partial_i f(q_{u_1}, \dots, q_{u_k})$$
- If $\|\nabla f\|_1 \leqslant 1 - \epsilon$, we'd be done.

- Remains to show on v-rooted tree:

$$\sum_u |\mathcal{I}[v \to u]| = O(1)$$

- Influences multiply on the tree:



$$\mathcal{I}[v \to u] = \mathcal{I}[v \to w]\,\mathcal{I}[w \to u]$$

- If we track $t_\ell = \sum_{u \in \text{level } \ell} |\mathcal{I}[v \to u]|$, and show for some $\ell = O(1)$, we have $t_\ell \leqslant 1 - \epsilon$, we are done. 🙂
- Let p_u be $\mathbb{P}[X_u = 0]$ in u's subtree.
- Using $q_u = \log((1 - p_u)/p_u)$, we have recursion

$$q_v = f(q_{u_1}, \dots, q_{u_k})$$

where $u_1, \dots, u_k$ are children of v.

- Claim:

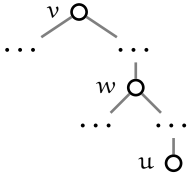
$$\mathcal{I}[v \to u_i] = \partial_i f(q_{u_1}, \dots, q_{u_k})$$

- If $\|\nabla f\|_1 \leqslant 1 - \epsilon$, we'd be done.
- Unfortunately, this is not the case. But we can use previous trick

$$g = \psi \circ f(\psi^{-1}(\cdot), \dots, \psi^{-1}(\cdot))$$

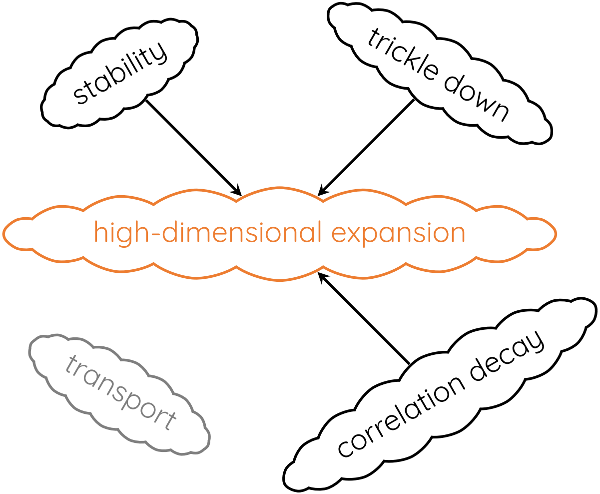
and g will have $\|\nabla g\|_1 \leqslant 1 - \epsilon$ everywhere for hardcore.

- Remains to show on v-rooted tree:
$$\sum_u |\mathcal{I}[v \to u]| = O(1)$$
- Influences multiply on the tree:



$$\mathcal{I}[v \to u] = \mathcal{I}[v \to w]\, \mathcal{I}[w \to u]$$

- If we track $t_\ell = \sum_{u \in \text{level } \ell} |\mathcal{I}[v \to u]|$, and show for some $\ell = O(1)$, we have $t_\ell \leqslant 1 - \epsilon$, we are done. ☺
- Let p_u be $\mathbb{P}[X_u = 0]$ in u's subtree.
- Using $q_u = \log((1 - p_u)/p_u)$, we have recursion
$$q_v = f(q_{u_1}, \ldots, q_{u_k})$$
where $u_1, \ldots, u_k$ are children of v.
- Claim:
$$\mathcal{I}[v \to u_i] = \partial_i f(q_{u_1}, \ldots, q_{u_k})$$
- If $\|\nabla f\|_1 \leqslant 1 - \epsilon$, we'd be done.
- Unfortunately, this is not the case. But we can use previous trick
$$g = \psi \circ f(\psi^{-1}(\cdot), \ldots, \psi^{-1}(\cdot))$$
and g will have $\|\nabla g\|_1 \leqslant 1 - \epsilon$ everywhere for hardcore.
- Derivatives of ψ are "nice". Contraction of f after $O(1)$ levels.



stability

trickle down

high-dimensional expansion

transport

correlation decay

Fast sampling

- Spectral independence shows Glauber mixing in $n^{O_{\lambda,\Delta}(1)}$. Impractical polynomial.

Fast sampling

- Spectral independence shows Glauber mixing in $n^{O_{\lambda,\Delta}(1)}$. Impractical polynomial.
- [Chen-Liu-Vigoda] showed mixing in $O_\Delta(n \log n)$. ☺

Fast sampling

- Spectral independence shows
  Glauber mixing in $n^{O_{\lambda,\Delta}(1)}$.
  Impractical polynomial.
- [Chen-Liu-Vigoda] showed mixing in
  $O_{\Delta}(n \log n)$. 🙂
- Now we know $O(n \log n)$
  [A-Jain-Koehler-Pham-Vuong,
  Chen-Eldan, Chen-Feng-Yin-Zhang] 🙂

Fast sampling

- Spectral independence shows Glauber mixing in $n^{O_{\lambda,\Delta}(1)}$. Impractical polynomial.
- [Chen-Liu-Vigoda] showed mixing in $O_\Delta(n \log n)$. 🙂
- Now we know $O(n \log n)$ [A-Jain-Koehler-Pham-Vuong, Chen-Eldan, Chen-Feng-Yin-Zhang] 🙂
- C-spectral independence implies that $n \leftrightarrow \ell$ block dynamics has relaxation time

$$\binom{n}{C} / \binom{n-\ell}{C}$$

Fast sampling

- Spectral independence shows Glauber mixing in $n^{O_{\lambda,\Delta}(1)}$. Impractical polynomial.
- [Chen-Liu-Vigoda] showed mixing in $O_\Delta(n \log n)$. 🙂
- Now we know $O(n \log n)$ [A-Jain-Koehler-Pham-Vuong, Chen-Eldan, Chen-Feng-Yin-Zhang] 🙂
- C-spectral independence implies that $n \leftrightarrow \ell$ block dynamics has relaxation time
$$\binom{n}{C}/\binom{n-\ell}{C}$$
- Note that $C = f(\Delta, \lambda)$ 🙁

Fast sampling

- Spectral independence shows Glauber mixing in $n^{O_{\lambda,\Delta}(1)}$. Impractical polynomial.
- [Chen-Liu-Vigoda] showed mixing in $O_{\Delta}(n \log n)$. 🙂
- Now we know $O(n \log n)$ [A-Jain-Koehler-Pham-Vuong, Chen-Eldan, Chen-Feng-Yin-Zhang] 🙂
- C-spectral independence implies that $n \leftrightarrow \ell$ block dynamics has relaxation time

$$\binom{n}{C} / \binom{n-\ell}{C}$$

- Note that $C = f(\Delta, \lambda)$ 🙁
- Observation: if $n - \ell = \Omega_{\Delta,\lambda}(n)$, relaxation time is $O_{\Delta,\lambda}(1)$!

Fast sampling

- Spectral independence shows Glauber mixing in $n^{O_{\lambda,\Delta}(1)}$. Impractical polynomial.
- [Chen-Liu-Vigoda] showed mixing in $O_\Delta(n \log n)$. ☺
- Now we know $O(n \log n)$ [A-Jain-Koehler-Pham-Vuong, Chen-Eldan, Chen-Feng-Yin-Zhang] ☺
- C-spectral independence implies that $n \leftrightarrow \ell$ block dynamics has relaxation time

$$\binom{n}{C}/\binom{n-\ell}{C}$$

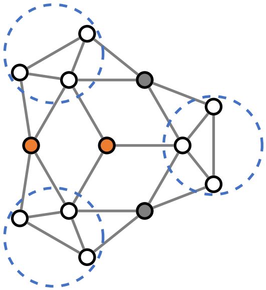
- Note that $C = f(\Delta, \lambda)$ ☹
- Observation: if $n - \ell = \Omega_{\Delta,\lambda}(n)$, relaxation time is $O_{\Delta,\lambda}(1)$!
- How to implement?

Fast sampling

- Spectral independence shows Glauber mixing in $n^{O_{\lambda,\Delta}(1)}$. Impractical polynomial.
- [Chen-Liu-Vigoda] showed mixing in $O_\Delta(n \log n)$. 🙂
- Now we know $O(n \log n)$ [A-Jain-Koehler-Pham-Vuong, Chen-Eldan, Chen-Feng-Yin-Zhang] 🙂
- C-spectral independence implies that $n \leftrightarrow \ell$ block dynamics has relaxation time

$$\binom{n}{C} / \binom{n-\ell}{C}$$

- Note that $C = f(\Delta, \lambda)$ 🙁
- Observation: if $n - \ell = \Omega_{\Delta,\lambda}(n)$, relaxation time is $O_{\Delta,\lambda}(1)$!
- How to implement?
- If $n - \ell = \delta n$ for sufficiently small δ, cond on ℓ random verts, we get islands of size $\simeq O_{\Delta,\lambda}(\log n)$. 🙂

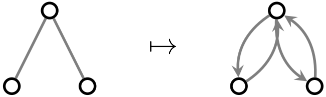


Proof:

Proof:

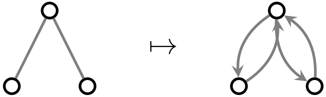
loose bound

- There are at most $n \cdot \Delta^{2(k-1)}$ connected subgraphs of size k. Covered by walks of len $2(k-1)$.



Proof:

- There are at most $n \cdot \Delta^{\overset{\text{loose bound}}{\downarrow} 2(k-1)}$ connected subgraphs of size k. Covered by walks of len $2(k-1)$.



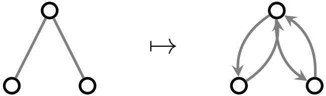
- If random δ fraction of verts are free, probability of a k-sized subgraph being free is

$$\simeq \delta^k$$

Proof:

loose bound ↓

- There are at most $n \cdot \Delta^{2(k-1)}$ connected subgraphs of size k. Covered by walks of len $2(k-1)$.



- If random δ fraction of verts are free, probability of a k-sized subgraph being free is

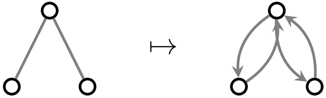
$$\simeq \delta^k$$

- Using union bound, prob of any surviving will be

$$n \cdot \Delta^{2(k-1)} \cdot \delta^k$$

Proof:

- There are at most $n \cdot \Delta^{\overset{\substack{\text{loose bound} \\ \downarrow}}{2(k-1)}}$ connected subgraphs of size k. Covered by walks of len $2(k-1)$.



- If random δ fraction of verts are free, probability of a k-sized subgraph being free is

$$\simeq \delta^k$$

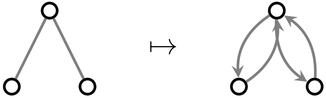
- Using union bound, prob of any surviving will be

$$n \cdot \Delta^{2(k-1)} \cdot \delta^k$$

- Set δ small and $k \simeq \log n$. 🙂

Proof:

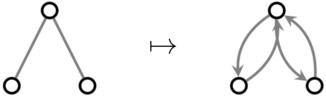
- There are at most $n \cdot \Delta^{\overset{\substack{\text{loose bound}\\ \downarrow}}{2(k-1)}}$ connected subgraphs of size k. Covered by walks of len $2(k-1)$.



- If random δ fraction of verts are free, probability of a k-sized subgraph being free is
$$\simeq \delta^k$$
- Using union bound, prob of any surviving will be
$$n \cdot \Delta^{2(k-1)} \cdot \delta^k$$
- Set δ small and $k \simeq \log n$. 🙂
- Since islands are small, we can sample from them much faster, in $\operatorname{poly}\log(n)$ time. 🙂

Proof:

- There are at most $n \cdot \Delta^{\overset{\text{loose bound}}{\downarrow}{2(k-1)}}$ connected subgraphs of size k. Covered by walks of len $2(k-1)$.

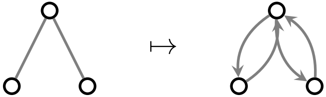


- If random δ fraction of verts are free, probability of a k-sized subgraph being free is
$$\simeq \delta^k$$
- Using union bound, prob of any surviving will be
$$n \cdot \Delta^{2(k-1)} \cdot \delta^k$$
- Set δ small and $k \simeq \log n$. ☺
- Since islands are small, we can sample from them much faster, in $\text{poly} \log(n)$ time. ☺
- Alternatively, you can use some form of comparison to prove $\widetilde{O}(n)$ relaxation time for Glauber dynamics itself [Chen-Liu-Vigoda]. ☺

Proof:

loose bound

- There are at most $n \cdot \Delta^{\overset{\downarrow}{2(k-1)}}$ connected subgraphs of size k. Covered by walks of len $2(k-1)$.



- If random δ fraction of verts are free, probability of a k-sized subgraph being free is
$$\simeq \delta^k$$
- Using union bound, prob of any surviving will be
$$n \cdot \Delta^{2(k-1)} \cdot \delta^k$$
- Set δ small and $k \simeq \log n$. 🙂

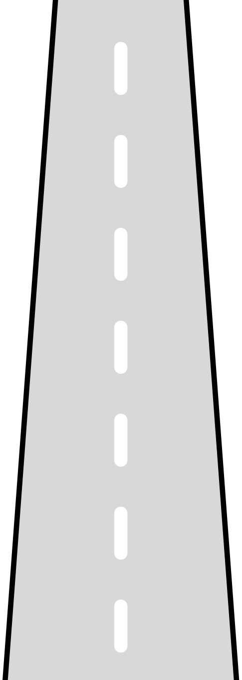
- Since islands are small, we can sample from them much faster, in $\mathbf{poly}\log(n)$ time. 🙂
- Alternatively, you can use some form of comparison to prove $\widetilde{O}(n)$ relaxation time for Glauber dynamics itself [Chen-Liu-Vigoda]. 🙂
- Note that spectral independence itself only gives us $\widetilde{O}(n^2)$ time algs. We need entropy contraction to get $\widetilde{O}(n)$. This was done for $\Delta = O(1)$ by [Chen-Liu-Vigoda], and for general Δ by [A-Jain-Koehler-Pham-Vuong].

HDX via Correlation Decay

- Influences
- Fast sampling

HDX via Transport

- Universality

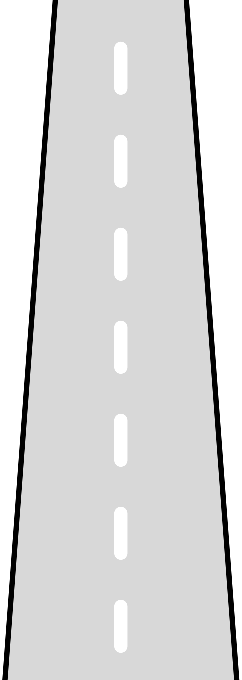


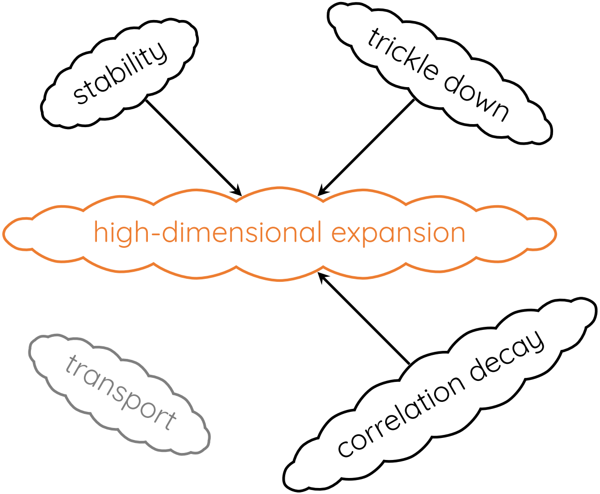
HDX via Correlation Decay

- Influences
- Fast sampling

HDX via Transport

- Universality





stability

trickle down

high-dimensional expansion

transport

correlation decay

HDX from transport

- We proved optimal mixing of colorings for $q \geqslant 2\Delta + 1$.

HDX from transport

- We proved optimal mixing of colorings for $q \geqslant 2\Delta + 1$.
- The proof was based on contraction of transport distance.

HDX from transport

- We proved optimal mixing of colorings for $q \geqslant 2\Delta + 1$.
- The proof was based on contraction of transport distance.
- Even the state-of-the-art bounds of $q \simeq \frac{11}{6}\Delta - \epsilon$ rely on transport contraction, although on a different chain. This resulted in slow but polynomial mixing.

HDX from transport

- We proved optimal mixing of colorings for $q \geqslant 2\Delta + 1$.
- The proof was based on contraction of transport distance.
- Even the state-of-the-art bounds of $q \simeq \frac{11}{6}\Delta - \epsilon$ rely on transport contraction, although on a different chain. This resulted in slow but polynomial mixing.
- Does transport contraction imply spectral independence?

HDX from transport

- We proved optimal mixing of colorings for $q \geqslant 2\Delta + 1$.
- The proof was based on contraction of transport distance.
- Even the state-of-the-art bounds of $q \simeq \frac{11}{6}\Delta - \epsilon$ rely on transport contraction, although on a different chain. This resulted in slow but polynomial mixing.
- Does transport contraction imply spectral independence?
- This give us fast sampling. ☺

HDX from transport

- We proved optimal mixing of colorings for $q \geqslant 2\Delta + 1$.
- The proof was based on contraction of transport distance.
- Even the state-of-the-art bounds of $q \simeq \frac{11}{6}\Delta - \epsilon$ rely on transport contraction, although on a different chain. This resulted in slow but polynomial mixing.
- Does transport contraction imply spectral independence?
- This give us fast sampling. 🙂
- Spectral independence from transport contraction was first dervied by [Liu, Blanca-Caputo-Chen-Parisi-Štefankovič-Vigoda].

HDX from transport

- We proved optimal mixing of colorings for $q \geqslant 2\Delta + 1$.
- The proof was based on contraction of transport distance.
- Even the state-of-the-art bounds of $q \simeq \frac{11}{6}\Delta - \epsilon$ rely on transport contraction, although on a different chain. This resulted in slow but polynomial mixing.
- Does transport contraction imply spectral independence?
- This give us fast sampling. 🙂
- Spectral independence from transport contraction was first dervied by [Liu, Blanca-Caputo-Chen-Parisi-Štefankovič-Vigoda].
- I will show a different argument based on universality [A-Jain-Koehler-Pham-Vuong].

HDX from transport

- We proved optimal mixing of colorings for $q \geqslant 2\Delta + 1$.
- The proof was based on contraction of transport distance.
- Even the state-of-the-art bounds of $q \simeq \frac{11}{6}\Delta - \epsilon$ rely on transport contraction, although on a different chain. This resulted in slow but polynomial mixing.
- Does transport contraction imply spectral independence?
- This give us fast sampling. ☺
- Spectral independence from transport contraction was first dervied by [Liu, Blanca-Caputo-Chen-Parisi-Štefankovič-Vigoda].
- I will show a different argument based on universality [A-Jain-Koehler-Pham-Vuong].

Universality of SI

If down-up walk for dist μ on $\binom{[n]}{k}$ has relaxation time $O(k)$, then μ is $O(1)$-SI.

HDX from transport

- We proved optimal mixing of colorings for $q \geqslant 2\Delta + 1$.
- The proof was based on contraction of transport distance.
- Even the state-of-the-art bounds of $q \simeq \frac{11}{6}\Delta - \epsilon$ rely on transport contraction, although on a different chain. This resulted in slow but polynomial mixing.
- Does transport contraction imply spectral independence?
- This give us fast sampling. 🙂

- Spectral independence from transport contraction was first dervied by [Liu, Blanca-Caputo-Chen-Parisi-Štefankovič-Vigoda].
- I will show a different argument based on universality [A-Jain-Koehler-Pham-Vuong].

Universality of SI

If down-up walk for dist μ on $\binom{[n]}{k}$ has relaxation time $O(k)$, then μ is $O(1)$-SI.

- So Dobrushin++ implies SI.

Proof:

Proof:

- Embedding $\binom{[n]}{k} \hookrightarrow \{0,1\}^n$, we need to prove

$$\mathrm{cov}(\mu) \preceq O(1) \cdot \mathrm{diag}(\mathrm{mean}(\mu))$$

Proof:

- Embedding $\binom{[n]}{k} \hookrightarrow \{0,1\}^n$, we need to prove
$$\mathrm{cov}(\mu) \preceq O(1) \cdot \mathrm{diag}(\mathrm{mean}(\mu))$$
- So for vector u, we need to show
$$u^\intercal \mathrm{cov}(\mu) u \leqslant O(1) \cdot u^\intercal \mathrm{diag}(\mathrm{mean}(\mu)) u$$

Proof:

- Embedding $\binom{[n]}{k} \hookrightarrow \{0,1\}^n$, we need to prove

$$\mathsf{cov}(\mu) \preceq O(1) \cdot \mathsf{diag}(\mathsf{mean}(\mu))$$

- So for vector u, we need to show

$$u^\intercal \mathsf{cov}(\mu) u \leqslant O(1) \cdot u^\intercal \mathsf{diag}(\mathsf{mean}(\mu)) u$$

- Now define function f on $\{0,1\}^n$ as

$$f(x) = \langle u, x \rangle$$

and let $\nu = f\mu$.

Proof:

- Embedding $\binom{[n]}{k} \hookrightarrow \{0,1\}^n$, we need to prove

$$\mathrm{cov}(\mu) \preceq O(1) \cdot \mathrm{diag}(\mathrm{mean}(\mu))$$

- So for vector u, we need to show

$$u^\intercal \mathrm{cov}(\mu) u \leqslant O(1) \cdot u^\intercal \mathrm{diag}(\mathrm{mean}(\mu)) u$$

- Now define function f on $\{0,1\}^n$ as

$$f(x) = \langle u, x \rangle$$

  and let $\nu = f\mu$.

- We have $\chi^2(\nu \parallel \mu) = \mathrm{Var}_\mu[f] = \mathbb{E}_{x\sim\mu}[(u^\intercal x)^2] - \mathbb{E}_{x\sim\mu}[u^\intercal x]^2 = u^\intercal \mathrm{cov}(\mu) u$

Proof:

- Embedding $\binom{[n]}{k} \hookrightarrow \{0,1\}^n$, we need to prove
$$\mathrm{cov}(\mu) \preceq O(1) \cdot \mathrm{diag}(\mathrm{mean}(\mu))$$
- So for vector u, we need to show
$$u^\intercal \,\mathrm{cov}(\mu) u \leqslant O(1) \cdot u^\intercal \mathrm{diag}(\mathrm{mean}(\mu)) u$$
- Now define function f on $\{0,1\}^n$ as
$$f(x) = \langle u, x \rangle$$
and let $\nu = f\mu$.
- We have $\chi^2(\nu \parallel \mu) = \mathrm{Var}_\mu[f] = \mathbb{E}_{x\sim\mu}[(u^\intercal x)^2] - \mathbb{E}_{x\sim\mu}[u^\intercal x]^2 = u^\intercal \,\mathrm{cov}(\mu) u$
- Because of relaxation time of $O(k)$, we have $\frac{\mathrm{Var}_\mu[f]}{\Omega(k)} \leqslant \mathbb{E}_{y\sim\mu D_{k\to k-1}}\left[\mathrm{Var}_{U_{k-1\to k}(y,\cdot)}[f]\right]$

Proof:

- Embedding $\binom{[n]}{k} \hookrightarrow \{0,1\}^n$, we need to prove
$$\mathrm{cov}(\mu) \preceq O(1) \cdot \mathrm{diag}(\mathrm{mean}(\mu))$$
- So for vector u, we need to show
$$u^\intercal \,\mathrm{cov}(\mu) u \leqslant O(1) \cdot u^\intercal \mathrm{diag}(\mathrm{mean}(\mu)) u$$
- Now define function f on $\{0,1\}^n$ as
$$f(x) = \langle u, x \rangle$$
and let $\nu = f\mu$.
- We have $\chi^2(\nu \parallel \mu) = \mathrm{Var}_\mu[f] =$
$$\mathbb{E}_{x \sim \mu}[(u^\intercal x)^2] - \mathbb{E}_{x\sim\mu}[u^\intercal x]^2 = u^\intercal \,\mathrm{cov}(\mu) u$$
- Because of relaxation time of $O(k)$, we have $\frac{\mathrm{Var}_\mu[f]}{\Omega(k)} \leqslant$
$$\mathbb{E}_{y \sim \mu D_{k \to k-1}}\left[\mathrm{Var}_{U_{k-1\to k}(y,\cdot)}[f]\right]$$
- But $\mathrm{Var}_{U_{k-1\to k}(y,\cdot)}[f] \leqslant$
$$\mathbb{E}_{x \sim U_{k-1\to k}(y,\cdot)}[(f(x)-f(y))^2]$$

Proof:

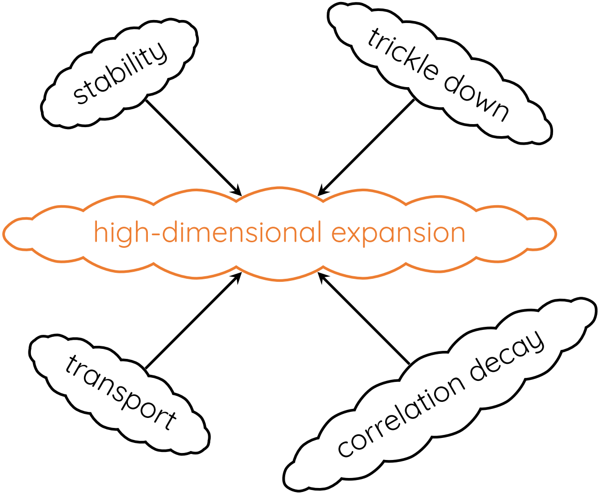
- Embedding $\binom{[n]}{k} \hookrightarrow \{0,1\}^n$, we need to prove
$$\mathrm{cov}(\mu) \preceq O(1) \cdot \mathrm{diag}(\mathrm{mean}(\mu))$$
- So for vector u, we need to show
$$u^\intercal \mathrm{cov}(\mu) u \leqslant O(1) \cdot u^\intercal \mathrm{diag}(\mathrm{mean}(\mu)) u$$
- Now define function f on $\{0,1\}^n$ as
$$f(x) = \langle u, x \rangle$$
and let $\nu = f\mu$.
- We have $\chi^2(\nu \parallel \mu) = \mathrm{Var}_\mu[f] = \mathbb{E}_{x\sim\mu}[(u^\intercal x)^2] - \mathbb{E}_{x\sim\mu}[u^\intercal x]^2 = u^\intercal \mathrm{cov}(\mu) u$
- Because of relaxation time of $O(k)$, we have $\frac{\mathrm{Var}_\mu[f]}{\Omega(k)} \leqslant$
$$\mathbb{E}_{y\sim\mu D_{k\to k-1}}\left[\mathrm{Var}_{U_{k-1\to k}(y,\cdot)}[f]\right]$$
- But $\mathrm{Var}_{U_{k-1\to k}(y,\cdot)}[f] \leqslant$
$$\mathbb{E}_{x\sim U_{k-1\to k}(y,\cdot)}[(f(x)-f(y))^2]$$
- The inside is simply $u^\intercal(x-y)(x-y)^\intercal u$. Note that $x - y = \mathbb{1}_i$ for some index i.

Proof:

- Embedding $\binom{[n]}{k} \hookrightarrow \{0,1\}^n$, we need to prove
$$\mathrm{cov}(\mu) \preceq O(1) \cdot \mathrm{diag}(\mathrm{mean}(\mu))$$
- So for vector u, we need to show
$$u^\intercal \mathrm{cov}(\mu) u \leqslant O(1) \cdot u^\intercal \mathrm{diag}(\mathrm{mean}(\mu)) u$$
- Now define function f on $\{0,1\}^n$ as
$$f(x) = \langle u, x \rangle$$
and let $\nu = f\mu$.
- We have $\chi^2(\nu \parallel \mu) = \mathrm{Var}_\mu[f] =$
$$\mathbb{E}_{x\sim\mu}[(u^\intercal x)^2] - \mathbb{E}_{x\sim\mu}[u^\intercal x]^2 = u^\intercal \mathrm{cov}(\mu) u$$
- Because of relaxation time of $O(k)$, we have $\frac{\mathrm{Var}_\mu[f]}{\Omega(k)} \leqslant$
$$\mathbb{E}_{y\sim\mu D_{k\to k-1}}\left[\mathrm{Var}_{U_{k-1\to k}(y,\cdot)}[f]\right]$$
- But $\mathrm{Var}_{U_{k-1\to k}(y,\cdot)}[f] \leqslant$
$$\mathbb{E}_{x\sim U_{k-1\to k}(y,\cdot)}[(f(x)-f(y))^2]$$
- The inside is simply $u^\intercal (x-y)(x-y)^\intercal u$. Note that $x - y = \mathbb{1}_i$ for some index i.
- Taking expectations we get
$$\frac{\mathrm{Var}_\mu[f]}{\Omega(k)} \leqslant \frac{u^\intercal \mathrm{diag}(\mathrm{mean}(\mu)) u}{k}$$

Proof:

- Embedding $\binom{[n]}{k} \hookrightarrow \{0,1\}^n$, we need to prove
$$\mathrm{cov}(\mu) \preceq O(1) \cdot \mathrm{diag}(\mathrm{mean}(\mu))$$
- So for vector u, we need to show
$$u^\intercal \mathrm{cov}(\mu) u \leqslant O(1) \cdot u^\intercal \mathrm{diag}(\mathrm{mean}(\mu)) u$$
- Now define function f on $\{0,1\}^n$ as
$$f(x) = \langle u, x \rangle$$
and let $\nu = f\mu$.
- We have $\chi^2(\nu \parallel \mu) = \mathrm{Var}_\mu[f] = \mathbb{E}_{x\sim\mu}[(u^\intercal x)^2] - \mathbb{E}_{x\sim\mu}[u^\intercal x]^2 = u^\intercal \mathrm{cov}(\mu) u$
- Because of relaxation time of $O(k)$, we have $\frac{\mathrm{Var}_\mu[f]}{\Omega(k)} \leqslant$
$$\mathbb{E}_{y\sim\mu D_{k\to k-1}}\left[\mathrm{Var}_{U_{k-1\to k}(y,\cdot)}[f]\right]$$
- But $\mathrm{Var}_{U_{k-1\to k}(y,\cdot)}[f] \leqslant$
$$\mathbb{E}_{x\sim U_{k-1\to k}(y,\cdot)}[(f(x) - f(y))^2]$$
- The inside is simply $u^\intercal(x-y)(x-y)^\intercal u$. Note that $x - y = \mathbb{1}_i$ for some index i.
- Taking expectations we get
$$\frac{\mathrm{Var}_\mu[f]}{\Omega(k)} \leqslant \frac{u^\intercal \mathrm{diag}(\mathrm{mean}(\mu)) u}{k}$$
- This finishes the proof. 🙂



stability

trickle down

high-dimensional expansion

transport

correlation decay