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
INTRODUCTION

• Problem: given a non-negative integer sequence $d$ of even sum, generate a graph $G \in G(d)$ 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)

- *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*
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 Blitzstein 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 Blitzstein and Diaconis: the distribution is known but not uniform, unknown variance (quality of the sample is unknown)

- Monte Carlo Markov Chain (MCMC) methods ⇒
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.

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

$t = 1$
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.

\[ \forall i \sum_j p_{i,j} = 1 \]

$t = 3$
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.

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

$t = 3$
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.

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

$t = 4$
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.

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

$t = 4$
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 $\varepsilon$ close to the uniform distribution in variation ($\ell_1$-)distance in $\text{poly}(n) \cdot \log \varepsilon^{-1}$ steps (rapidly mixing)
State space: $\mathcal{G}(d) \cup \bigcup_{i, j \in V} \mathcal{G}(d - 1_i - 1_j)$.

Transitions: u.a.r. choose $a, b \in V(G)$, then
State space: $\mathcal{G}(d) \cup \bigcup_{i, j \in V} \mathcal{G}(d - 1_i - 1_j)$.

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.
Jerrum-Sinclair chain

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

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

To get a sample from $G(d)$ in reasonable time, we must have

(where $n = \dim(d)$)

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

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.
Theorem (Jerrum and Sinclair 1992)

The class of degree sequences satisfying

\[(\Delta - \delta + 1)^2 \leq 4\delta(n - \Delta + 1)\]

is $P$-stable.
\( h_0(n) = (1, 2, \ldots, n - 1, n, n, n + 1, \ldots, 2n - 1) \)

has a unique realization

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

Can be blown up to a non-pathological non-\( P \)-stable class.
A (SEEMINGLY PATHOLOGICAL) OBSTACLE

Breaker of (Markov) Chains
Half-Graph of Erdős and Hajnal
Queen of Split Graphs
Protector of Irregularity

\[ h_0(n) = (1, 2, \ldots, n-1, n, n, n+1, \ldots, 2n-1) \]

has a unique realization

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

Can be blown up to a non-pathological non-\( P \)-stable class.
The cardinality of the state space of the JS chain can easily be a factor of $n^8$ larger than $\mathcal{G}(d)$. 

$\mathcal{G}(7, 4, 1, \ldots)$

$\mathcal{G}(6, 4, 1, \ldots)$

$\mathcal{G}(7, 3, 1, \ldots)$

$\mathcal{G}(7, 3, 2, \ldots)$

$\mathcal{G}(6, 3, 2, \ldots)$
**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)

\[
\text{swap:}
\]

\[
\begin{array}{c}
\text{swap:} \\
\hline
\end{array}
\]
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)
## Previous results on the Swap chain

Rapid mixing of the Swap Markov chain shown by

<table>
<thead>
<tr>
<th></th>
<th>simple</th>
<th>bipartite</th>
<th>directed</th>
</tr>
</thead>
<tbody>
<tr>
<td>regular</td>
<td>Cooper, Dyer, Greenhill 2007</td>
<td>Erdős et al. 2013</td>
<td>Greenhill 2011</td>
</tr>
<tr>
<td>( \Delta \leq c\sqrt{m} )</td>
<td>Greenhill and Sfragara 2018</td>
<td>Erdős, Miklós, M, Soltész 2018</td>
<td></td>
</tr>
<tr>
<td>Interval</td>
<td>( (\Delta - \delta)^2 \leq \delta(n - \Delta) )</td>
<td>Erdős, Miklós, M, Soltész 2018</td>
<td>similar</td>
</tr>
<tr>
<td>strongly stable</td>
<td>Amanatidis and Kleer 2019</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A UNIFYING RESULT

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...?
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

\[
\mathcal{H}_k := \left\{ h_k(n) : n \geq k \right\},
\]

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 $H_0(n) \bigtriangleup G$ look like?
**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?

\( H_0(n) \triangle G \)

\[
\begin{array}{ccccccccc}
& & & & & & & & \\
& 1 & & 2 & & i & & n-1 & n-1 \\
& & \bullet & & \bullet & & \bullet & & \bullet & \bullet \\
n-1 & n-1 & & & & n-i+1 & 2 & 1 & \\
& \bullet & & \bullet & & \bullet & & \bullet & \\
& & & & & & & & \\
\end{array}
\]
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?

$H_0(n) \triangle G$ 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

\[ h_0(n) := \begin{pmatrix} 1 & 2 & 3 & \ldots & n-2 & n-1 & n \\ n & n-1 & n-2 & \ldots & 3 & 2 & 1 \end{pmatrix} \]

\[ h_1(n) := \begin{pmatrix} 1 & 2 & 3 & \ldots & n-2 & n-1 & n-1 \\ n-1 & n-1 & n-2 & \ldots & 3 & 2 & 1 \end{pmatrix} \]

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 $\Gamma$ 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 $X, Y \in \mathcal{G}(d)$.

$\mathcal{G}(d)$

\[ \rho(\Gamma): \text{max number of sequences through a } Z \]

\[ \ell(\Gamma): \text{max length of a sequence in } \Gamma \]

**Theorem (follows from Jerrum and Sinclair 1990)**

\[ \tau_{\text{swap}}(\varepsilon) \leq \text{poly}(n) \cdot \frac{\rho(\Gamma)}{|\mathcal{G}(d)|} \cdot \ell(\Gamma) \cdot (\log(|\mathcal{G}(d)|) + \log(\varepsilon^{-1})) \]
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 $Y$!

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

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

bipartite

add a clique

orient upwards

simple

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

\[ \tau_{\text{swap}}(\varepsilon) \leq n^4 \cdot \max_{G \in \mathcal{G}(d)} \sum_{G \in \gamma} \frac{f(\gamma)|\gamma|}{|\mathcal{G}(d)|} \cdot (\log(|\mathcal{G}(d)|) + \log(\varepsilon^{-1})) \]
• 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
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
• 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
Elementary alternating circuit: each vertex on the trail is either visited once, or twice but with different parity.
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 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 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 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