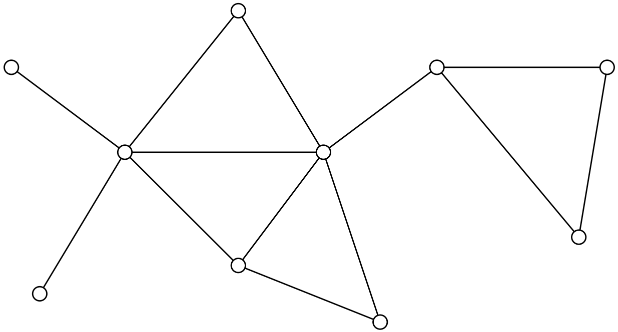
Rapid mixing of hypergraph independent sets

Yumeng Zhang
University of California, Berkeley

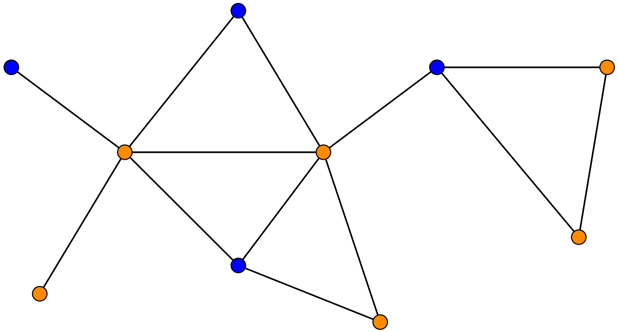
Simons Insitute
June 6th, 2017

Joint work with Jonathan Hermon and Allan Sly

Independent sets on graph



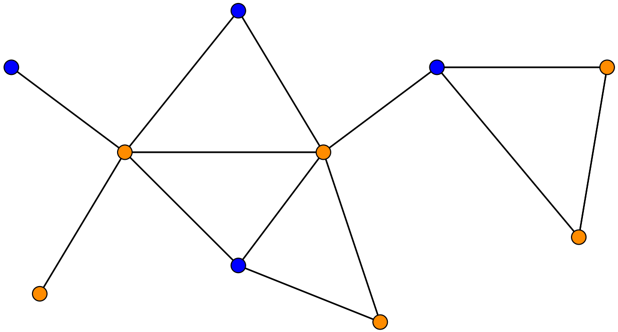
Independent sets on graph



An independent set ⇔ No edge has more than one 1.

⇔ Every edge has at least one 0.

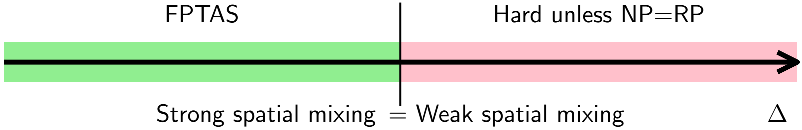
Independent sets on graph



An independent set $\Leftrightarrow$ No edge has more than one 1.
$\Leftrightarrow$ Every edge has at least one 0.

Want to approximate the partition function $Z(G)$

Phase Transition of Counting Complexity



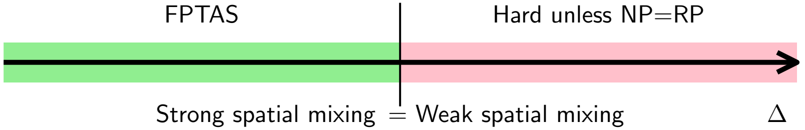
[Weitz 06], [Sly 10]

FPTAS for all graphs with maximum degree d

if and only if

Correlation decay on d-regular tree.

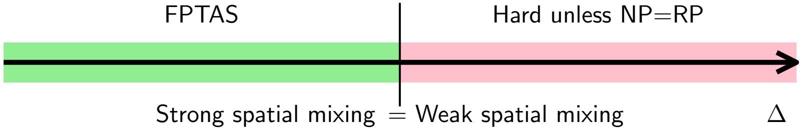
Phase Transition of Counting Complexity



[Weitz 06], [Sly 10], [Li-Lu-Yin 12], [Sly-Sun 12]

Similar result holds for other anti-ferromagnetic 2-spin systems.

Phase Transition of Counting Complexity

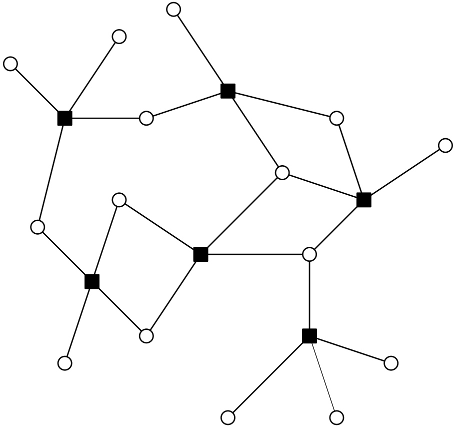


[Weitz 06], [Sly 10], [Li-Lu-Yin 12], [Sly-Sun 12]

Similar result holds for other anti-ferromagnetic 2-spin systems.

What about hypergraphs?

Independent Set on Hypergraphs



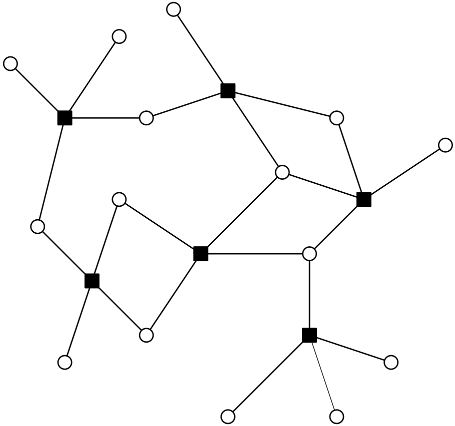
$G = (V, F, E)$

V: vertices(circles).

F: hyperedges(squares).

Degree Δ, Size k

Independent Set on Hypergraphs



$G = (V, F, E)$

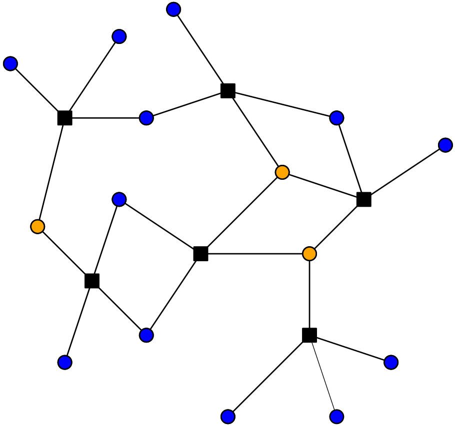
V: vertices(circles).

F: hyperedges(squares).

Degree Δ, Size k

No edge has more than one 1 $\not\Leftrightarrow$ Every edge has at least one 0.

Independent Set on Hypergraphs



$G = (V, F, E)$

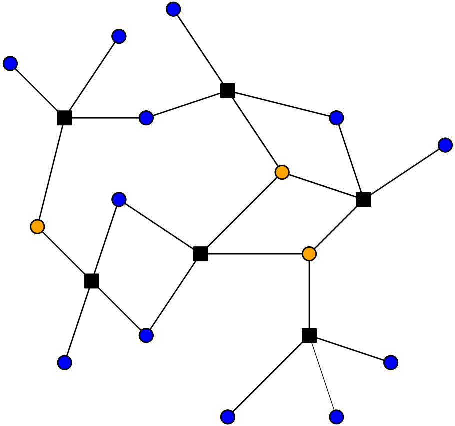
V: vertices(circles).

F: hyperedges(squares).

Degree Δ, Size k

Hypergraph independent set: Every edge has at least one 0.

Independent Set on Hypergraphs



$G = (V, F, E)$

V: vertices(circles).

F: hyperedges(squares).

Degree Δ, Size k

Hypergraph independent set: Every edge has at least one 0.

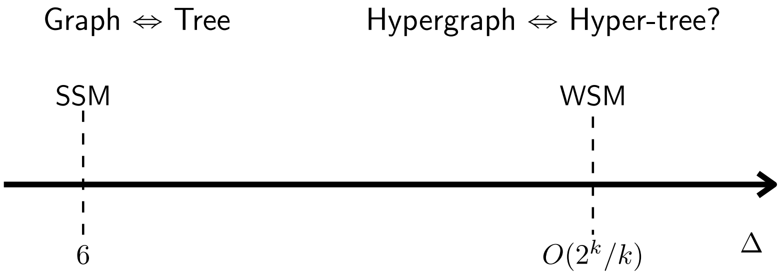
—Also known as monotone CNF.

Phase Transition of Counting Complexity

Graph $\Leftrightarrow$ Tree

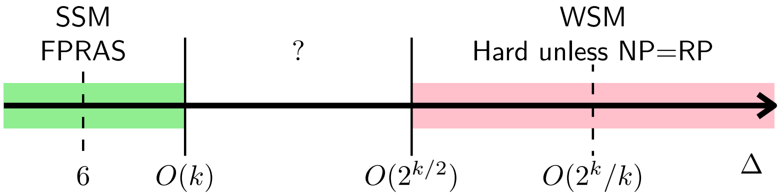
Hypergraph $\Leftrightarrow$ Hyper-tree?

Phase Transition of Counting Complexity



- SSM: correlation decay under boundary of arbitrary shape.
- WSM: correlation decay under boundary of complete tree.

Phase Transition of Counting Complexity

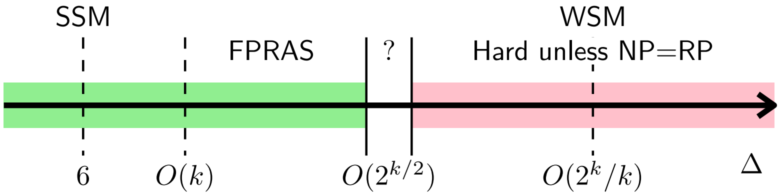


[BDK06] MCMC algo. for $\Delta \leqslant (k-1)/2$.

[BGGGS16] FPTAS for $\Delta \leqslant k$.

[BGGGS16] No approx. algo. unless NP=RP for $\Delta \geqslant 5 \cdot 2^{k/2}$.

Phase Transition of Counting Complexity



[Moitra17] FPTAS for $\Delta \leqslant e^{k/20}$.

[GJL17] Exact sampling for $\Delta \leqslant c2^{k/2}$ with min overlap.

Today MCMC algo. for $\Delta \leqslant c2^{k/2}$ (with monotonicity).

Main result

Glauber dynamics Pick a vertex, flip a coin,
set to new value whenver possible.

Mixing time $t_{\text{mix}} = \min\{t : d_{\text{TV}}(\mathbb{P}^t(\sigma, \cdot), \pi) < 1/4, \forall \sigma \in \Omega\}$.

Main result

Glauber dynamics Pick a vertex, flip a coin,
set to new value whenver possible.

Mixing time $t_{\mathrm{mix}} = \min\{t : d_{\mathrm{TV}}(\mathbb{P}^t(\sigma, \cdot), \pi) < 1/4, \forall \sigma \in \Omega\}$.

Main Theorem

For any k-hypergraph with maximum degree $\Delta \leqslant c2^{k/2}$,
the Glauber dynamics mixes in $O(n \log n)$ time.

Main result

Glauber dynamics Pick a vertex, flip a coin,
set to new value whenver possible.

Mixing time $t_{\mathrm{mix}} = \min\{t : d_{\mathrm{TV}}(\mathbb{P}^t(\sigma, \cdot), \pi) < 1/4, \forall \sigma \in \Omega\}$.

Main Theorem

For any k-hypergraph with maximum degree $\Delta \leqslant c2^{k/2}$,
the Glauber dynamics mixes in $O(n \log n)$ time.

Corollary

Rapid mixing + self-reducibility $\Rightarrow$ FPRAS

Path Coupling

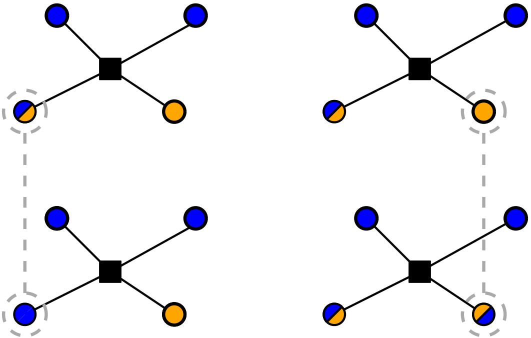
Theorem [Bubley-Dyer 97]

Let X_t, X'_t be two runs of the Glauber dynamics. If there exists a coupling of X, X' such that for **any** $d(X_0, X'_0) = 1$,

$$\mathbb{E}[d(X_1, X'_1) \mid X_0, X'_0] \leqslant 1 - \alpha/n$$

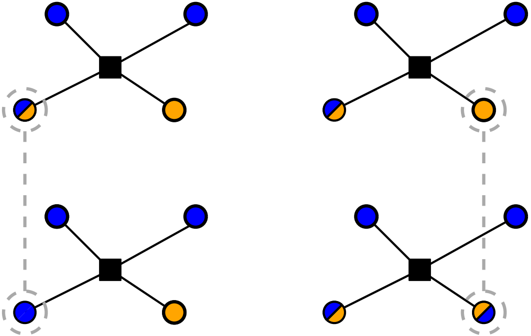
then the mixing time is $O(n \log n)$.

Path Coupling



$$\mathbb{E}[d(X_1, X_1') \mid X_0, X_0'] = 1 - \frac{1}{n} + \Delta \cdot \frac{1}{n} \cdot \frac{1}{2}.$$

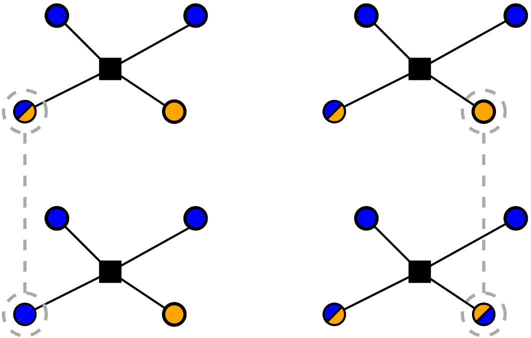
Path Coupling



$$\mathbb{E}[d(X_1, X_1') \mid X_0, X_0'] = 1 - \frac{1}{n} + \Delta \cdot \frac{1}{n} \cdot \frac{1}{2}.$$

New discrepancy is created if the other vertices in the same hyperedge are all 1,

Path Coupling



$$\mathbb{E}[d(X_1, X_1') \mid X_0, X_0'] = 1 - \frac{1}{n} + \Delta \cdot \frac{1}{n} \cdot \frac{1}{2}.$$

New discrepancy is created if the other vertices in the same hyperedge are all 1, which is **increasingly unlikely** (2^{-k}).

Path Coupling

There are lots of interactions between bad events, but each individual interaction is weak.

Path Coupling

There are lots of interactions between bad events, but each individual interaction is weak.

Union bounds such as the “vanilla” path coupling will blow up.

Path Coupling

There are lots of interactions between bad events, but each individual interaction is weak.

Union bounds such as the "vanilla" path coupling will blow up.

In this talk, we will take a more direct approach:

- Characterize the "interactions" – via analyzing the propagation of the discrepancies in the space-time graph.
- Bound their appearance – via an auxillary percolation.

Grand Coupling

and

the Propagation of Discrepancies

Grand Coupling

Continuous-time Glauber dynamics

- Place an independent Pois(1) clock at each vertex.
- At each alarm, flip a coin, set the value whenever possible.

Grand Coupling

Continuous-time Glauber dynamics

- Place an independent $\text{Pois}(1)$ clock at each vertex.
- At each alarm, flip a coin, set the value whenever possible.

Let $(X_t^\sigma)_{t \geqslant 0}$ be the process on G starting from $\sigma \in \Omega_G$;
$(Y_t^s)_{t \geqslant s}$ be the process on empty graph starting at time s from the all 1 configuration.

Grand Coupling

Continuous-time Glauber dynamics

- Place an independent Pois(1) clock at each vertex.
- At each alarm, flip a coin, set the value whenever possible.

Let $(X_t^\sigma)_{t\geqslant 0}$ be the process on G starting from $\sigma \in \Omega_G$;
$(Y_t^s)_{t\geqslant s}$ be the process on empty graph starting at time s from the all 1 configuration.

$(Y_t^s)_{t\geqslant s}$ is identical to the lazy simple random walk on $\{0, 1\}^n$.

Grand Coupling

If we use the same clocks and coins for updating X_t's and Y_t's, then

$$X_t^{\sigma}(v) \leqslant Y_t^{0}(v) \leqslant Y_t^{s}(v), \quad \forall \sigma \in \Omega_G, 0 \leqslant s \leqslant t.$$

Grand Coupling

If we use the same clocks and coins for updating X_t's and Y_t's, then

$$X_t^\sigma(v) \leqslant Y_t^0(v) \leqslant Y_t^s(v), \quad \forall \sigma \in \Omega_G, 0 \leqslant s \leqslant t.$$

Let $t_{\text{coup}} \equiv \min\{t : X_t^\sigma = X_t^\tau,\ \forall \sigma, \tau \in \Omega_G\}$. It follows that

$$t_{\text{mix}} \leqslant \min\{T : \mathbb{P}(t_{\text{coup}} > T) \leqslant 1/4\}.$$

Grand Coupling

If we use the same clocks and coins for updating X_t's and Y_t's, then

$$X_t^\sigma(v) \leqslant Y_t^0(v) \leqslant Y_t^s(v), \quad \forall \sigma \in \Omega_G, 0 \leqslant s \leqslant t.$$

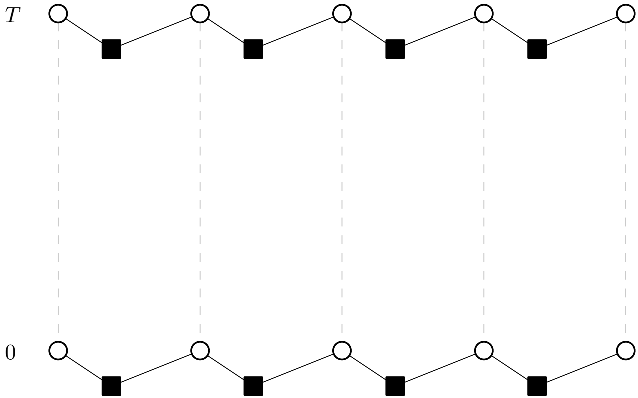
Let $t_{\text{coup}} \equiv \min\{t : X_t^\sigma = X_t^\tau,\ \forall \sigma, \tau \in \Omega_G\}$. It follows that

$$t_{\text{mix}} \leqslant \min\{T : \mathbb{P}(t_{\text{coup}} > T) \leqslant 1/4\}.$$

It is enough to bound $\mathbb{P}(t_{\text{coup}} > T)$ for $T = O(n \log n)$.

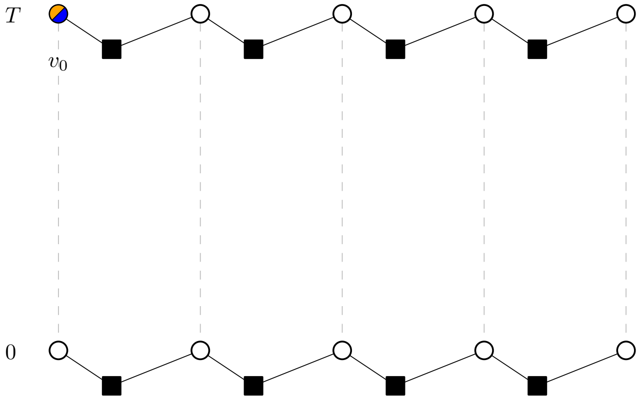
Tracing back the discrepancies

$t_{\text{coup}} \geqslant T$



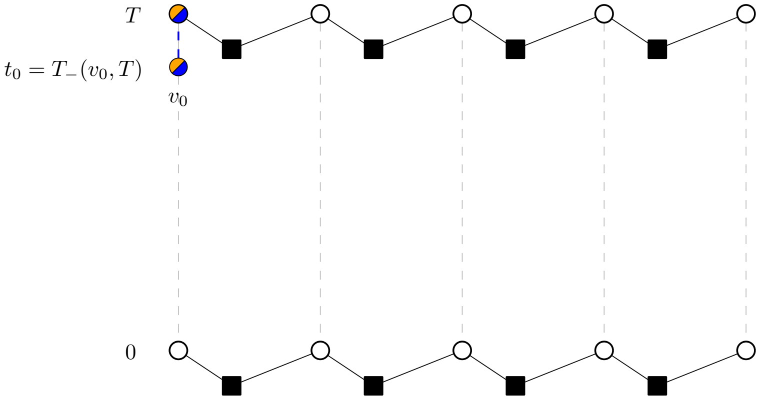
Tracing back the discrepancies

$$t_{\mathsf{coup}} \geqslant T \Rightarrow \exists \sigma, \tau \in \Omega_G, v_0 \in V \text{ such that } X_T^{\sigma}(v_0) \neq X_T^{\tau}(v_0).$$



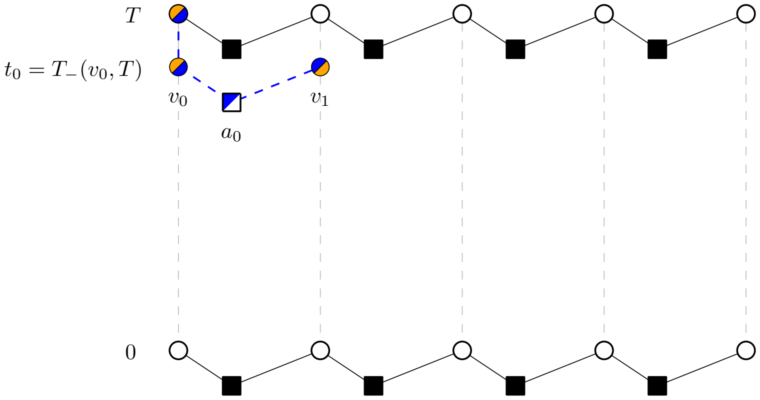
Tracing back the discrepancies

$$t_{\mathsf{coup}} \geqslant T \Rightarrow \exists \sigma, \tau \in \Omega_G, v_0 \in V \text{ such that } X_T^{\sigma}(v_0) \neq X_T^{\tau}(v_0).$$



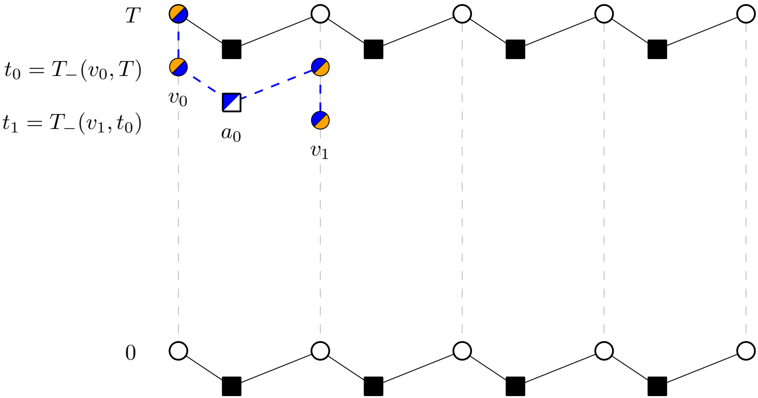
Tracing back the discrepancies

$$t_{\mathsf{coup}} \geqslant T \Rightarrow \exists \sigma, \tau \in \Omega_G, v_0 \in V \text{ such that } X_T^\sigma(v_0) \neq X_T^\tau(v_0).$$



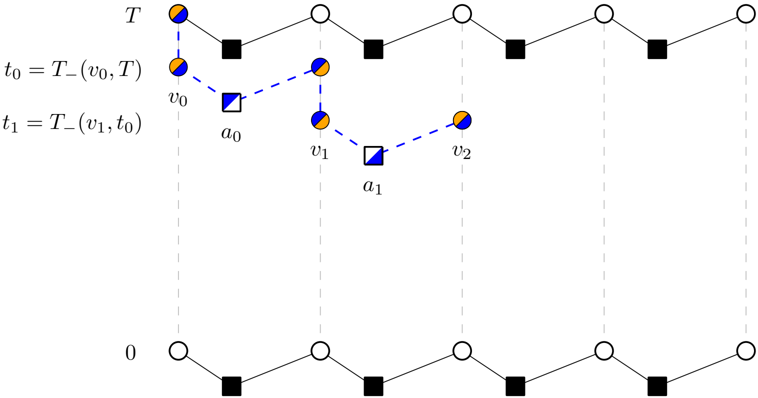
Tracing back the discrepancies

$$t_{\mathsf{coup}} \geqslant T \Rightarrow \exists \sigma, \tau \in \Omega_G, v_0 \in V \text{ such that } X_T^\sigma(v_0) \neq X_T^\tau(v_0).$$



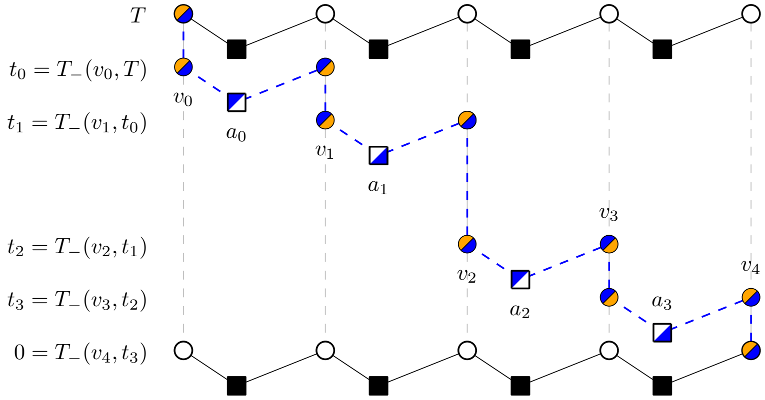
Tracing back the discrepancies

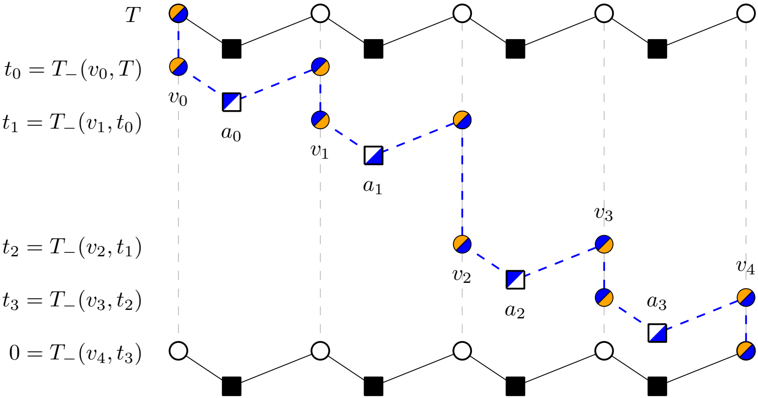
$$t_{\mathsf{coup}} \geqslant T \Rightarrow \exists \sigma, \tau \in \Omega_G, v_0 \in V \text{ such that } X_T^\sigma(v_0) \neq X_T^\tau(v_0).$$



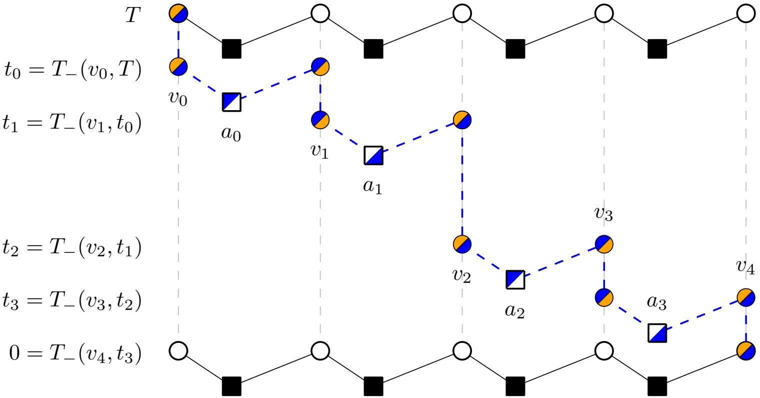
Tracing back the discrepancies

$$t_{\mathsf{coup}} \geqslant T \Rightarrow \exists \sigma, \tau \in \Omega_G, v_0 \in V \text{ such that } X_T^\sigma(v_0) \neq X_T^\tau(v_0).$$



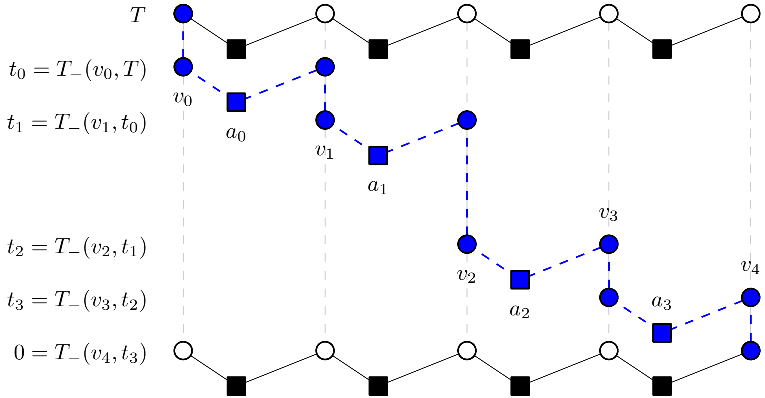


T
$t_0 = T_-(v_0, T)$
$t_1 = T_-(v_1, t_0)$
$t_2 = T_-(v_2, t_1)$
$t_3 = T_-(v_3, t_2)$
$0 = T_-(v_4, t_3)$
v_0
a_0
v_1
a_1
v_2
a_2
v_3
a_3
v_4



$$t_{\mathrm{coup}} \geqslant T \;\Rightarrow\; \exists\big((v_\ell, a_\ell, t_\ell)\big)_{0\leqslant \ell \leqslant L} \text{ such that for all } 0 \leqslant \ell \leqslant L,$$

$$(v_\ell, t_\ell) \in \text{Updates}, \quad t_\ell = T_-(v_\ell, t_{\ell-1}), \quad X^{\sigma}_{t_\ell} \vee X^{\tau}_{t_\ell}(\partial a_\ell \setminus \{v_\ell\}) = \underline{1}$$



$$t_{\text{coup}} \geqslant T \;\Rightarrow\; \exists\big((v_\ell, a_\ell, t_\ell)\big)_{0\leqslant \ell \leqslant L} \text{ such that for all } 0 \leqslant \ell \leqslant L,$$

$$(v_\ell, t_\ell) \in \text{Updates}, \quad t_\ell = T_-(v_\ell, t_{\ell-1}), \quad Y^0_{t_\ell}(\partial a_\ell) = \underline{1}$$

Discrepancy Paths

We say that $((v_\ell, a_\ell, t_\ell))_{0 \leqslant \ell \leqslant L}$ is a **discrepancy path** if

$$(v_\ell, t_\ell) \in \text{Updates}, \quad t_\ell = T_-(v_\ell, t_{\ell-1}), \quad Y^0_{t_\ell}(\partial a_\ell) = \underline{1}$$

It follows that

$$t_{\text{coup}} > T \Rightarrow \exists \textbf{ discrepancy path} \text{ with } t_{-1} = T.$$

Discrepancy Paths

To control the existence of **discrepancy paths** with $t_{-1} = T$:

- Construct a space-time percolation such that every discrepancy path maps to an open path in the percolation.
- Bound the open cluster of the percolation containing $t = 0$.

Discrepancy Paths

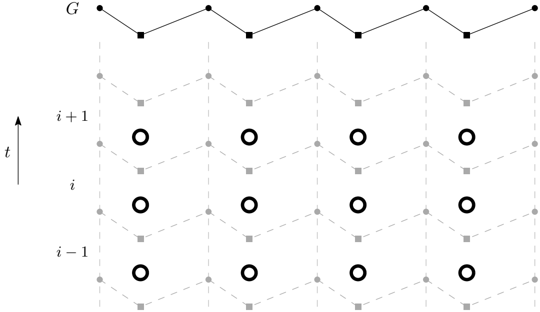
To control the existence of **discrepancy paths** with $t_{-1} = T$:

- Construct a space-time percolation such that every discrepancy path maps to an open path in the percolation.
- Bound the open cluster of the percolation containing $t = 0$.

Will show:
A weaker proof in 10 mins + modifications to the tight result.

Auxillary Percolation

Auxillary Percolation

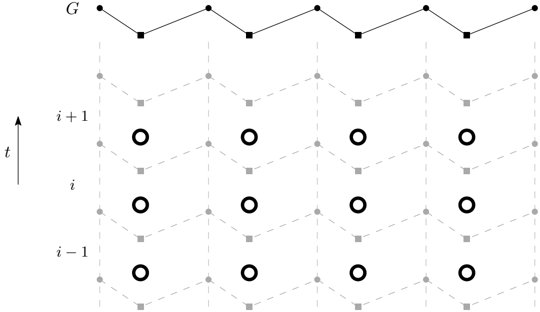


Break the time line into chunks of length $2k$ and define

$$H \equiv F \times \mathbb{Z}_+ \equiv F \times \{0, 1, \dots\}.$$

Each site $(a, i) \in H$ represents the block $\partial a \times [2ik, 2(i+1)k)$.

Auxillary Percolation

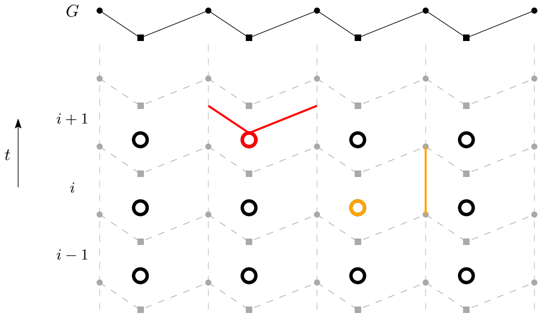


We say that $(a, i) \in H$ is **bad** if $Y_t^{2(i-1)k}(v) = 1$ for

Horizontal: some $t \in [2ik, 2(i+1)k)$ and all $v \in \partial a$.

or **Vertical:** all $t \in [2ik, 2(i+1)k)$ and some $v \in \partial a$.

Auxillary Percolation

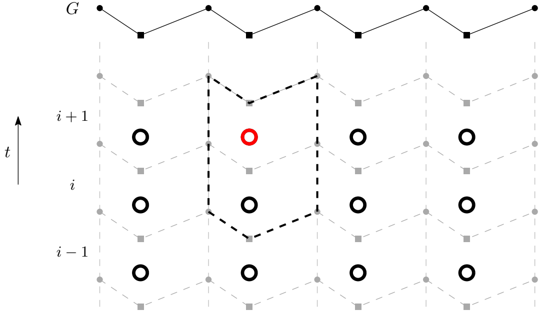


We say that $(a,i) \in H$ is **bad** if $Y_t^{2(i-1)k}(v) = 1$ for

Horizontal: some $t \in [2ik, 2(i+1)k)$ and all $v \in \partial a$.

or **Vertical:** all $t \in [2ik, 2(i+1)k)$ and some $v \in \partial a$.

Auxillary Percolation



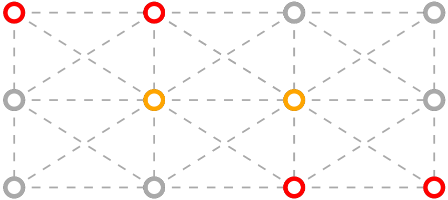
We say that $(a, i) \in H$ is **bad** if $Y_t^{2(i-1)k}(v) = 1$ for

Horizontal: some $t \in [2ik, 2(i+1)k)$ and all $v \in \partial a$.

or **Vertical:** all $t \in [2ik, 2(i+1)k)$ and some $v \in \partial a$.

This definition is local.

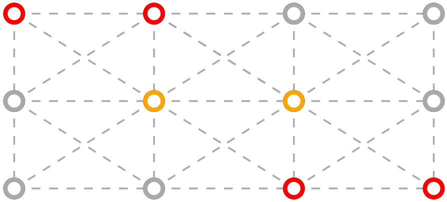
Control the Percolation



Draw an edge between $(a, i), (b, j) \in H$ if and only if

$$\partial a \cap \partial b \neq \varnothing, \quad |i - j| \leqslant 1.$$

Control the Percolation

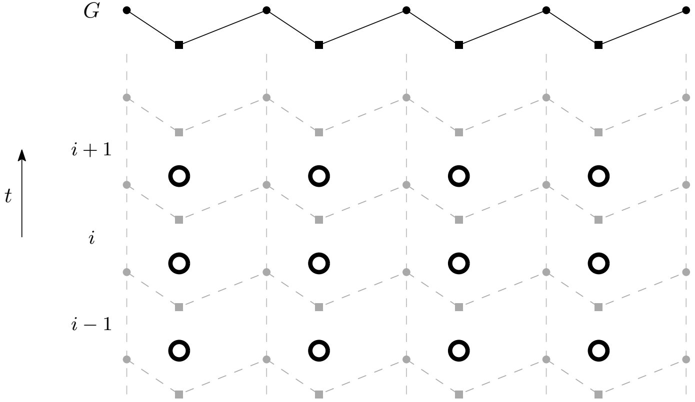


Draw an edge between $(a, i), (b, j) \in H$ if and only if

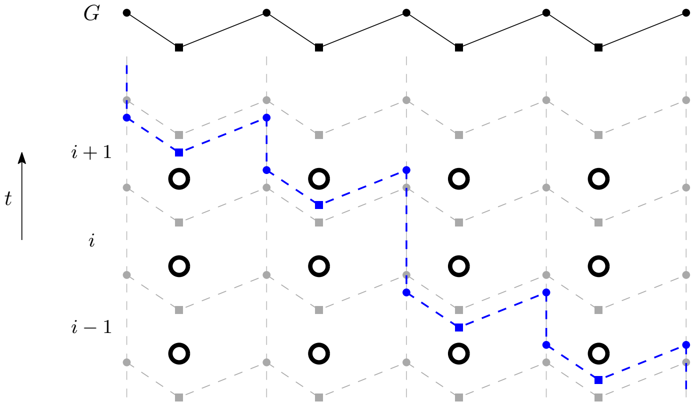
$$\partial a \cap \partial b \neq \varnothing, \quad |i - j| \leqslant 1.$$

Claim: If (a, i) is not adjacent to (b, j), then

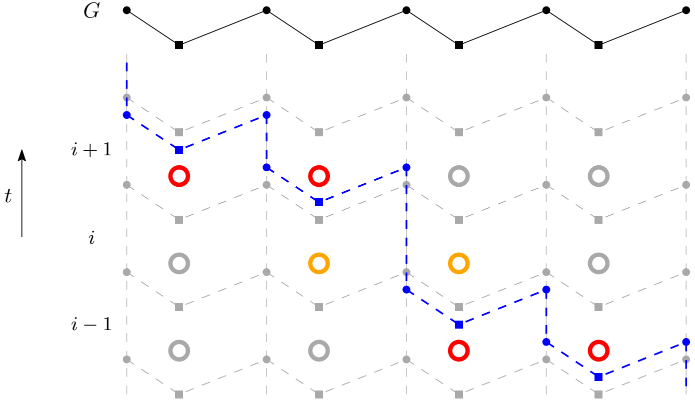
$$\{(a, i) \text{ is } \textbf{bad}\} \perp\!\!\!\perp \{(b, j) \text{ is } \textbf{bad}\}.$$



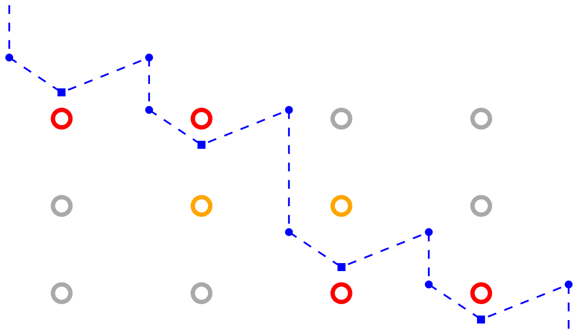
G
i + 1
t
i
i − 1



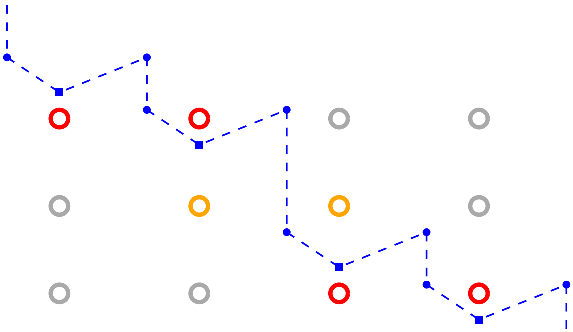
Fix a **discrepancy path**, analyze the sites (a, i)'s along it.



Fix a **discrepancy path**, analyze the sites (a, i)'s along it.

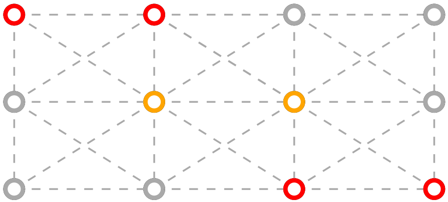


Fix a **discrepancy path**, analyze the sites (a, i)'s along it.
$\Rightarrow$ Each (a, i) is either **bad** or next to some (b, i) that is bad.

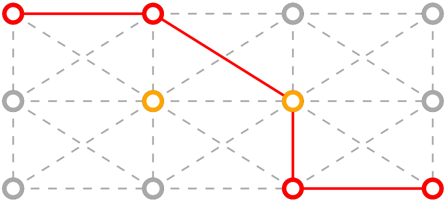


Fix a **discrepancy path**, analyze the sites (a, i)'s along it.
$\Rightarrow$ Each (a, i) is either **bad** or next to some (b, i) that is bad.
$\Rightarrow$ **discrepancy path** $\subseteq$ **open cluster** of the percolation
if we allow the cluster to connect through diagonals.

Control the Percolation



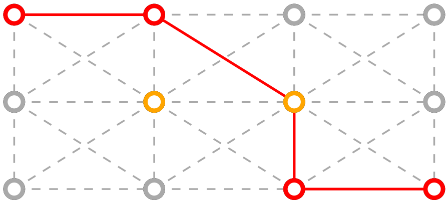
Control the Percolation



We say that a path $\gamma = ((a_\ell, i_\ell))_{0 \leqslant \ell \leqslant L}$ in H is a **minimal path** if (a_ℓ, i_ℓ) is not adjacent to $(a_{\ell'}, i_{\ell'})$ for all $\ell \geqslant \ell' + 2$.

$$\Rightarrow \ \{(a_\ell, i_\ell) \text{ is } \textbf{bad}\} \perp\!\!\!\perp \{(a_{\ell'}, i_{\ell'}) \text{ is } \textbf{bad}\}, \ \forall \ell \geqslant \ell' + 2.$$

Control the Percolation

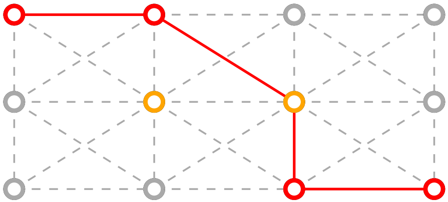


We say that a path $\gamma = ((a_\ell, i_\ell))_{0 \leqslant \ell \leqslant L}$ in H is a **minimal path** if (a_ℓ, i_ℓ) is not adjacent to $(a_{\ell'}, i_{\ell'})$ for all $\ell \geqslant \ell' + 2$.

$$\Rightarrow \ \{(a_\ell, i_\ell) \text{ is } \textbf{bad}\} \perp\!\!\!\perp \{(a_{\ell'}, i_{\ell'}) \text{ is } \textbf{bad}\}, \ \forall \ell \geqslant \ell' + 2.$$

Any cluseter contains a **minimal path** as subset.

Control the Percolation



$t_{\text{coup}} > T$

$\Rightarrow \exists$ **discrepancy path** with $t_{-1} = T$.

$\Rightarrow \exists$ **open cluster** in H connecting $F \times \{0\}$ and $F \times \{\lfloor T/k \rfloor\}$.

$\Rightarrow \exists$ **open minimal path** in H with length $L = \lfloor T/k \rfloor$.

Control the Percolation

Recall that $Y_t(\partial a)$ is just the LSRW on $\{0, 1\}^k$.

$$\Rightarrow \mathbb{P}((a, i) \text{ is } \textbf{bad}) \leqslant 2k^2 2^{-k} + (k+1)e^{-k} \approx 2k^2 2^{-k}.$$

Control the Percolation

Recall that $Y_t(\partial a)$ is just the LSRW on $\{0, 1\}^k$.

$$\Rightarrow \mathbb{P}((a, i) \text{ is } \textbf{bad}) \leqslant 2k^2 2^{-k} + (k+1)e^{-k} \approx 2k^2 2^{-k}.$$

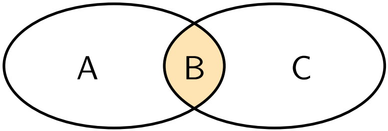
Note that the maximum degree of H is $(2\Delta k + 1)$

$$\begin{aligned}
\mathbb{P}(t_{\text{coup}} > T) &\leqslant \mathbb{P}(\exists \textbf{ open minimal path} \text{ with length } \lfloor T/k \rfloor) \\
&\leqslant \sum_{((a_\ell, i_\ell))_{0 \leqslant \ell \leqslant L}} \prod_{r=0}^{\lfloor L/2 \rfloor} \mathbb{P}((a_{2r}, i_{2r}) \text{ is } \textbf{bad}) \\
&\leqslant n(2\Delta k + 1)^L (2k^2 2^{-k})^{\lfloor L/2 \rfloor} \\
&\to 0, \quad \text{if } \Delta \leqslant O(k^{-2} 2^{k/2}),\ L = \Omega(\log n).
\end{aligned}$$

Removing the factor of k^2

Removing the factor of k^2

For a discrepancy passes to an adjacent hyperedge before being killed by an update, we need:



a vertex in C is updated:

1. when C is all 1;
2. when B is all 1 before each vertex in B is updated once.

Removing the factor of k^2

Recall the key step from previous slide:

$$\mathbb{P}(t_{\text{coup}} \leqslant T) \leqslant Cn(\Delta k)^L (k^2 2^{-k})^{\lfloor L/2 \rfloor}$$

The "branching number" consists of two part:

Removing the factor of k^2

Recall the key step from previous slide:

$$\mathbb{P}(t_{\text{coup}} \leqslant T) \leqslant Cn(\Delta k)^L (k^2 2^{-k})^{\lfloor L/2 \rfloor}$$

The "branching number" consists of two part:

1. The number of adjacent hyperedges —$\leqslant \Delta k$

Removing the factor of k^2

Recall the key step from previous slide:

$$\mathbb{P}(t_{\text{coup}} \leqslant T) \leqslant Cn(\Delta k)^L (k^2 2^{-k})^{\lfloor L/2 \rfloor}$$

The “branching number” consists of two part:

1. The number of adjacent hyperedges —$\leqslant \Delta k$
2. The probability that a discrepancy passes to an adjacent hyperedge before being killed by an update.

Removing the factor of k^2

Recall the key step from previous slide:

$$\mathbb{P}(t_{\text{coup}} \leqslant T) \leqslant Cn(\Delta k)^L (k^2 2^{-k})^{\lfloor L/2 \rfloor}$$

The "branching number" consists of two part:

1. The number of adjacent hyperedges —$\leqslant \Delta k$
2. The probability that a discrepancy passes to an adjacent hyperedge before being killed by an update.
   —i.e. an vertex in C is updated —$\leqslant k$

Removing the factor of k^2

Recall the key step from previous slide:

$$\mathbb{P}(t_{\text{coup}} \leqslant T) \leqslant Cn(\Delta k)^L (k^2 2^{-k})^{\lfloor L/2 \rfloor}$$

The "branching number" consists of two part:

1. The number of adjacent hyperedges —$\leqslant \Delta k$
2. The probability that a discrepancy passes to an adjacent hyperedge before being killed by an update.
   —i.e. an vertex in C is updated —$\leqslant k$

Removing the factor of k^2

Recall the key step from previous slide:

$$\mathbb{P}(t_{\text{coup}} \leqslant T) \leqslant Cn(\Delta k)^L (k^2 2^{-k})^{\lfloor L/2 \rfloor}$$

The "branching number" consists of two part:

1. The number of adjacent hyperedges —$\leqslant \Delta k$
2. The probability that a discrepancy passes to an adjacent hyperedge before being killed by an update.
   —i.e. an vertex in C is updated —$\leqslant k$
   a. when C is all one; —avg. $2^{-k/2}$

Removing the factor of k^2

Recall the key step from previous slide:

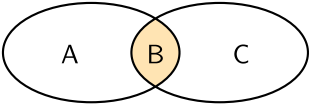
$$\mathbb{P}(t_{\text{coup}} \leqslant T) \leqslant Cn(\Delta k)^L (k^2 2^{-k})^{\lfloor L/2 \rfloor}$$

The "branching number" consists of two part:

1. The number of adjacent hyperedges —$\leqslant \Delta k$
2. The probability that a discrepancy passes to an adjacent hyperedge before being killed by an update.
   —i.e. an vertex in C is updated —$\leqslant k$
   a. when C is all one; —avg. $2^{-k/2}$
   b. before each vertex in B is updated once. —$\leqslant k$

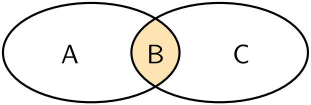
Removing the factor of k^2

Assume that adj. hyperedges share $|B| = m \leqslant k/2$ vertices.



Removing the factor of k^2

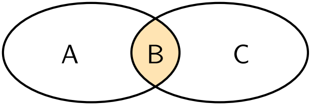
Assume that adj. hyperedges share $|B| = m \leqslant k/2$ vertices.



1. The number of adjacent hyperedges —$\Delta k/m$

Removing the factor of k^2

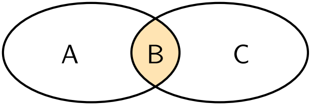
Assume that adj. hyperedges share $|B| = m \leqslant k/2$ vertices.



1. The number of adjacent hyperedges —$\Delta k/m$
2. An vertex in C is updated —$(k - m)$

Removing the factor of k^2

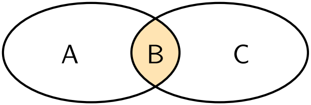
Assume that adj. hyperedges share $|B| = m \leqslant k/2$ vertices.



1. The number of adjacent hyperedges —$\Delta k/m$
2. An vertex in C is updated —$(k-m)$
   a. when C is all one; —$\approx 2^{-(k-m)}$

Removing the factor of k^2

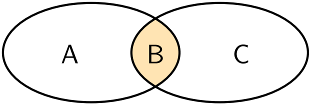
Assume that adj. hyperedges share $|B| = m \leqslant k/2$ vertices.



1. The number of adjacent hyperedges —$\Delta k/m$
2. An vertex in C is updated —$(k-m)$
   a. when C is all one; —$\approx 2^{-(k-m)}$
   b. when B is all one and before every vertex in B is updated at least once. —$\approx m^{-1}$

Removing the factor of k^2

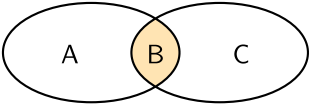
Assume that adj. hyperedges share $|B| = m \leqslant k/2$ vertices.



1. The number of adjacent hyperedges —$\Delta k/m$
2. An vertex in C is updated —$(k - m)$
   a. when C is all one; —$\approx 2^{-(k-m)}$
   b. when B is all one and before every vertex in B is updated at least once. —$\approx m^{-1}$
   (which corresponds to the time a SRW on hypercube $\{0, 1\}^m$ stays at $(1, \ldots, 1)$ before mixing.)

Removing the factor of k^2

Assume that adj. hyperedges share $|B| = m \leqslant k/2$ vertices.

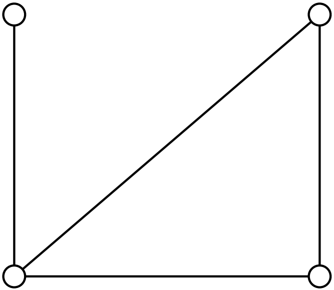


1. The number of adjacent hyperedges —$\Delta k/m$
2. An vertex in C is updated —$(k - m)$
   a. when C is all one; —$\approx 2^{-(k-m)}$
   b. when B is all one and before every vertex in B is updated at least once. —$\approx m^{-1}$
   (which corresponds to the time a SRW on hypercube $\{0, 1\}^m$ stays at $(1, \ldots, 1)$ before mixing.)

The branching factor is $\max_{m \leqslant k/2} \Delta \frac{k(k-m)}{m^2} 2^{-(k-m)} \lesssim \Delta 2^{-k/2}$.

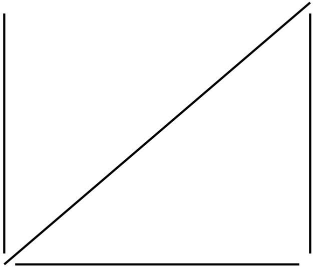
Worst Case

The worst case $m = k/2$ corresponds to reduction construction in the proof of hardness.



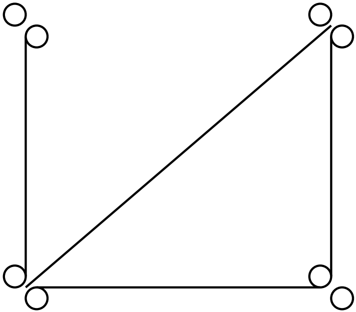
Worst Case

The worst case $m = k/2$ corresponds to reduction construction in the proof of hardness.



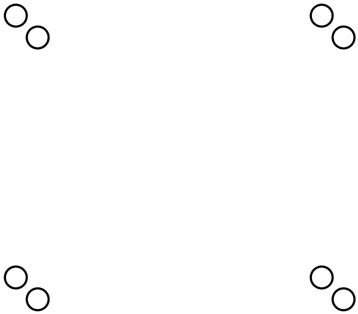
Worst Case

The worst case $m = k/2$ corresponds to reduction construction in the proof of hardness.



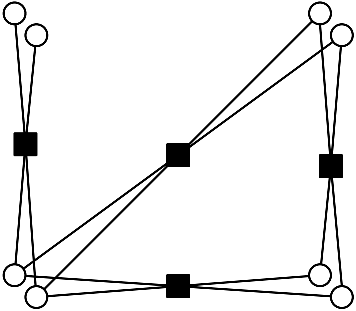
Worst Case

The worst case $m = k/2$ corresponds to reduction construction in the proof of hardness.



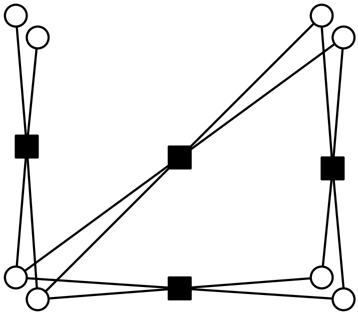
Worst Case

The worst case $m = k/2$ corresponds to reduction construction in the proof of hardness.



Worst Case

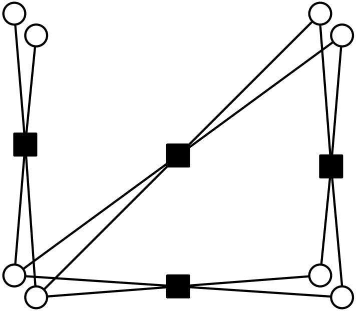
The worst case $m = k/2$ corresponds to reduction construction in the proof of hardness.



In fact, restricting to linear hypergraphs (i.e. max overlap 1) improves the bound to $\Delta \leqslant 2^k/\text{poly}(k)$

Worst Case

The worst case $m = k/2$ corresponds to reduction construction in the proof of hardness.



Meanwhile, in [GJL17] the best case is achieved when the minimum overlap is $\Omega(k)$.

Open problem

There are three results uploaded to ArXiv around the same time last year.

	Monotone	Min overlap	Δ	Type
[Moitra17]	No	No	$\leqslant e^{ck}$	FPTAS
[GJL17]	No	Yes	$\leqslant 2^{k/2}$	Exact
[HSZ17]	Yes	No	$\leqslant 2^{k/2}$	FPRAS

Any chance combining the results?

Open problem

1. Any "physical" interperatation to the threshold $O(2^{k/2})$?
2. How does the mixing time/sampling complexity changes when we impose extra restrictions to G?
   We also showed that rapid mixing holds for
   linear hypergraphs with $\Delta \leqslant O(2^k/k^3)$.
   random regular graph with $\Delta \leqslant O(2^k/k) = O(\mathrm{WSM})$.
3. General models?

Thank you!