Mixing time of the swap Markov chain and P-stability of degree sequences

Péter L. Erdős, Catherine S. Greenhill, Ervin Győri, István Miklós, Tamás Róbert Mezei, Dániel Soltész, Lajos Soukup

Bratislava, Aug 26-30, EUROCOMB 2019

Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences
https://renyi.hu/?en

Personal homepage: https://trm.hu

- *Problem*: given a non-negative integer sequence d of even sum, generate a graph $G \in \mathcal{G}(d)$ with degree sequence d, uniformly at random (labeled vertices)
- *Motivation*:
    - network science: generating graphs from a null model for hypothesis testing
    - testing software, algorithms
    - simulations

Possible ways to sample $\mathcal{G}(d)$

- Enumerate elements of $\mathcal{G}(d)$: the set is huge (exponential in n), generally not feasible

Possible ways to sample $\mathcal{G}(d)$

- Enumerate elements of $\mathcal{G}(d)$: the set is huge (exponential in n), generally not feasible
- *Stub pairing* (configuration model)

Possible ways to sample $\mathcal{G}(d)$

- Enumerate elements of $\mathcal{G}(d)$: the set is huge (exponential in n), generally not feasible

- *Stub pairing* (configuration model)
    - the probability of a multiedge or loop appearing tends to 1 exponentially quickly for regular graphs of degree $\Omega((\log n)^{\frac{1}{2}+\varepsilon})$.
      Booster shot: Rejection schemes (eg. Wormald et al.)

Possible ways to sample $\mathcal{G}(d)$

- Enumerate elements of $\mathcal{G}(d)$: the set is huge (exponential in n), generally not feasible
- *Stub pairing* (configuration model)
  - the probability of a multiedge or loop appearing tends to 1 exponentially quickly for regular graphs of degree $\Omega((\log n)^{\frac{1}{2}+\varepsilon})$.
    Booster shot: Rejection schemes (eg. Wormald et al.)
  - Importance sampling by Bliztstein and Diaconis: the distribution is known but not uniform, unknown variance (quality of the sample is unknown)

Possible ways to sample $\mathcal{G}(d)$

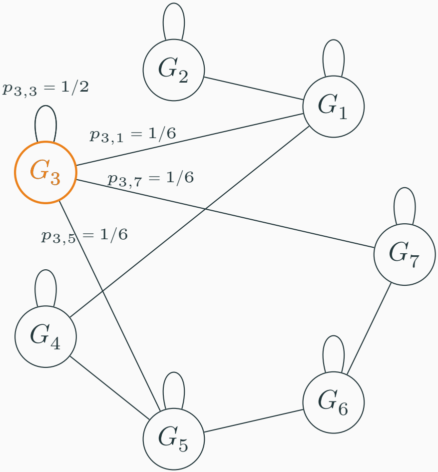
- Enumerate elements of $\mathcal{G}(d)$: the set is huge (exponential in n), generally not feasible
- *Stub pairing* (configuration model)
  - the probability of a multiedge or loop appearing tends to 1 exponentially quickly for regular graphs of degree $\Omega((\log n)^{\frac{1}{2}+\varepsilon})$.
    Booster shot: Rejection schemes (eg. Wormald et al.)
  - Importance sampling by Bliztstein and Diaconis: the distribution is known but not uniform, unknown variance (quality of the sample is unknown)
- Monte Carlo Markov Chain (MCMC) methods $\Rightarrow$

Markov Chains - a reminder

Our chains transition from state i to state j with some probability $p_{i,j} = p_{j,i}$ (symmetric), independently of time and previous steps

$\forall i \ \sum_j p_{i,j} = 1$

$t = 1$

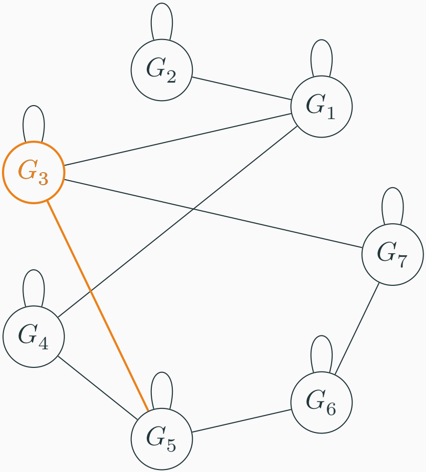


Markov Chains - a reminder

Our chains transition from state i to state j with some probability $p_{i,j} = p_{j,i}$ (symmetric), independently of time and previous steps

$\forall i \; \sum_j p_{i,j} = 1$

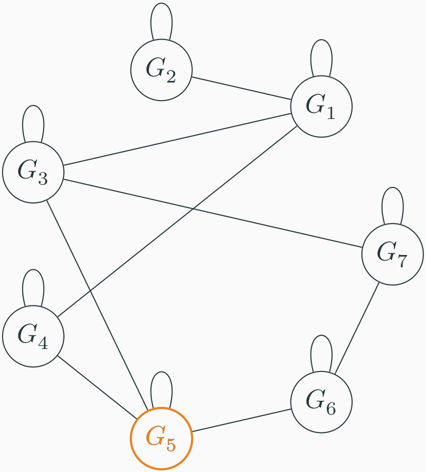
$t = 1$



Our chains transition from state i to state j with some probability $p_{i,j} = p_{j,i}$ (symmetric), independently of time and previous steps

$\forall i \ \sum_j p_{i,j} = 1$

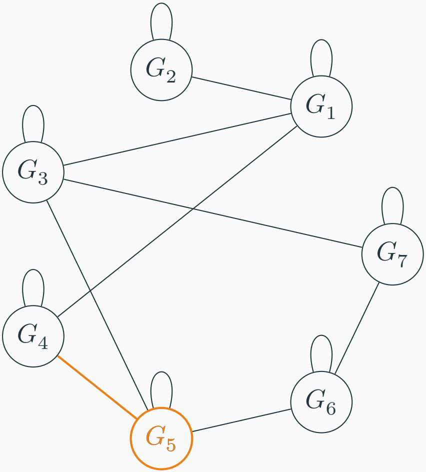
$t = 2$



Our chains transition from state i to state j with some probability $p_{i,j} = p_{j,i}$ (symmetric), independently of time and previous steps

$\forall i \ \sum_j p_{i,j} = 1$

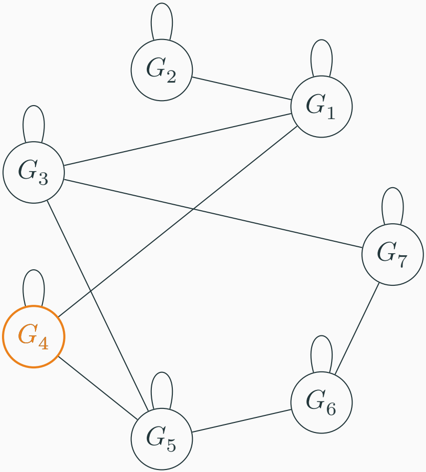
$t = 2$



Our chains transition from state i to state j with some probability $p_{i,j} = p_{j,i}$ (symmetric), independently of time and previous steps

$\forall i \ \sum_j p_{i,j} = 1$

$t = 3$

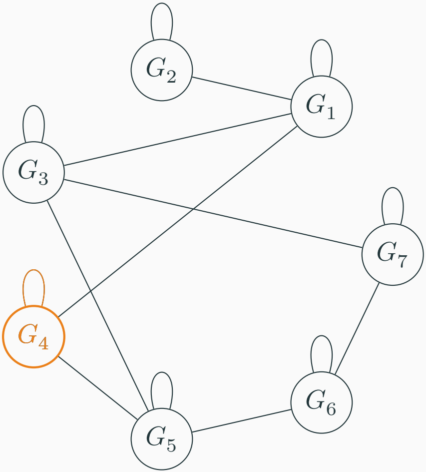


Markov Chains - a reminder

Our chains transition from state i to state j with some probability $p_{i,j} = p_{j,i}$ (symmetric), independently of time and previous steps

$\forall i \ \sum_j p_{i,j} = 1$

$t = 3$

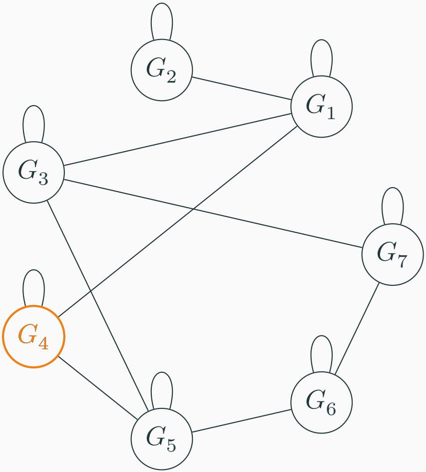


Markov Chains - a reminder

Our chains transition from state i to state j with some probability $p_{i,j} = p_{j,i}$ (symmetric), independently of time and previous steps

$\forall i \ \sum_j p_{i,j} = 1$

$t = 4$

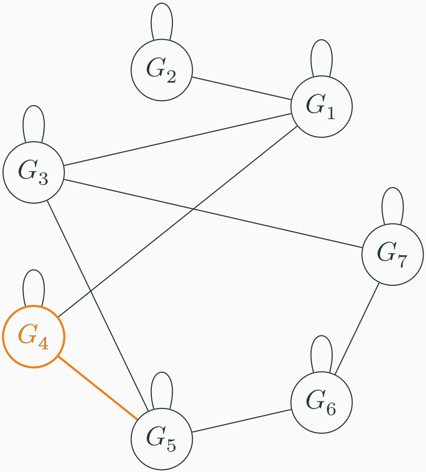


Markov Chains - a reminder

Our chains transition from state i to state j with some probability $p_{i,j} = p_{j,i}$ (symmetric), independently of time and previous steps

$\forall i \ \sum_j p_{i,j} = 1$

$t = 4$

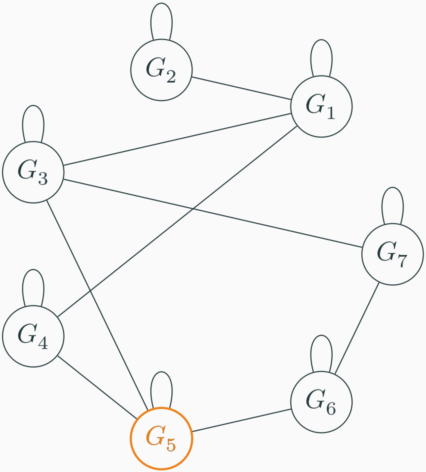


Markov Chains - a reminder

Our chains transition from state i to state j with some probability $p_{i,j} = p_{j,i}$ (symmetric), independently of time and previous steps

$\forall i \ \sum_j p_{i,j} = 1$

$t = 5$



- If the Markov-chain is irreducible, symmetric, and aperiodic then the MC converges to the uniform distribution
- Instead of exact, only require *approximate* sampling: the sampled distribution is ε close to the uniform distribution in variation (ℓ_1-)distance in $\mathrm{poly}(n) \cdot \log \varepsilon^{-1}$ steps (rapidly mixing)

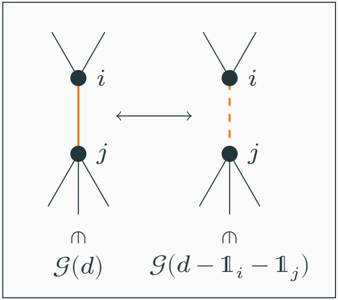
State space: $\mathcal{G}(d) \cup \bigcup_{i,j \in V} \mathcal{G}(d - \mathbb{1}_i - \mathbb{1}_j)$.

Transitions: u.a.r. choose $a, b \in V(G)$, then

State space: $\mathcal{G}(d) \cup \bigcup_{i,j \in V} \mathcal{G}(d - \mathbb{1}_i - \mathbb{1}_j)$.

Transitions: u.a.r. choose $a, b \in V(G)$, then

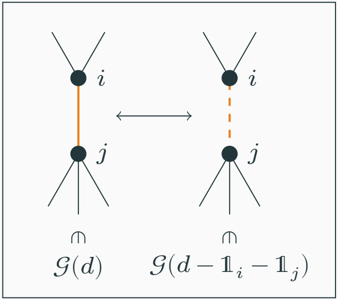
- if $G \in \mathcal{G}(d)$, delete $ab \in E(G)$ if it exists.

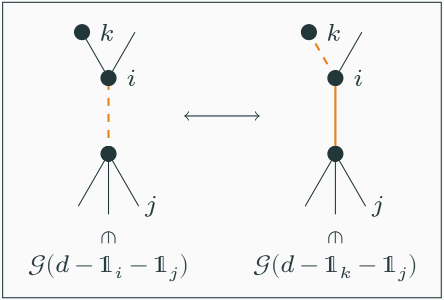


State space: $\mathcal{G}(d) \cup \bigcup_{i,j \in V} \mathcal{G}(d - \mathbb{1}_i - \mathbb{1}_j)$.

Transitions: u.a.r. choose $a, b \in V(G)$, then

- if $G \in \mathcal{G}(d)$, delete $ab \in E(G)$ if it exists.
- if $G \in \mathcal{G}(d - \mathbb{1}_i - \mathbb{1}_j)$ and $\deg_G(a) < d(a)$, try to add ab to $E(G)$. If $\deg_{G+ab}(b) > d(b)$, then delete u.a.r. an edge of b.





The state space of JS chain: $\mathcal{G}(d) \cup \bigcup_{i,j \in V} \mathcal{G}(d - \mathbb{1}_i - \mathbb{1}_j)$

To get a sample from $\mathcal{G}(d)$ in reasonable time, we must have (where $n = \dim(d)$)

$$\frac{\sum_{i,j} |\mathcal{G}(d - \mathbb{1}_i - \mathbb{1}_j)|}{|\mathcal{G}(d)|} \leq \mathrm{poly}(n) \quad \forall d \in \mathcal{D}.$$

In this case, we call $\mathcal{D}$ a P-*stable* class of degree sequences.

Theorem (Jerrum and Sinclair 1990)

The JS chain is rapidly mixing on degree sequences from a P*-stable class.*

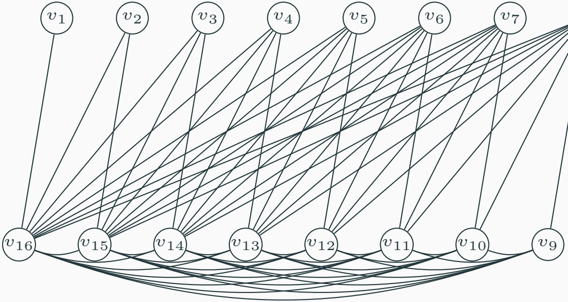
Example for a P-stable region

Theorem (Jerrum and Sinclair 1992)

The class of degree sequences satisfying $(\Delta - \delta + 1)^2 \leq 4\delta(n - \Delta + 1)$ is P-stable.

A (seemingly pathological) obstacle

$n = 8$



$$h_0(n) = (1, 2, \dots, n-1, n, n, n+1, \dots, 2n-1)$$

has a unique realization

$\{h_0(n) \mid n \in \mathbb{N}\}$ not P-stable: $|\mathcal{G}(h_0(n) - \mathbb{1}_n - \mathbb{1}_{2n})| \approx \left(\frac{3+\sqrt{5}}{2}\right)^n$

Can be blown up to a non-pathological non-P-stable class.

A (SEEMINGLY PATHOLOGICAL) OBSTACLE

$$h_0(n) = (1, 2, \dots, n-1, n, n, n+1, \dots, 2n-1)$$

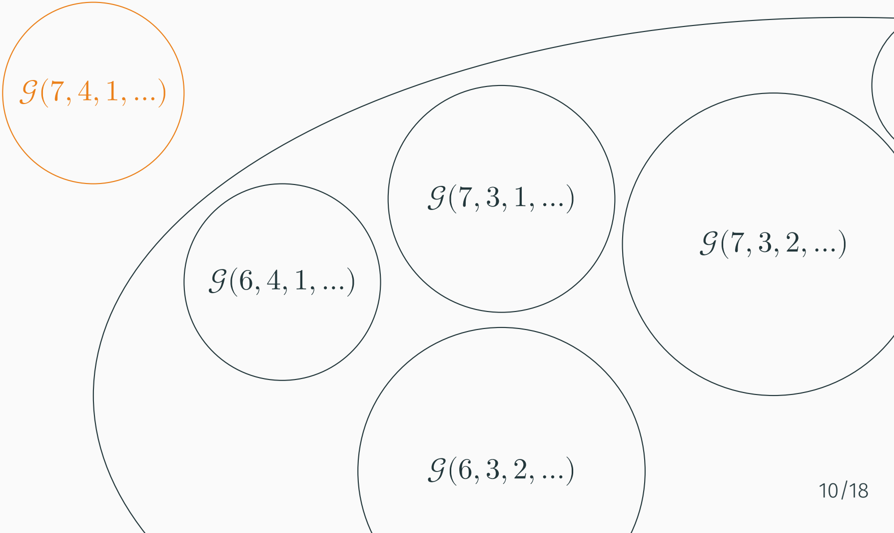
has a unique realization

$\{h_0(n) \mid n \in \mathbb{N}\}$ not P-stable: $|\mathcal{G}(h_0(n) - \mathbb{1}_n - \mathbb{1}_{2n})| \approx \left(\frac{3+\sqrt{5}}{2}\right)^n$

Can be blown up to a non-pathological non-P-stable class.

Applicability remarks

The cardinality of the state space of the JS chain can easily be a factor of n^8 larger than $\mathcal{G}(d)$.

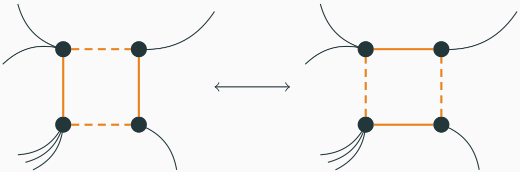


Proposed by Kannan, Tetali, Vempala (1997)

State space: *only* the set of realizations $\mathcal{G}(d)$ of a deg. seq. d

Transitions: exchange edges with non-edges along a randomly chosen alternating C_4 (least perturbation)

swap:



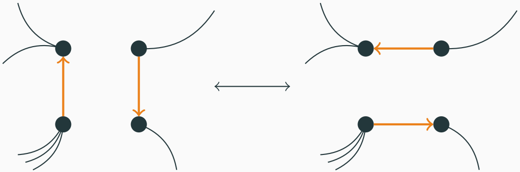
The Swap (Switch) Markov-chain

Proposed by Kannan, Tetali, Vempala (1997)

State space: *only* the set of realizations $\mathcal{G}(d)$ of a deg. seq. d

Transitions: exchange edges with non-edges along a randomly chosen alternating C_4 (least perturbation)

swap:



directed $\triangle$:

Previous results on the Swap chain

Rapid mixing of the Swap Markov chain shown by

	simple	*bipartite*	*directed*
regular	Cooper, Dyer, Greenhill 2007	Erdős et al. 2013	Greenhill 2011
$\Delta \leq c\sqrt{m}$	Greenhill and Sfragara 2018	Erdős, Miklós, M, Soltész 2018	
Interval	—	Erdős, Miklós, M, Soltész 2018 $(\Delta - \delta)^2 \leq \delta(n - \Delta)$	similar
strongly stable	Amanatidis and Kleer 2019		—

Theorem (Greenhill, Erdős, Miklós, M, Soltész, Soukup 2019+)

The swap Markov-chain is rapidly mixing on P-stable degree sequences (unconstrained, bipartite, directed)

- Proof: complex (based on the Jerrum-Sinclair method)
- Every previously known rapidly mixing region is P-stable
- Gao and Wormald (2016) describe several P-stable regions, including power-law distribution-bounded degree sequences for $\gamma > 1 + \sqrt{3}$
- Power-law degree sequences with $\gamma > 2$ are also conjectured to be P-stable

Beyond P-stability...?

P-stability is not necessary for rapid mixing

For all $n, k \in \mathbb{Z}^+$, let us define the bipartite degree sequence

$$h_k(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n-k \\ n-k & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

Theorem (Erdős, Győri, M, Milós, Soltész 2019+)

For any $k \in \mathbb{Z}^+$, the swap Markov chain is rapidly mixing on

$$\mathcal{H}_k := \Big\{ h_k(n) \ : \ n \geq k \Big\},$$

even though the class is not P-stable:

$$\frac{|\mathcal{G}(h_{k+1}(n))|}{|\mathcal{G}(h_k(n))|} = e^{\Omega_k(n)}$$

Remark: the proof works up to $k \leq c\sqrt{\log n}$ for some c.

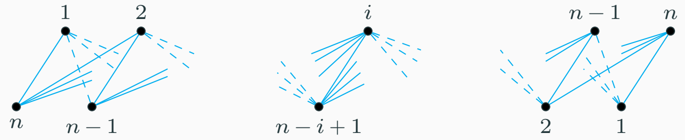
Proof of rapid mixing for $k = 1$; geometric representation

$$h_0(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n \\ n & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

$$h_1(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n-1 \\ n-1 & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

Suppose $G \in \mathcal{G}(h_1(n))$. What does $H_0(n) \triangle G$ look like?

$H_0(n)$

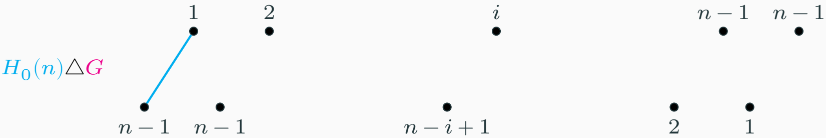


Proof of rapid mixing for $k = 1$; geometric representation

$$h_0(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n \\ n & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

$$h_1(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n-1 \\ n-1 & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

Suppose $G \in \mathcal{G}(h_1(n))$. What does $H_0(n) \triangle G$ look like?

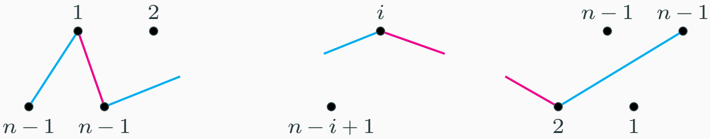


$$h_0(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n \\ n & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

$$h_1(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n-1 \\ n-1 & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

Suppose $G \in \mathcal{G}(h_1(n))$. What does $H_0(n) \triangle G$ look like?

$H_0(n) \triangle G$



$H_0(n) \triangle G$ is an x-monotone path!

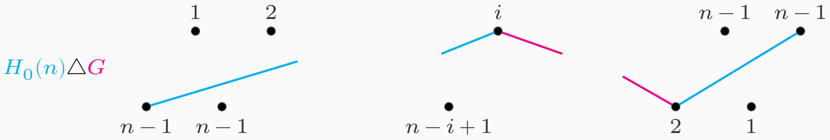
Swap in this representation: moves a vertex of the path or deletes/inserts a pair of adjacent vertices

Proof of rapid mixing for $k = 1$; geometric representation

$$h_0(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n \\ n & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

$$h_1(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n-1 \\ n-1 & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

Suppose $G \in \mathcal{G}(h_1(n))$. What does $H_0(n) \triangle G$ look like?



$H_0(n) \triangle G$ is an x-monotone path!

Swap in this representation: moves a vertex of the path or deletes/inserts a pair of adjacent vertices

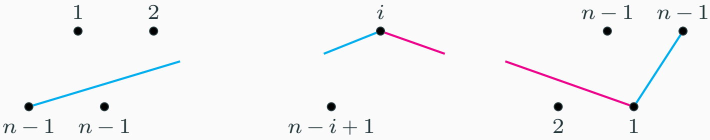
Proof of rapid mixing for $k = 1$; geometric representation

$$h_0(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n \\ n & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

$$h_1(n) := \begin{pmatrix} 1 & 2 & 3 & \cdots & n-2 & n-1 & n-1 \\ n-1 & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix}$$

Suppose $G \in \mathcal{G}(h_1(n))$. What does $H_0(n) \triangle G$ look like?

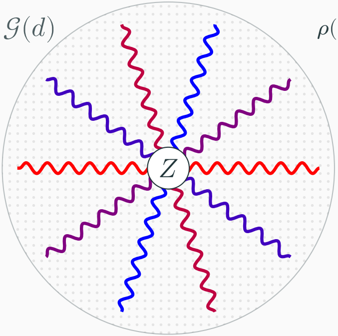
$H_0(n) \triangle G$



$H_0(n) \triangle G$ is an x-monotone path!

Swap in this representation: moves a vertex of the path or deletes/inserts a pair of adjacent vertices

Let Γ contain a swap sequence $X = Z_0^{X,Y}, Z_1^{X,Y}, Z_2^{X,Y}, \ldots, Z_\ell^{X,Y} = Y$ for each pair of realizations $X, Y \in \mathcal{G}(d)$.



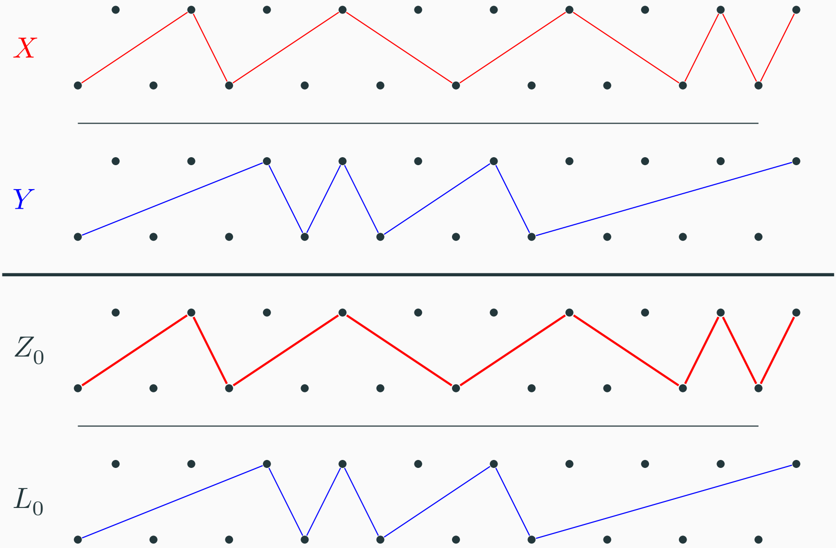
$\rho(\Gamma)$: max number of sequences through a Z

$\ell(\Gamma)$: max length of a sequence in Γ

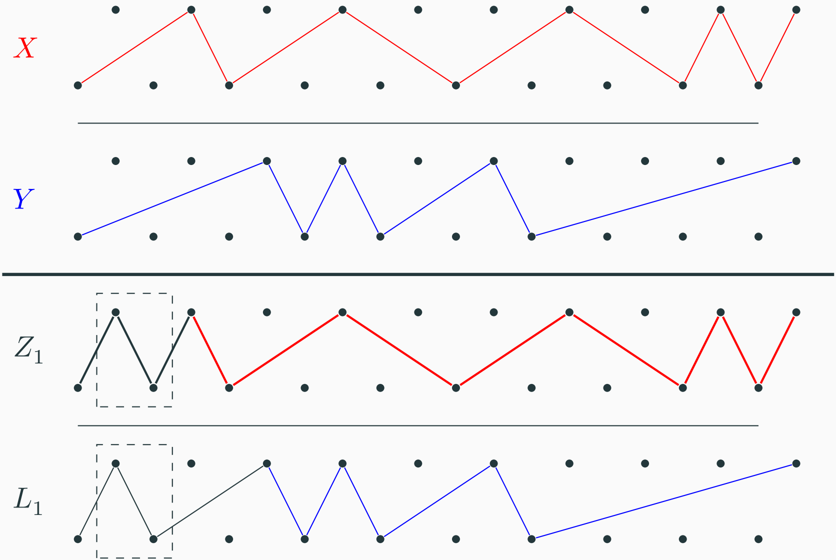
Theorem (follows from Jerrum and Sinclair 1990)

$$\tau_{swap}(\varepsilon) \leq \mathrm{poly}(n) \cdot \frac{\rho(\Gamma)}{|\mathcal{G}(d)|} \cdot \ell(\Gamma) \cdot \left(\log(|\mathcal{G}(d)|) + \log(\varepsilon^{-1})\right)$$

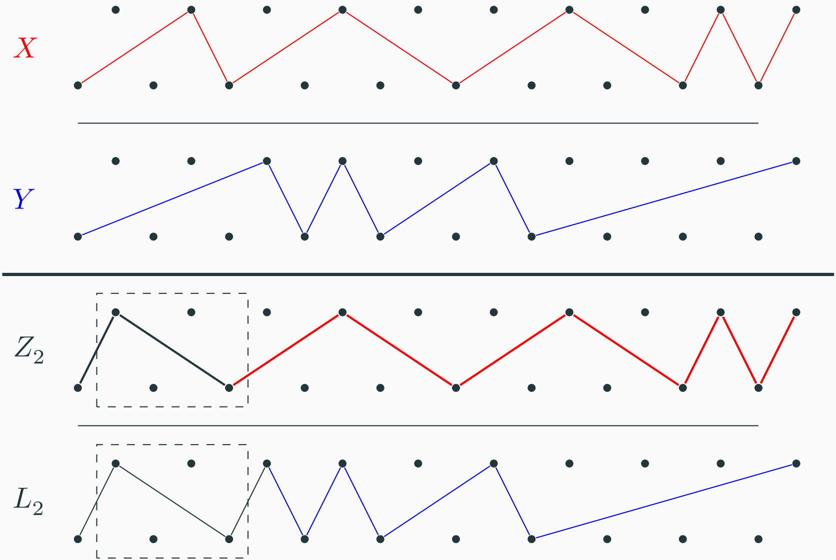
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



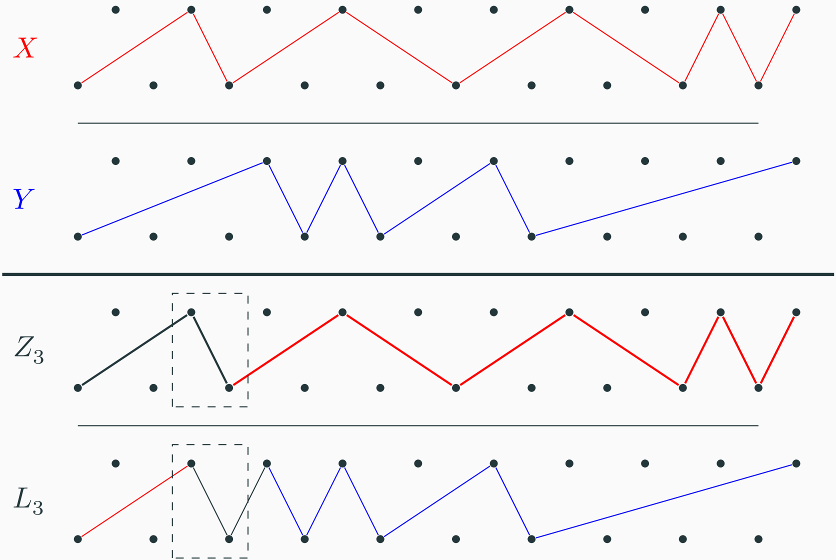
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



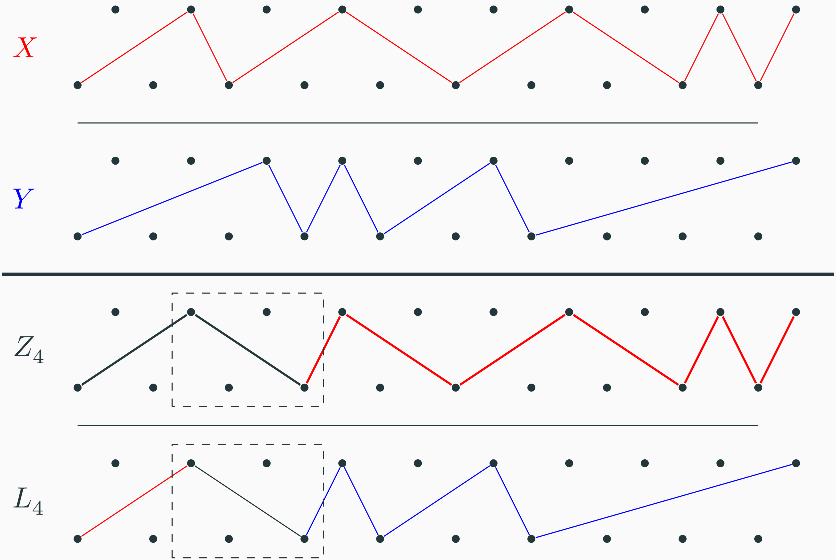
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



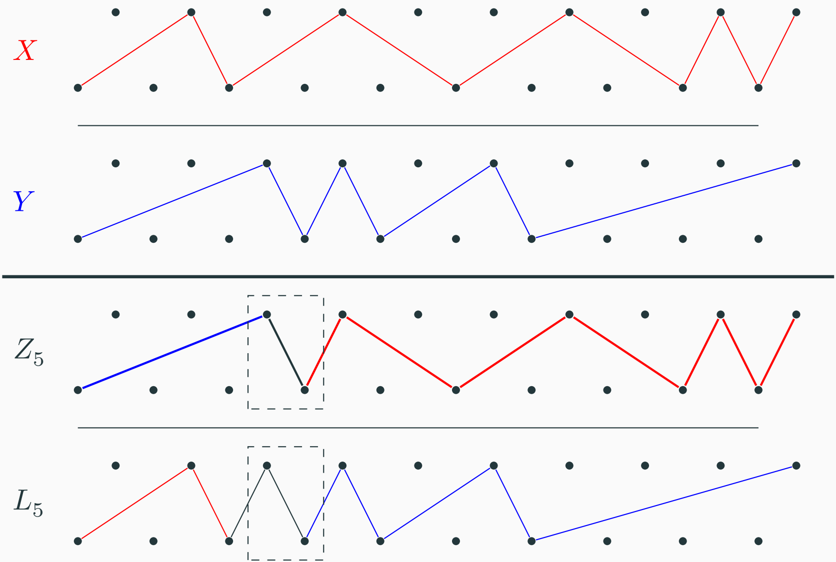
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



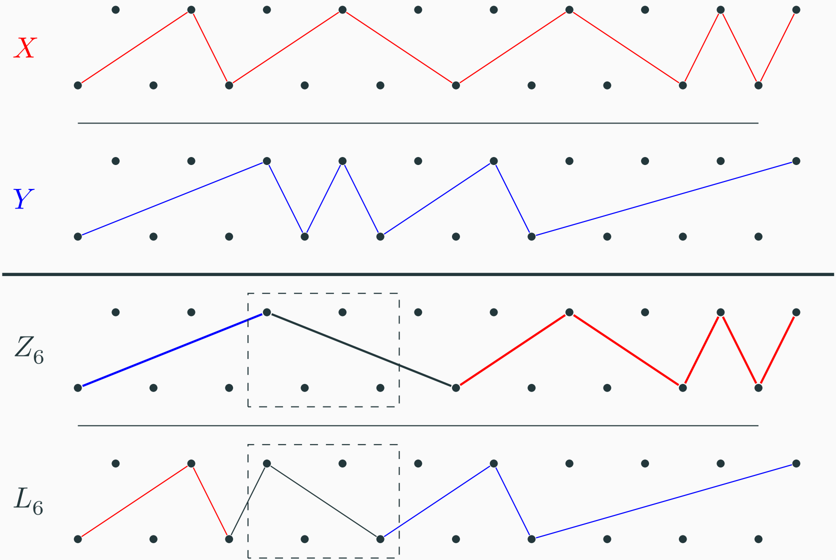
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



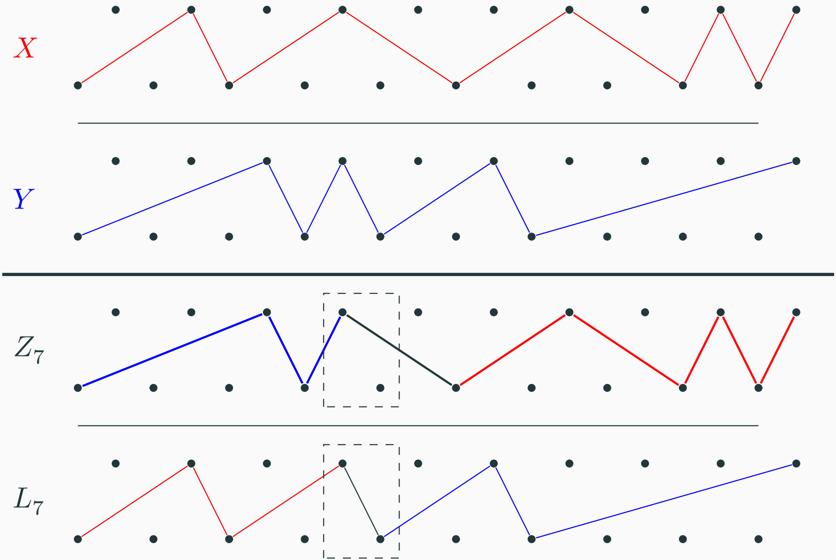
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



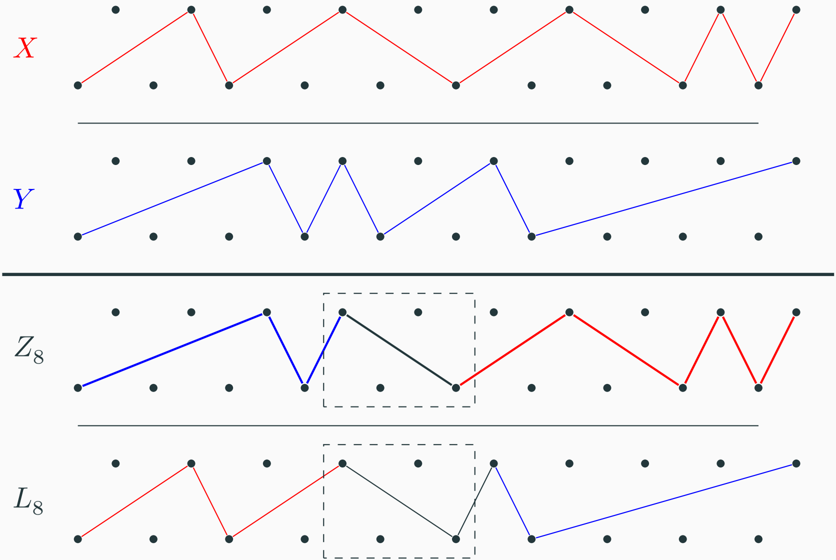
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



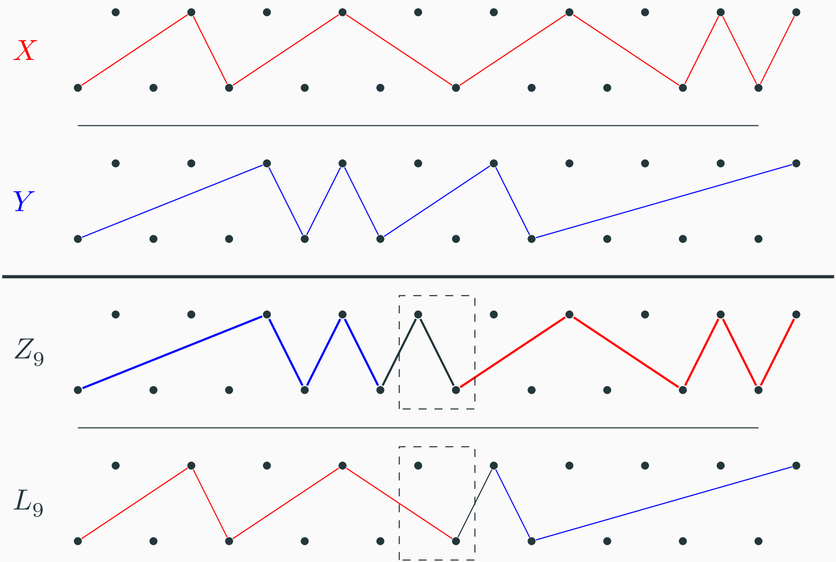
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



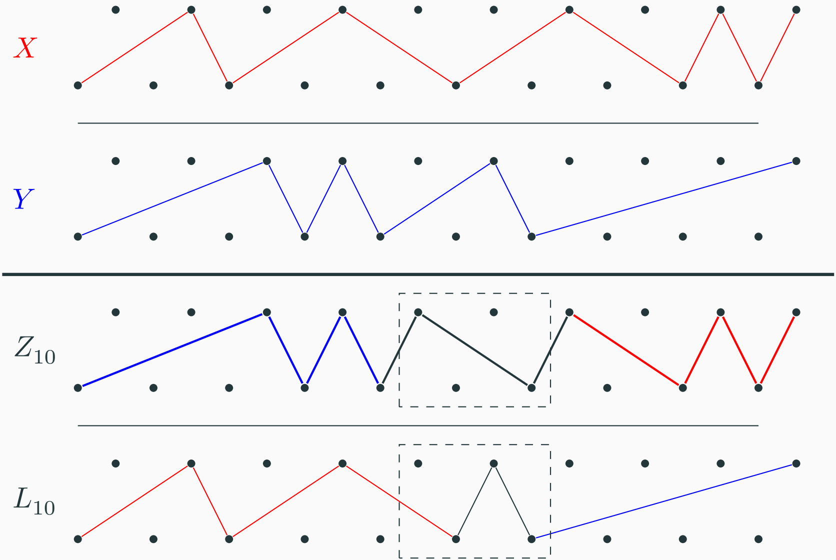
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



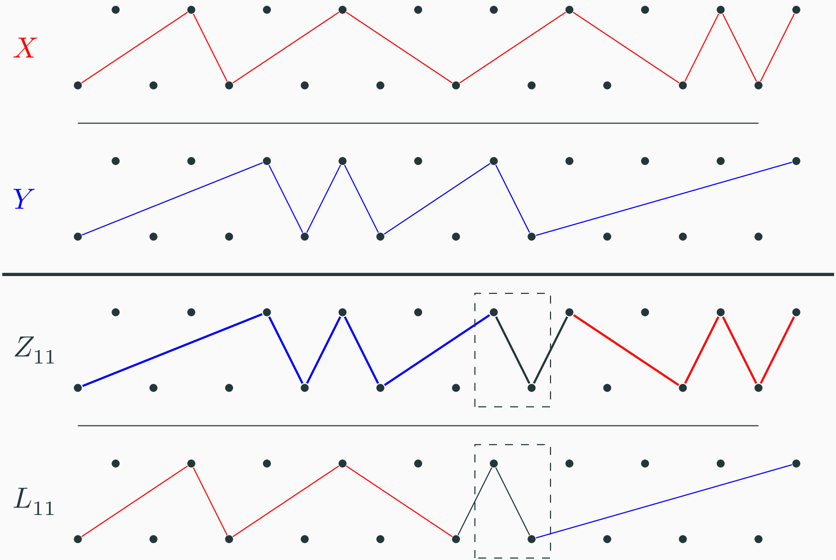
Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



Swap sequence between $X, Y \in \mathcal{G}(h_1(n))$



Clearly, $\ell(\Gamma) = \mathcal{O}(n)$.

From Z_i and L_i and a $\mathcal{O}(\log n)$ bits we can recover X and Y!

$\Longrightarrow$ For a fix $Z \in \mathcal{G}(d)$, the number of swap sequences of Γ passing through Z is at most the number of possible L_i times $\mathrm{poly}(n)$!

$\Longrightarrow \rho(\Gamma) = \mathrm{poly}(n) \cdot |\mathcal{G}(h_1(n))| \overset{\text{Jerrum-Sinclair}}{\Longrightarrow}$ Swap MC is rapidly mixing on $\mathcal{G}(h_1(n))$!

Thank you for listening to my presentation!

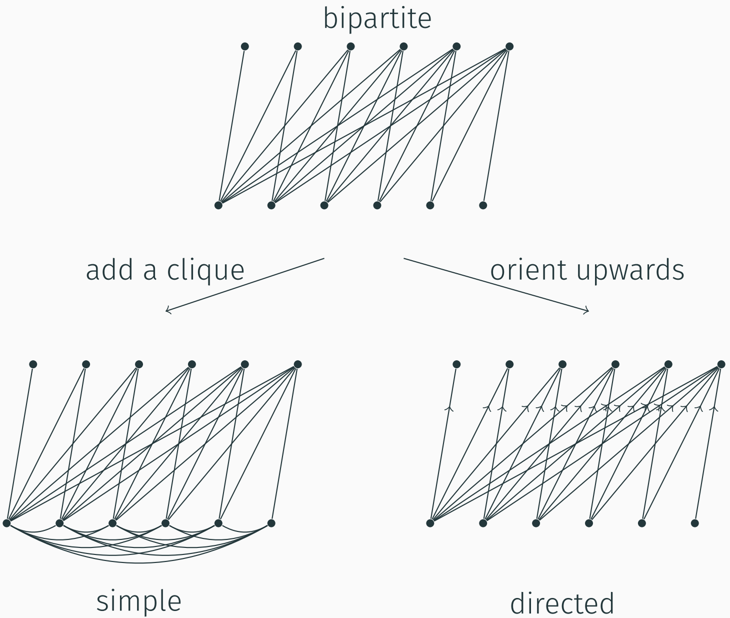
Homepage: `https://trm.hu`

Full papers

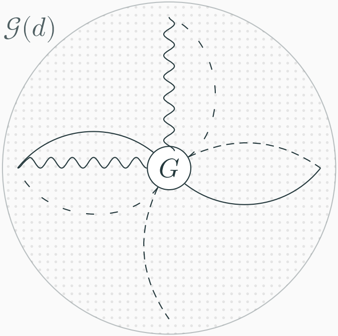
Unified approach:
`https://arxiv.org/abs/1903.06600`

Beyond P-stability: coming soon.

The simple and directed analogues follow immediately



Let f be a multicommodity-flow that sends 1 quantity of commodity between each two realizations in the swap graph on $\mathcal{G}(d)$.



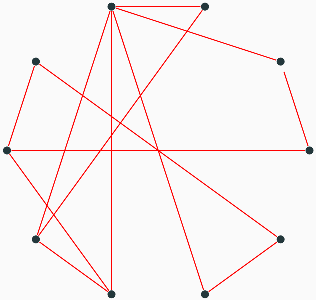
Theorem (follows from Jerrum and Sinclair 1990)

$$\tau_{swap}(\varepsilon) \leq n^4 \cdot \max_{G \in \mathcal{G}(d)} \sum_{G \in \gamma} \frac{f(\gamma)|\gamma|}{|\mathcal{G}(d)|} \cdot (\log(|\mathcal{G}(d)|) + \log(\varepsilon^{-1}))$$

Designing the multicommodity-flow

- Let $G_1, G_2 \in \mathcal{G}(d) \Rightarrow$
  $E(G_1) \triangle E(G_2)$ is a balanced red-blue graph
- Fix an alternating closed trail on $E(G_1) \triangle E(G_2)$
- Crucial decision: decomposing the alternating trail into "simpler" alternating circuits

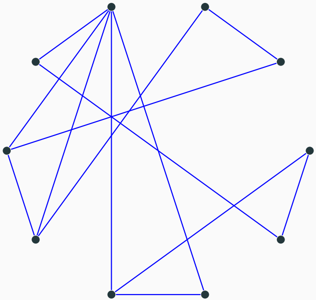
G_1



Designing the multicommodity-flow

- Let $G_1, G_2 \in \mathcal{G}(d) \Rightarrow$
  $E(G_1) \triangle E(G_2)$ is a balanced red-blue graph
- Fix an alternating closed trail on $E(G_1) \triangle E(G_2)$
- Crucial decision: decomposing the alternating trail into "simpler" alternating circuits

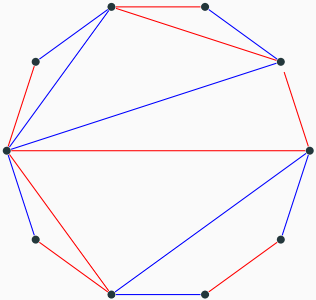
G_2



Designing the multicommodity-flow

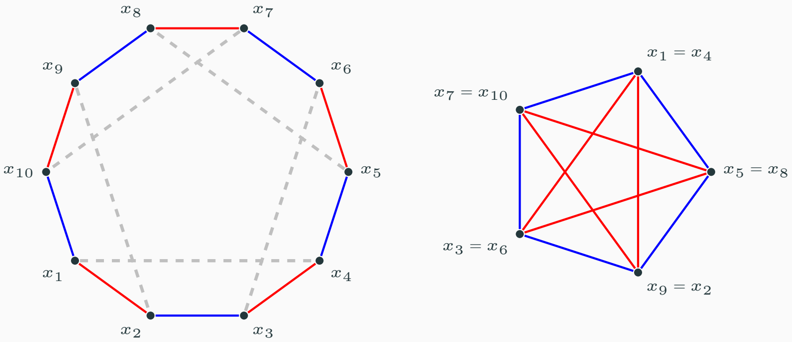
- Let $G_1, G_2 \in \mathcal{G}(d) \Rightarrow$
  $E(G_1) \triangle E(G_2)$ is a balanced red-blue graph
- Fix an alternating closed trail on $E(G_1) \triangle E(G_2)$
- Crucial decision: decomposing the alternating trail into "simpler" alternating circuits

$E(G_1) \triangle E(G_2)$



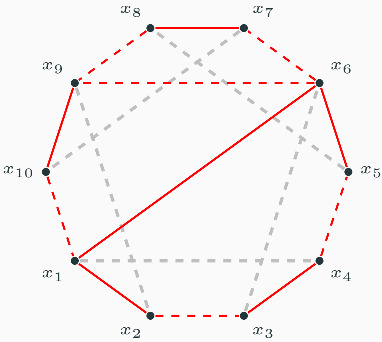
Elementary alternating circuits

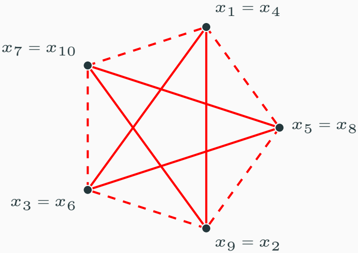
Elementary alternating circuit: each vertex on the trail is either visited once, or twice but with different parity



Elementary alternating circuits

Elementary alternating circuit: each vertex on the trail is either visited once, or twice but with different parity

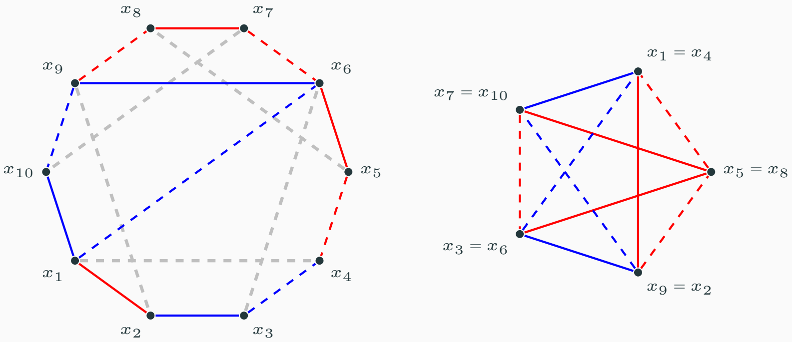




Sweep: the edges are exchanged with non-edges and vica versa from x_1 in counter-clockwise order

Elementary alternating circuits

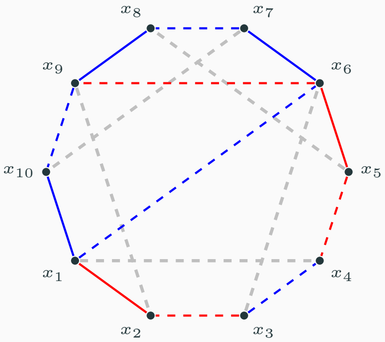
Elementary alternating circuit: each vertex on the trail is either visited once, or twice but with different parity

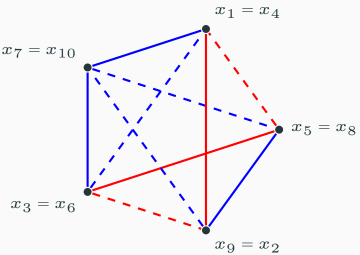


Sweep: the edges are exchanged with non-edges and vica versa from x_1 in counter-clockwise order

Elementary alternating circuits

Elementary alternating circuit: each vertex on the trail is either visited once, or twice but with different parity

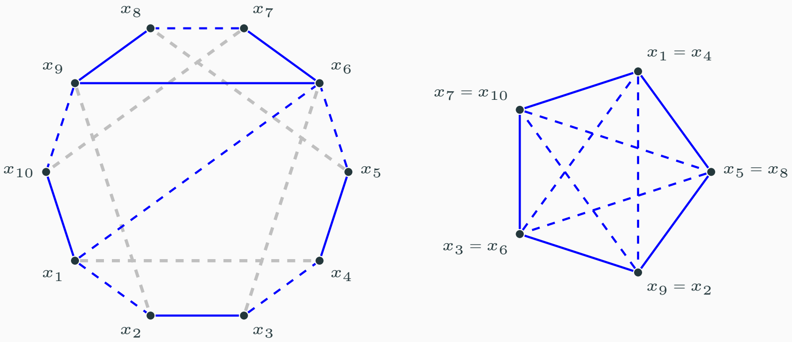




Sweep: the edges are exchanged with non-edges and vica versa from x_1 in counter-clockwise order

Elementary alternating circuits

Elementary alternating circuit: each vertex on the trail is either visited once, or twice but with different parity



Sweep: the edges are exchanged with non-edges and vica versa from x_1 in counter-clockwise order