A non-P-stable class of degree sequences for which the swap Markov chain is rapidly mixing

Péter L. Erdős[*,†], Ervin Győri[*], Tamás Róbert Mezei[*,†], István Miklós[†,‡], Dániel Soltész[†,§]

<erdos.peter,gyori.ervin,mezei.tamas.robert,miklos.istvan,soltesz.daniel>@renyi.hu

Alfréd Rényi Institute of Mathematics, Reáltanoda u. 13–15, 1053 Budapest, Hungary

Abstract. One of the simplest methods of generating a random graph with a given degree sequence is provided by the Monte Carlo Markov Chain method using **swaps**. The swap Markov chain converges to the uniform distribution, but generally it is not known whether this convergence is rapid or not. After a number of results concerning various degree sequences, rapid mixing was established for so-called P-stable degree sequences (including that of directed graphs), which covers every previously known rapidly mixing region of degree sequences.

In this paper we give a non-trivial family of degree sequences that are not P-stable and the swap Markov chain is still rapidly mixing on them. The proof uses the 'canonical paths' method of Jerrum and Sinclair. The limitations of our proof are not fully explored.

1. Introduction

An important problem in network science is to sample simple graphs with a given degree sequence (almost) uniformly. In this paper we study a **Markov Chain Monte Carlo** (MCMC) approach to this problem. The MCMC method can be successfully applied in many special cases. A vague description of this approach is that we start from an arbitrary graph with a given degree sequence and sequentially apply small random modifications that preserve the degree sequence of the graph. This can be viewed as a random walk on the space of **realizations** (graphs) of the given degree sequence. It is well-known that after sufficiently many steps the distribution over the state space is close to the uniform distribution. The goal is to prove that the necessary number of steps to take (formally, the mixing time of the Markov chain) is at most a polynomial of the length of the degree sequence.

In this paper we study the so-called **swap Markov chain** (also known as the switch Markov chain). For clarity, we refer to the degree sequence of a simple graph as an **unconstrained** degree sequence. From now on, a degree sequence may refer to unconstrained, bipartite, and directed degree sequences.

Definition 1.1 (Swap). *For a bipartite or an unconstrained degree sequence* $\boldsymbol{d}$*, we say that two realizations* $G_1, G_2 \in \mathcal{G}(\boldsymbol{d})$ *are connected by a swap, if*

$$|E(G_1) \triangle E(G_2)| = 4.$$

A swap can be seen in Figure 1; for the precise definition of the swap Markov chain, see Definition 2.1. Clearly, if G_1 and G_2 are two simple graphs joined by a swap, then $F = E(G_1) \triangle E(G_2)$ is a C_4, and $E(G_2) = E(G_1) \Delta F$. Hence, the term swap is also used

[*]The authors were supported in part by the National Research, Development and Innovation Office – NKFIH grant K-116769

[†]The authors were supported in part by the National Research, Development and Innovation Office – NKFIH grant KH-126853

[‡]IM was supported in part by the National Research, Development and Innovation Office – NKFIH grant SNN-116095

[§]DS was supported in part by the National Research, Development and Innovation Office – NKFIH grans K-120706 and KH-130371.

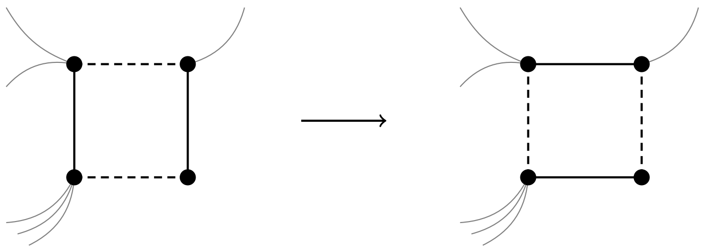

FIGURE 1. A swap (dashed lines emphasize missing edges)

to refer to the operation of taking the symmetric difference with a given C_4. It should be noted, though, that only a minority of C_4's define a (valid) swap, either because they do not preserve the degree sequence (if the C_4 does not alternate between edges of G_1 and G_2), or because they introduce an edge which violates the constraints of the model (say, an edge inside one of the color classes in the bipartite case).

The question whether the mixing time of the swap Markov chain is short enough is interesting from both a practical and a theoretical point of view (although short enough depends greatly on the context). The swap Markov chain is already used in applications, hence rigorous upper bounds on its mixing time are much needed, even for special cases.

The swap Markov chain uses transitions which correspond to minimal perturbations: let G_1, G_2 be two distinct graphs with the same degree sequence; the smallest possible size of the symmetric difference of the edges of G_1 and G_2 is at least 4, and equality holds if and only if G_1 can be transformed into G_2 by a single swap. There are many instances where the Markov chain of the smallest perturbations turned out to have polynomial mixing time, see [13]. However, it is unknown whether the mixing time of the swap Markov chain is uniformly bounded by a polynomial for every (unconstrained) degree sequence. Hence from a theoretical point of view, even an upper bound of $O(n^{10})$ on the mixing time of the swap Markov chain would be considered a great success, even though in practice it is only slightly better than no upper bound at all.

The present paper is written from a theoretical point of view and should be considered as a step towards answering the following question.

Question 1.2 (Kannan, Tetali, and Vempala [11])**.** *Is the swap Markov chain rapidly mixing on the realizations of all graphic degree sequences?*

There is a long line of results where the rapid mixing of the swap Markov chain is proven for certain degree sequences, see [2, 12, 8, 6, 7, 9]. Recently some of these results were unified, first by Amanatidis and Kleer [1], who established rapid mixing for so-called strongly stable classes of degree sequences of simple and bipartite graphs. The most general result at the time of writing is provided in [5], where rapid mixing is established for those classes of unconstrained, bipartite, and directed degree sequences that are ***P*-stable** (see Definition 1.3).

We will denote graphs with upper case letters (e.g. G), degree sequences (which are non-negative integer vectors) with bold lower case letters (e.g. $\boldsymbol{d}$). The closed neighborhood of a subset of vertices $U \subseteq V(G)$ in a graph G is denoted by $N_G[U] \supseteq U$. Classes of graphs and classes of degree sequences are both denoted by upper case calligraphic letters (e.g. $\mathcal{G}$ or $\mathcal{H}$). We say that a graph G is a realization of a degree sequence $\boldsymbol{d}$, if the degree sequence of G is $\boldsymbol{d}$. For a degree sequence $\boldsymbol{d}$, we denote the set of all realizations of $\boldsymbol{d}$ by $\mathcal{G}(\boldsymbol{d})$. The L_1-norm of a vector x is denoted by $\|x\|_1$.

Definition 1.3. *Let $\mathcal{D}$ be a set of degree sequences. We say that $\mathcal{D}$ is **P-stable**, if there exists a polynomial $p \in \mathbb{R}[x]$ such that for any $n \in \mathbb{N}$ and any degree sequence $\boldsymbol{d} \in \mathcal{D}$ on n vertices, any degree sequence d' with $\|d' - d\|_1 = 2$ satisfies $|\mathcal{G}(\boldsymbol{d}')| \le p(n) \cdot |\mathcal{G}(\boldsymbol{d})|$.*

The main contribution in the present paper is that we make the first non-trivial steps beyond P-stability.

Definition 1.4. *For all $n, k \in \mathbb{Z}^+$ and*

$$\boldsymbol{h}_k(n) := \begin{pmatrix} n-k & n-1 & n-2 & \cdots & 3 & 2 & 1 \\ n-k & n-1 & n-2 & \cdots & 3 & 2 & 1 \end{pmatrix} \qquad \mathcal{H}_k := \bigcup_{n>k} \boldsymbol{h}_k(n)$$

be a bipartite degree sequence and a class of bipartite degree sequences respectively.

We first show that methods of [5] do not apply to sequences from $\mathcal{H}_k$.

Theorem 1.5. *For any fixed k, the class of degree sequences $\mathcal{H}_k$ is not P-stable.*

Our main result is that we prove the rapid mixing for the degree sequences in these classes.

Theorem 1.6. *For any fixed k, the swap Markov chain is rapidly mixing for the bipartite degree sequences in $\mathcal{H}_k$.*

We extend the region rapid mixing of the swap Markov chain to a family of non-P-stable degree sequence classes in the directed and the unconstrained case by simple reductions to the bipartite case. There are directed and unconstrained degree sequence classes corresponding to $\mathcal{H}_k$, they will be defined in Section 3, and be denoted by $\mathcal{H}_k^{\mathrm{dir}}$ and $\mathcal{H}_k^{\mathrm{UC}}$ respectively.

Corollary 1.7. *Neither $\mathcal{H}_k^{dir}$ nor $\mathcal{H}_k^{UC}$ are P-stable degree sequence classes. For any fixed k, the swap Markov chain is rapidly mixing on the realizations of both $\mathcal{H}_k^{dir}$ and $\mathcal{H}_k^{UC}$.*

The paper is organized as follows. In Section 2 we introduce the swap Markov chains, some related definitions, and Sinclair's result on mixing time. Section 3 shows why it is sufficient to analyze only the bipartite swap Markov chain. Section 4 describes how to decompose realizations of $\boldsymbol{h}_k(n)$, which is then used by Section 5 and Section 6 to prove Theorem 1.5 and Theorem 1.6, respectively. Finally, in Section 7 we a propose a possible extension of $\mathcal{H}_k$ such that the two theorems still hold.

2. Definitions, notation, and simple proofs

2.1. **The swap Markov chain.** For the precise definition of Markov chains and an introduction to their theory, the reader is referred to Durrett [3]. To define the unconstrained, bipartite, and directed swap Markov chains, it is sufficient to define their transition matrices.

Definition 2.1 (unconstrained/bipartite swap Markov chain)**.** *Let $\boldsymbol{d}$ be an unconstrained or bipartite degree sequence on n vertices. The state space of the swap Markov chain $\mathcal{M}(\boldsymbol{d})$ is $\mathcal{G}(\boldsymbol{d})$. The transition probability between two different states of the chain is nonzero if and only if the corresponding realizations are connected by a swap, and in this case this probability is $\frac{1}{6}\binom{n}{4}^{-1}$. The probability that the chain stays at a given state is one minus the probability of leaving the given state.*

It is well-known that any two realizations of an unconstrained or bipartite degree sequence can be transformed into one-another through a series of swaps, for example through Havel and Hakimi's canonical realizations: the proof the theorem actually proves the existence of a swap sequence between any realization and the canonical realization. The property of connectedness had been discovered even earlier, probably first by Petersen.

The degree sequence of a directed graph consists of a pair of in-degree and out-degree sequences. A realization of a directed degree sequence is a graph with the required in- and out-degrees, without loops and parallel edges, but cycles of length two are allowed.

A simple example shows that the space of all realizations of a directed degree sequence is not necessarily connected by swaps: $(1, 1, 1; 1, 1, 1)$. The two realizations are oppositely directed triangles, and there is no possible swap between them, since there are only three vertices. There are other, arbitrarily large, directed degree sequences whose realizations are not connected by swaps. For this reason, the directed swap Markov chain utilizes an additional operation besides swaps. We shall use the so-called **triple-swap**.

Definition 2.2 (Triple-swap)**.** *For a directed degree sequence $\boldsymbol{d}$, we say that two realizations $G_1, G_2 \in \mathcal{G}(\boldsymbol{d})$ are connected by a triple-swap if there is a cyclically directed triangle in G_1 with the property that reversing its edges transforms G_1 into G_2.*

Now we are ready to define the swap Markov chain for directed degree sequences.

Definition 2.3 (directed swap Markov chain)**.** *Let $\boldsymbol{d}$ be a directed degree sequence on n vertices. The state space of the swap Markov chain $\mathcal{M}(\boldsymbol{d})$ for $\boldsymbol{d}$ is $\mathcal{G}(\boldsymbol{d})$. The transition probability between two different states of the chain is $\frac{1}{12}\binom{n}{4}^{-1}$ if they are connected by a swap, $\frac{1}{4}\binom{n}{3}^{-1}$ if they are connected by a triple-swap, and zero otherwise. The probability that the chain stays at a given state is one minus the probability of leaving the given state.*

The swap Markov chains defined are irreducible (connected), symmetric, reversible, and lazy. Their unique stationary distribution is the uniform distribution $\pi \equiv |\mathcal{G}(\boldsymbol{d})|^{-1}$.

Definition 2.4. *The mixing time of a Markov chain is*

$$\tau(\varepsilon) = \min\left\{t_0 \ : \ \forall x \ \forall t \geq t_0 \ \|P^t(x, \cdot) - \pi\|_1 \leq 2\varepsilon\right\},$$

where $P^t(x, y)$ is the probability that the Markov chain started from x is at state y after t steps.

Definition 2.5. *The swap Markov chain is said to be rapidly mixing on an infinite set of degree sequences $\mathcal{D}$ if there exists a fixed polynomial $\mathrm{poly}(n, \log \varepsilon^{-1})$ which bounds the mixing time of the swap Markov chain on $\mathcal{G}(\boldsymbol{d})$ for any $\boldsymbol{d} \in \mathcal{D}$ (where n is the length of $\boldsymbol{d}$).*

Sinclair's seminal paper describes a combinatorial method to bound the mixing time.

Definition 2.6 (Markov graph)**.** *Let $G(\mathcal{M}(\boldsymbol{d}))$ be the graph whose vertices are realizations of $\boldsymbol{d}$ and two vertices are connected by an edge if the swap Markov chain on $\mathcal{G}(\boldsymbol{d})$ has a positive transition probability between the two realizations.*

We say that $\gamma_{G,H}$ is a **canonical path** if it is a path connecting two distinct realizations $G, H \in \mathcal{G}(\boldsymbol{d})$ in the Markov graph. The load of Γ of a set of canonical paths is defined as

$$\rho(\Gamma) = \max_{P(e) \neq 0} \frac{|\{\gamma \in \Gamma \ : \ e \in E(\gamma)\}|}{|\mathcal{G}(\boldsymbol{d})| \cdot P(e)},$$

where $P(e)$ is the transition probability assigned to the edge e of the Markov graph (this is well-defined because the studied Markov chains are symmetric). The next lemma follows from Proposition 1 and Corollary 4 of Sinclair [14].

Lemma 2.7. *If Γ consists of canonical paths $\gamma_{G,H}$ for each pair of distinct realizations, then*

$$\tau(\varepsilon) \leq \rho(\Gamma) \cdot \ell(\Gamma) \cdot \left(\log(|\mathcal{G}(\boldsymbol{d})|) + \log(\varepsilon^{-1})\right),$$

where $\ell(\Gamma)$ is the length of the longest path in Γ.

Obviously, $\log(|\mathcal{G}(\boldsymbol{d})| \leq n^2$, so from now on our focus is on bounding ρ by a polynomial of n. Before turning to the proofs of Theorem 1.5 and Theorem 1.6, we show how they imply Corollary 1.7.

3. Reduction to the bipartite case

A large part of the present paper deals with the realizations of $\boldsymbol{h}_k(n)$, hence it will be convenient to introduce the following notation.

Definition 3.1. *For all k, let the two color classes of the realizations of $\boldsymbol{h}_k(n)$ be denoted by A and B, where $A := \{a_1, \dots, a_n\}$ and $B := \{b_1, \dots, b_n\}$. For $1 \leq i \leq n-1$ the degree of a_i and b_i is i, and the degree of a_n and b_n is $n-k$.*

$$\begin{pmatrix} a_n & a_{n-1} & a_{n-2} & \cdots & a_3 & a_2 & a_1 \\ b_n & b_{n-1} & b_{n-2} & \cdots & b_3 & b_2 & b_1 \end{pmatrix}$$

The reason why an analogue of Theorem 1.6 holds for some unconstrained and directed degree sequences is that for each bipartite degree sequence (class) we can associate a corresponding unconstrained and directed degree sequence (class) such that this map induces a graph homomorphism from the Markov graph of the bipartite swap Markov chain to the Markov graph of the mapped degree sequence.

3.1. Reducing the unconstrained case to the bipartite case.

Definition 3.2. *Let $\boldsymbol{d} := (d_1^A, \dots, d_n^A; d_1^B, \dots, d_n^B)$ be a bipartite degree sequence. We define a corresponding unconstrained degree sequence as $\mathrm{UC}(\boldsymbol{d}) := (d_1^A, \dots, d_n^A, d_1^B + (n-1), d_2^B + (n-1), \dots, d_n^B + (n-1))$ on $(a_1, \dots, a_n, b_1, \dots, b_n)$ as vertices.*

Lemma 3.3. *For any realization G of $\mathrm{UC}(\boldsymbol{d})$, the induced subgraph $G[B]$ is a clique.*

Proof. Trivial, also follows from the Erdős-Gallai theorem:

$$e(G[B]) \geq \frac{1}{2}\left(\sum_{i=1}^{n}(d_i^B + n - 1) - \sum_{i=1}^{n} d_i^A\right) = \frac{1}{2}n(n-1).$$

□

Clearly, inserting/removing the edges of the clique on B produces a bijection between the realizations of $\boldsymbol{d}$ and $\mathrm{UC}(\boldsymbol{d})$. In particular, if one is P-stable, the other is as well. The bijection preserves swaps, since the symmetric difference of the edge sets of two realizations does not change as a result of the described operation.

By using the described mapping, a set of canonical paths Γ on $G(\mathcal{M}(\boldsymbol{d}))$ can be mapped to a set of canonical paths Γ_{UC} on $G(\mathcal{M}(\mathrm{UC}(\boldsymbol{d})))$ such that $\rho(\Gamma_{\mathrm{UC}}) \leq \rho(\Gamma)$.

Definition 3.4. *Let $\mathcal{H}_k^{\mathrm{UC}} := \{\mathrm{UC}(\boldsymbol{h}_k(n)) \ : \ n > k\}$.*

The statements of Corollary 1.7 on $\mathcal{H}_k^{\mathrm{UC}}$ will follow once Theorem 1.5 and Theorem 1.6 is proved (using Lemma 2.7).

3.2. **Reducing the directed case to the bipartite case.**

Definition 3.5. *Let* $\boldsymbol{d} := (d_1^A, \dots, d_n^A; d_1^B, \dots, d_n^B)$ *be a bipartite degree sequence. We define a corresponding directed degree sequence on* $A \cup B$ *as vertices as the* $4n$*-dimensional vector*

$$\begin{aligned}\vec{\boldsymbol{d}} &= (\,\vec{\boldsymbol{d}}_{in}(a_1), \dots, \vec{\boldsymbol{d}}_{in}(a_n), \vec{\boldsymbol{d}}_{in}(b_1), \dots, \vec{\boldsymbol{d}}_{in}(b_n); \vec{\boldsymbol{d}}_{out}(a_1), \dots, \vec{\boldsymbol{d}}_{out}(a_n), \vec{\boldsymbol{d}}_{out}(b_1), \dots, \vec{\boldsymbol{d}}_{out}(b_n)\,) \\ &= (\quad d_1^A \quad, \dots, \quad d_n^A \quad, \quad 0 \quad, \dots, \quad 0 \quad; \quad 0 \quad, \dots, \quad 0 \quad, \quad d_1^B \quad, \dots, \quad d_n^B \quad).\end{aligned}$$

It is obvious that realizations of $\boldsymbol{d}$ and $\vec{\boldsymbol{d}}$ are in a 1-to-1 correspondence with each other through the operations of undirecting the edges and directing them from B to A, respectively. As in the unconstrained case, the statements of Corollary 1.7 on $\mathcal{H}_k^{\mathrm{dir}}$ will follow from the bipartite case.

4. A geometric representation of the realizations of a degree sequence

In this section we introduce a representation of the realizations of a degree sequence. This representation will turn out to be quite simple for the degree sequences $\boldsymbol{h}_i(n)$.

Definition 4.1 (standard drawing)**.** *For all graphs* G *on vertices* $\{a_1, \dots, a_n, b_1, \dots, b_n\}$, *we say that a drawing of* G *in the plane is a standard drawing if all the edges are straight lines and for all* $1 \le j \le n$, *the coordinates of* a_j, b_j *are* $(j - 0.5, 1), (n - j, 0)$ *respectively, see Figure 2.*

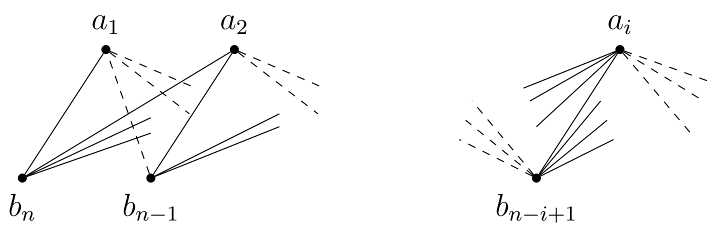

FIGURE 2. The standard drawing of the single realization of $H_0(n)$.

For two graphs G_1, G_2 on the same vertex set, we define their *symmetric difference* $G_1 \triangle G_2$ with $V(G_1 \triangle G_2) = V(G_1) = V(G_2)$ and $E(G_1 \triangle G_2) = E(G_1) \triangle E(G_2)$.

Definition 4.2 (geometric representation)**.** *The geometric representation of a realization* G *of a degree sequence* $\boldsymbol{h}_k(n)$ *is the standard drawing of the symmetric difference* $G \triangle H_0(n)$.

It is not hard to see that every geometric representation of a realization of $\boldsymbol{h}_1(n)$ is a single path connecting b_n to a_n that is x-monotone (its first coordinate changes monotonically), and y-alternating (its second coordinate alternates between 0 and 1), see for example Figure 3.

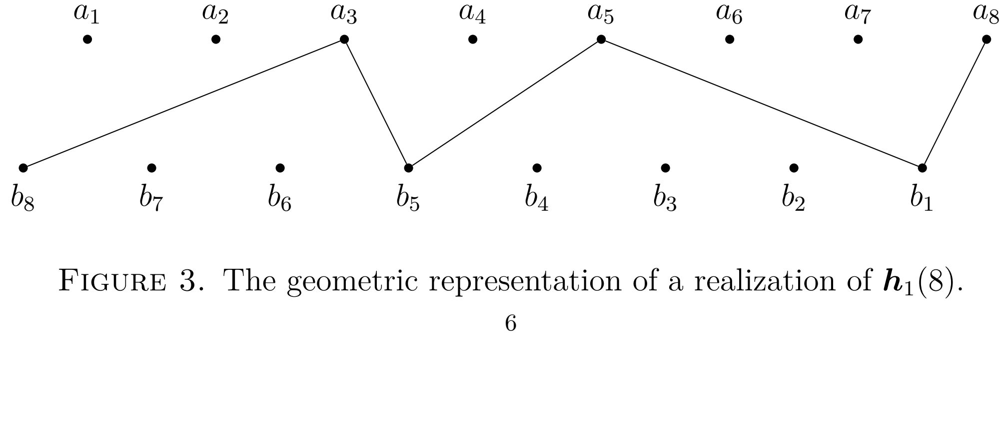

FIGURE 3. The geometric representation of a realization of $\boldsymbol{h}_1(8)$.

Definition 4.3 (k-path system)**.** *Let G be a graph with $V(G) = \{a_1, \ldots, a_n, b_1, \ldots, b_n\}$, we say that the geometric representation of G is a k-path system on $2n$ vertices if there are k edge disjoint graphs $G_1, \ldots, G_k$ on the same vertex set as G, each consisting of a single x-monotone and y-alternating path connecting b_n to a_n, and $E(G) = E(G_1) \uplus \ldots \uplus E(G_k)$.*

Lemma 4.4. *For all $k < n$, the geometric representation of every realization of $\boldsymbol{h}_k(n)$ is a k-path system. And conversely, for every k-path system G, there is a realization of $\boldsymbol{h}_k(n)$ whose geometric representation is G.*

Proof. Let G be a realization of $\boldsymbol{h}_k(n)$ and H_0 be the single realization of $\boldsymbol{h}_0(n)$. For every i, j, the vertices $a_j b_i \in E(H_0)$ if and only if the x-coordinate of b_i is smaller than the x-coordinate of a_j. If $j \neq n$, the degree of a_j is the same in H_0 as in G, hence in the geometric representation of G, the number of neighbors of a_j that have a smaller x coordinate than a_j is the same as the number of neighbors having a larger x-coordinate. In the case of $j = n$, the degree of a_n in the geometric representation of G is clearly k.

A similar argument shows that for all $j \neq n$, the neighbors of b_j whose x-coordinate is smaller than the x-coordinate of b_j is the same as the number of neighbors whose x-coordinate is larger, and the degree of b_n is again k. From these properties the first statement of Lemma 4.4 follows by a simple induction on k.

The second statement follows readily from the observation that taking the symmetric difference of a k-path system with the standard drawing of H_0 results in a realization of $\boldsymbol{h}_k(n)$. □

5. Non-P-stability results

The main results of this section make use of the fact that it is relatively straightforward to get the asymptotic growth rate of the number of k-path systems when k is a constant and n tends to infinity. We first illustrate this in case of $k = 1$. For any $1 \leq \ell \leq n$ let $U_\ell := \{a_1, \ldots, a_\ell, b_n, \ldots, b_{n-\ell+1}\}$. We always assume that the vertices in U_ℓ are drawn in the plane as in Figure 2.

Lemma 5.1. *The number of 1-path systems on $2n$ vertices is*

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{n-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Proof. Let $S_a(\ell)$ be the number of x-monotone, y-alternating paths in U_ℓ that start at b_n and end in one of the vertices in $\{a_1, \ldots, a_\ell\}$. Similarly, let $S_b(\ell)$ be the number of x-monotone, y-alternating paths in U_ℓ that start at b_n and end in one of the vertices in $\{b_n, \ldots, b_{n-\ell+1}\}$.

Observe that the number of 1-path systems on $2n$ vertices is exactly $S_b(n)$ (delete the edge incident to a_n of the path). It is easy to see that

$$\begin{aligned} S_a(\ell+1) &= 2S_a(\ell) + S_b(\ell) \\ S_b(\ell+1) &= S_a(\ell) + S_b(\ell). \end{aligned} \tag{1}$$

Since $S_a(1) = S_b(1) = 1$, from (1) we get that $S_b(n)$ is the quantity in the statement of the Lemma and the proof is complete. □

Corollary 5.2. *The number of realizations of $\boldsymbol{h}_1(n)$ is $\Theta\left(\left(\frac{3+\sqrt{5}}{2}\right)^n\right)$*

Proof. It is easy to show that the quantity in Lemma 5.1 is $\Theta(c^n)$ where c is the largest eigenvalue of the matrix

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

which is $\frac{3+\sqrt{5}}{2}$. □

The proof of Lemma 5.1 can be interpreted as follows. We count 1-path systems by looking at their induced subgraphs on the vertices in U_ℓ (the number of these is precisely $S_a(\ell) + S_b(\ell)$). The main observation is that the number of ways an induced subgraph of a 1-path system on the vertices of U_ℓ can be extended to an induced subgraph on the vertices $U_{\ell+1}$ only depends on whether the path's endpoint lies in $\{a_1, \dots, a_\ell\}$ or in $\{b_n \dots, b_{n-\ell+1}\}$.

We shall mimic the argument of Lemma 5.1 with k-path systems where $k > 1$, in the sense that we shall consider induced subgraphs on he vertices U_ℓ and some kind of "endpoint structure" on it.

Definition 5.3 (type)**.** *Let $\mathcal{P}_k$ be the set of partitions of k (the set of multisets of positive integers whose sum of elements is exactly k) and $\mathcal{P}_0 := \{\emptyset\}$. For all positive integers k, we define the set of types as*

$$\mathcal{T}_k := \{(R, Q) \mid \exists\, 0 \le m \le k : R \in \mathcal{P}_m, Q \in \mathcal{P}_{k-m}\}.$$

Definition 5.4 (type of a subgraph of a k-path system)**.** *Let X be a k-path system on $2n$ vertices and let X_ℓ be the induced subgraph of X on the vertices U_ℓ. We say that the type of X_ℓ is $T = (R, Q) \in \mathcal{T}_k$ if there is an injective function $f : R \to \{a_1, \dots, a_\ell\}$ such that for every $a_i \in f(R)$ we have*

$$\deg_X(a_i) - \deg_{X_\ell}(a_i) = f^{-1}(a_i),$$

and for all $a_i \in \{a_1, \dots, a_\ell\} \setminus f(R)$ we have $\deg_X(a_i) = \deg_{X_\ell}(a_i)$. *Similarly, there is an injective function $g : Q \to \{b_n, \dots, b_{n-\ell+1}\}$ such that for every $b_i \in g(Q)$*

$$\deg_X(b_i) - \deg_{X_\ell}(b_i) = g^{-1}(b_i),$$

and for all $b_i \in \{b_n, \dots, b_{n-\ell+1}\} \setminus g(Q)$ we have $\deg_X(b_i) = \deg_{X_\ell}(b_i)$.

Informally, the type of the subgraph X_ℓ describes the multiplicities of the incidences of the endpoints of the k paths on $\{a_1, \dots, a_\ell\}$ and $\{b_n, \dots, b_{n-\ell}\}$.

In the proof of Lemma 5.1, the functions $S_a(\ell), S_b(\ell)$ were actually the number of $(\{1\}, \emptyset)$ and $(\emptyset, \{1\})$ type subgraphs of 1-path systems on U_ℓ, respectively. The next definition is the analogue of the matrix in the proof of Corollary 5.2 for large k.

Definition 5.5 (type matrix)**.** *For all k, let us fix an ordering of the types: $\mathcal{T}_k = (T_1, \dots, T_{|\mathcal{T}_k|})$. Let ℓ and n be so large, that there exists k-path system which has type T_i on U_ℓ. We define $p_{i,j}$ to be the number possible ways a k-path system on U_ℓ can be extended to a k-path system of type T_j on $U_{\ell+1}$. We define the type-matrix $\mathcal{P}_k$ to be the $|\mathcal{T}_k| \times |\mathcal{T}_k|$ matrix whose element in the i-th row and j-th column is $p_{i,j}$.*

It is not hard to see that $p_{i,j}$ is well-defined, in other words, $p_{i,j}$ does not depend on either ℓ, n, or the k-path system.

In the proof of Corollary 5.2, the type matrix

$$\mathcal{P}_1 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

corresponds to the ordering $\mathcal{T}_1 = \big((\{1\}, \emptyset), (\emptyset, \{1\})\big)$. Now we are ready to prove the analogue of Lemma 5.1 for k-path systems where $k > 1$.

Lemma 5.6. *For every $k > 1$, the number of k-path systems on $2n$ vertices is*

$$v^T \mathcal{P}_k^{n-1} w$$

where:

- *v is the vector of length $|\mathcal{T}_k|$ which contains 1 at the coordinates which correspond to the types $(\{1\},\{k-1\})$, $(\emptyset,\{k\}) \in \mathcal{T}_k$, and zero everywhere else,*
- *$\mathcal{P}_k$ is the type-matrix,*
- *w is the vector of length $|\mathcal{T}_k|$ that contains 1 at the coordinate that corresponds to the type $(\emptyset,\{1,1,\dots,1\})$ and zero everywhere else.*

Proof. With the appropriate substitutions, the proof is identical to the proof of Lemma 5.1. The type of k-path system induced on U_1 is either $(\{1\},\{k-1\})$ or $(\emptyset,\{k\})$. By the definition of $\mathcal{P}_k$, the vector $v^T\mathcal{P}^{n-1}$ contains the number of graphs on the vertices U_n with a given type. Of these, the k-path systems correspond to graphs with type $(\emptyset,\{1,1,\dots,1\})$ (deleting a_n from a k-path system transforms it to a graph of this type). Hence the statement of the lemma follows. □

The following simple property of the type matrix will be used.

Definition 5.7. *A matrix $\mathcal{P}$ is primitive, if $\exists m$ for which every entry of $\mathcal{P}^m$ is positive.*

Lemma 5.8. *For all k, the type matrix $\mathcal{P}_k$ is primitive.*

Proof. For every type $t \in \mathcal{T}_k$ it is easy to design a k-path system X such that the induced subgraph of X on U_ℓ (for some ℓ) is t and the induced subgraph of X on $U_{\ell+k}$ is of type $(\emptyset,\{1,1,\dots,1\})$. (A possible solution is what we call later a half-buffer, see Definition 6.1.) Hence in $\mathcal{P}_k^k$ the row and column that correspond to the type $(\emptyset,\{1,1,\dots,1\})$ are strictly positive. Since $\mathcal{P}_k$ is non-negative, it also follows that $\mathcal{P}_k^{2k}$ is positive. □

Now we are ready to prove the key lemma to refute the P-stability of the class of degree sequences $\mathcal{H}_k$.

Lemma 5.9. *For every k, the largest eigenvalue of the type-matrix $\mathcal{P}_k$ is smaller than the largest eigenvalue of the type matrix $\mathcal{P}_{k+1}$.*

Proof. By Lemma 5.8, both $\mathcal{P}_k$ and $\mathcal{P}_{k+1}$ are primitive. By the Perron-Frobenius theory, they both have a real positive eigenvalue r_k and r_{k+1}, respectively, that is larger in absolute value than all of their other eigenvalues. Moreover, both limits

$$\lim_{n\to\infty}\frac{\mathcal{P}_k^n}{r_k^n} \quad \text{and} \quad \lim_{n\to\infty}\frac{\mathcal{P}_{k+1}^n}{r_{k+1}^n}$$

exist and are one dimensional projections. Let the set of types $S \subset \mathcal{T}_{k+1}$ be defined as follows:

$$S := \{(R,Q) \in \mathcal{T}_{k+1} \,:\, 1 \in Q\}.$$

Let $M^{(n)}$ be the principal minor of $\mathcal{P}_{k+1}^n$ that is obtained by taking those rows and columns which correspond to types in S. Without loss of generality, we may assume that if the i-th row of $M^{(1)}$ corresponds to a type (R,Q), then the i-th row of $\mathcal{P}_k$ corresponds to the type $(R, Q\setminus\{1\})$. Moreover, we may assume that the ordering of $\mathcal{T}_k$ and $\mathcal{T}_{k+1}$ is compatible in the following sense: if $T = \{R,Q\}$ and $T' = (R',Q')$ are types in S and $T < T'$ according to the ordering on $\mathcal{T}_{k+1}$, then $(R, Q\setminus\{1\}) < (R', Q'\setminus\{1\})$ according to the ordering on $\mathcal{T}_k$.

First, we prove the following two properties of $M^{(1)}$.

(1) The matrix $M^{(1)}$ is elementwise larger than or equal to $\mathcal{P}_k$.
(2) The matrix $M^{(1)}$ is not equal to $\mathcal{P}_k$.

Since $|S| = |\mathcal{T}_k|$, the matrix $M^{(1)}$ is a $|\mathcal{T}_k|\times|\mathcal{T}_k|$ matrix. We start with proving the second statement. The entry of $\mathcal{P}_k$ in the intersection of the row and column that correspond to the type $(\emptyset,\{1,\dots,1\}) \in \mathcal{T}_k$ and $(\{k\},\emptyset) \in \mathcal{T}_k$, respectively, is clearly 1. On the other hand, the value of $M^{(1)}$ in this row and column corresponds to the number of transitions from $(\emptyset, \{1, \dots, 1\}) \in \mathcal{T}_{k+1}$ to $(\{k\}, \{1\})$ which is $k+1$ (the number of ways one can choose one of the $k+1$ paths which will not be extended). Therefore $M^{(1)} \neq \mathcal{P}_k$.

For the first statement, for any two types $(R, Q), (R', Q') \in \mathcal{T}_k$, if a type (R, Q) subgraph of a k-path system on the vertices U_ℓ can be extended to an another type (R', Q') subgraph on the vertices $U_{\ell+1}$ in p ways, then clearly a type $(R, Q \cup \{1\})$ subgraph of a $k+1$-path system on the vertices U_ℓ can be extended to a type $(R', Q' \cup \{1\})$ subgraph on the vertices $U_{\ell+1}$ in at least p ways. Therefore the first property is also proven.

Suppose to the contrary that $r_{k+1} \leq r_k$. Since the limit

$$\lim_{n\to\infty} \frac{\mathcal{P}_{k+1}^n}{r_{k+1}^n}$$

exists and is finite, both the limits

$$\lim_{n\to\infty} \frac{\mathcal{P}_{k+1}^n}{r_k^n} \quad \text{and} \quad \lim_{n\to\infty} \frac{M^{(n)}}{r_k^n}$$

exist and are finite. Since $M^{(1)}$ is a principal minor of $\mathcal{P}_{k+1}$, and every element of $\mathcal{P}_{k+1}$ is non-negative, for all k the matrix $M^{(k)}$ is elementwise larger than or equal to $(M^{(1)})^k$. Hence the sequence

$$\left\{ \frac{(M^{(1)})^n}{r_k^n} \right\}_{n=1}^{\infty}$$

is bounded. By the two properties of $M^{(1)}$ and the fact that $\mathcal{P}_k$ is primitive, it follows that there is an integer m such that $(M^{(1)})^m$ is elementwise strictly larger than $\mathcal{P}_k^m$. Thus there is a positive ε such that $(M^{(1)})^m$ is elementwise strictly larger than $(1+\varepsilon)\mathcal{P}_k^m$. Therefore the sequence

$$\left\{ \frac{((1+\varepsilon)\mathcal{P}_k^m)^n}{r_k^{mn}} \right\}_{n=1}^{\infty} = \left\{ (1+\varepsilon)^n \frac{\mathcal{P}_k^{mn}}{r_k^{mn}} \right\}_{n=1}^{\infty}$$

is bounded, but this clearly contradicts the fact that the limit

$$\lim_{n\to\infty} \frac{\mathcal{P}_{k+1}^n}{r_{k+1}^n}$$

is a one dimensional projection. □

Proof of Theorem 1.5. Observe, that $\|\boldsymbol{h}_{k+1}(n) - \boldsymbol{h}_k(n)\|_1 = 2$. However, according to Lemma 5.9

$$\frac{|\mathcal{G}(\boldsymbol{h}_{k+1}(n))|}{|\mathcal{G}(\boldsymbol{h}_k(n))|} = \Theta\left(\left(\frac{r_{k+1}}{r_k} \right)^n \right),$$

which grows exponentially as $n \to \infty$, so $\mathcal{H}_k$ is not P-stable. □

6. Rapid mixing of the Swap Markov chain

6.1. **Overview of the proof.** There are multiple similarities between the proof approach of Theorem 1.6 and that of previous results in [8, 6, 7, 9]. The rapid mixing results in all these papers, and the present one, apply the following approach. Let $\boldsymbol{d}$ be a degree sequence. They all construct a set of paths in $G(\mathcal{M}(\boldsymbol{d}))$ and bound the number of paths passing through each realization.

Let $X, Y \in \mathcal{G}(\boldsymbol{d})$ be any two distinct realizations. We will define a swap sequence $\gamma_{X,Y} : X = Z_0, Z_1, \dots, Z_t = Y$, which is also referred to as a canonical path in Lemma 2.7. We will also define a set of corresponding encodings $L_0(X,Y), L_1(X,Y), \dots, L_t(X,Y)$ with the following properties.

- **Reconstructable:** there is an algorithm that for each i, takes Z_i and $L_i(X, Y)$ as an input and outputs the realizations X and Y.
- **Encodable in $\mathcal{G}(\boldsymbol{d})$:** the total number of encodings on each vertex of $G(\mathcal{M}(\boldsymbol{d}))$ is not much larger than $|\mathcal{G}(\boldsymbol{d})|$.

In addition, the canonical path system $\Gamma := \{\gamma_{X,Y} \ : \ X, Y \in \mathcal{G}(\boldsymbol{d})\}$ on $G(\mathcal{M}(\boldsymbol{h}_k(n)))$ will satisfy the following two properties:

- No two canonical paths assign the same encoding to any given realization.
- The total number of encodings used at any vertex of $G(\mathcal{M}(\boldsymbol{h}_k(n)))$ is $O(|\mathcal{G}(\boldsymbol{h}_k(n))| \cdot \mathrm{poly}_k(n))$.

These properties ensure that $\rho(\Gamma) = O(\mathrm{poly}_k(n))$, thus proving them will immediately imply that the bipartite swap Markov chain is rapidly mixing because of Lemma 2.7.

Now we give a brief description of how the $X = Z_0, Z_1, \ldots, Z_{t+1} = Y$ canonical path is constructed. To avoid confusion with paths of the k-path system, we shall refer to $X = Z_0, Z_1, \ldots, Z_{t+1} = Y$ as the swap sequence that connects X to Y.

The main idea is to morph the k-path system of X into the k-path system of Y "from left to right": a **buffer**, which is just a region of constant width with a very specific structure, will be inserted on the left side of X, which is then moved from left-to-right peristaltically. In a typical intermediate realization Z_i along the canonical path, the vertices left of the buffer will induce the same graph on Z_i and Y, and the vertices right of the buffer will induce the same graph on Z_i and X, see Figure 4.

The encoding L_i will contain a k-path system whose structure is similar to Z_i, but the role of X and Y is interchanged. Moreover L_i will contain the position of the buffer and some additional information about the vertices in the buffer.

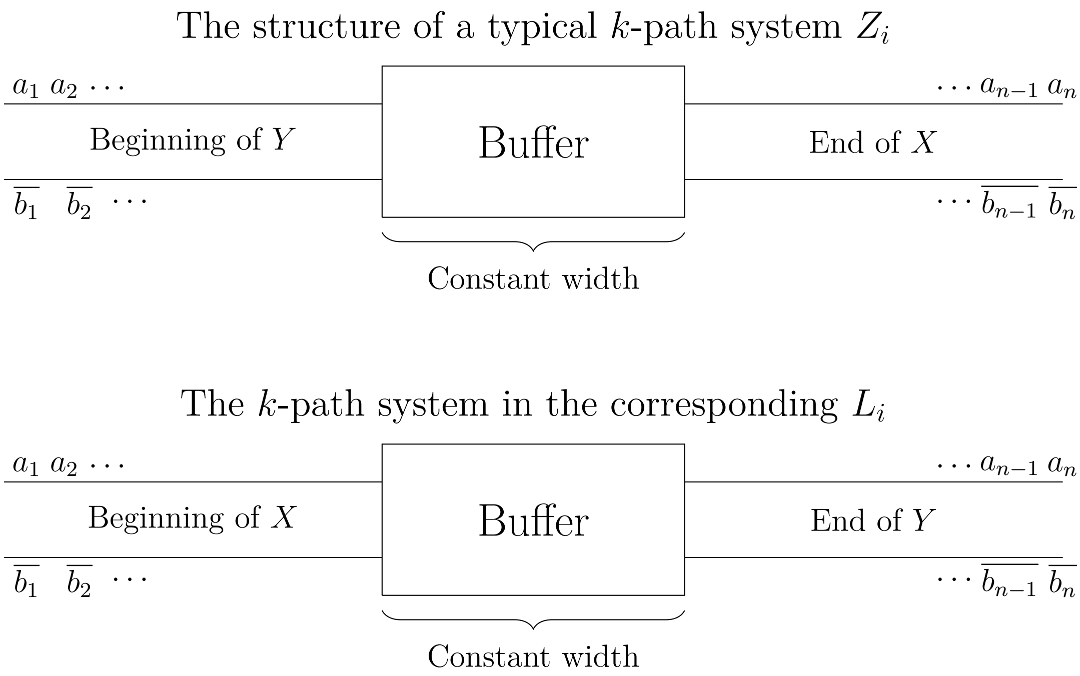

FIGURE 4. A realization along a canonical path and the associated encoding.

Recall that $U_i = \{a_1, \ldots, a_i\} \cup \{b_n, \ldots, b_{n-i+1}\}$ is the union of the first and last i vertices of A and B, respectively.

Definition 6.1 (half-buffer)**.** *Let X be a k-path system on $2n$ vertices and let X_i, X_{i+k} be its induced subgraphs on the vertices U_i, U_{i+k}, respectively. Let A_1, A_2, B_1, B_2 be the sets $\{a_1, \dots, a_i\}, \{a_{i+1} \dots, a_{i+k}\}, \{b_n \dots, b_{n-i+1}\}, \{b_{n-i}, \dots, b_{n-i-k+1}\}$ respectively, see Figure 5. Let ℓ be the number of edges connecting A_1 to B_2. We say that X contains a half-buffer starting at $i+1$ if X satisfies the following properties.*

(1) *There are exactly k edges $E := \{e_1, \dots, e_k\}$ in X that connect $A_1 \cup B_1$ to $A_2 \cup B_2$.*
(2) *There is at most one edge of E incident on any vertex in $A_2 \cup B_2$.*
(3) *The endpoints of the edges in E, in $A_2 \cup B_2$ are "left-compressed": the endpoints of the ℓ edges connecting A_1 to B_2, in B_2 are the vertices $\{a_{i+1}, \dots, a_{i+\ell}\}$. And similarly, the endpoints of the $k - \ell$ edges connecting B_1 to A_2, in A_2 are $\{b_{n-i}, \dots, b_{n-i-k+\ell+1}\}$.*
(4) *The edges of X between A_1 and B_2 are non-crossing. (Recall that X is a k-path system, hence all edges are drawn as straight lines.)*
(5) *The edges of X between B_1 and A_2 are non-crossing.*
(6) *The induced subgraph of X on the vertices $A_2 \cup B_2$ is the matching*

$$\{(a_{i+\ell+1}, b_{n-i}), (a_{i+\ell+2}, b_{n-i+2}) \dots, (a_{i+k}, b_{n-i-k+\ell+1})\}.$$

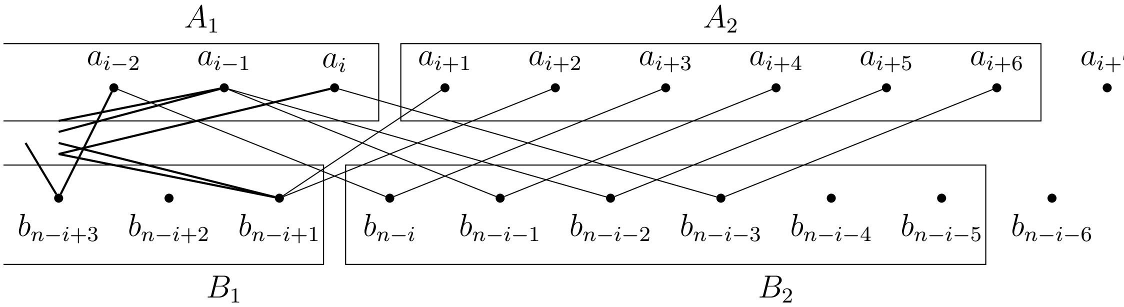

FIGURE 5. The induced subgraph of a 6-path system which contains a half-buffer starting at $i+1$, induced on the vertices $A_1 \cup A_2 \cup B_1 \cup B_2$.

Definition 6.2 (buffer)**.** *Let X be a k-path system on n vertices. Rotate every edge of X around the middle of the segment connecting a_n and b_n by $180°$ to obtain the k-path system X'. For any $0 \le i < i + 2k \le j \le n$, we say that X contains a buffer starting at $i+1$ and ending at j, if*

- *X contains a half-buffer starting at $i+1$,*
- *X' contains a half-buffer starting at j,*
- *X contains the matching*

$$\{(a_{i+1}, b_{n-j+k}), (a_{i+2}, b_{n-j+k-1}), \dots (a_{i+k-1}, b_{n-j+2}), (a_{i+k}, b_{n-j+1})\}.$$

Lemma 6.3. *Let k, z be constants such that $2k \le z$. For every k-path system X on n vertices, and every $0 \le i \le n - z$ there exists a k-path system $\widetilde{X} = \widetilde{X}[i+1, i+z]$ on n vertices and a constant $c = c(k, z)$ satisfying the following properties.*

(1) *The induced subgraphs of X and $\widetilde{X}$ on the vertices U_i are identical.*
(2) *The induced subgraphs of X and $\widetilde{X}$ on the vertices $(A \cup B) \setminus U_{i+z}$ are identical.*
(3) *$\widetilde{X}$ contains a buffer starting at $i+1$ and ending at $i+z$.*
(4) *There is a swap sequence connecting X and $\widetilde{X}$ using at most c swaps, all of which are using four vertices of the closed neighborhood $N_X[U_{i+z} \setminus U_i]$.*

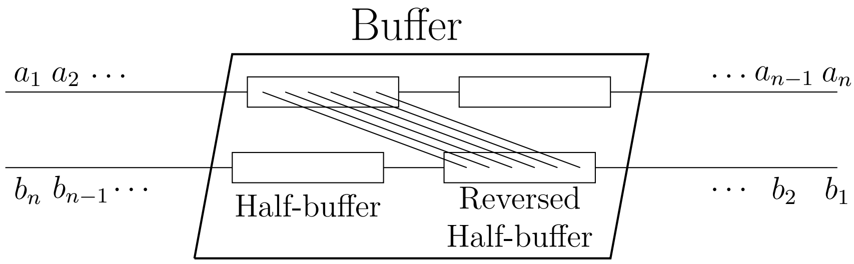

FIGURE 6. A buffer consists of a half-buffer and a "reversed" half-buffer that are connected by a matching of size k, plus a prescribed number of vertices between the ordinary and the reversed half-buffer.

Proof. We prove the existence of $\widetilde{X}$ by constructing it. Take the induced subgraph of X on the vertices of U_i, and extend it to the vertices U_{i+k} by appending a half-buffer starting from $i+1$. Let this graph be X_1. Now take the induced subgraph of X on the vertices $(A \cup B) \setminus U_{i+z}$, and extend it to a graph on the vertices $(A \cup B) \setminus U_{i+z-k}$ by appending a "reversed half-buffer" (see Definition 6.2). Let this graph be X_2. Since $2k \le z$, X_1 and X_2 can be joined by a matching to form a k-path system $\widetilde{X}$ that possesses the first three properties.

The fourth property follows from the observation that every edge in the symmetric difference $X \triangle \widetilde{X}$ has its endpoints in the closed neighborhood $N_X[U_{i+z} \setminus U_i]$. Now let H_0 be the single realization of $\boldsymbol{h}_0(n)$ and let G_1 and G_2 be the induced subgraphs of $X \triangle H_0$ and $\widetilde{X} \triangle H_0$ on the vertices in $N_X[U_{i+z} \setminus U_i]$, respectively. Since G_1 and G_2 are two different realizations of the same degree sequence, by the irreducibility of the swap Markov chain there is a swap sequence connecting them. Since the number of vertices of G_1 and G_2 is at most $2z+2k$, the length of the shortest swap sequence between them is clearly at most the total number of graphs on $2z+2k$ vertices, which only depends on z and k. Hence the proof of (4) is complete. □

Corollary 6.4. *Let k be a positive integer, X, Y be a pair of k-path systems. If for positive integers i, z the inequalities $0 \le i \le n-z$ and $2k \le z$ hold, then there is a unique k-path system $T_{X,Y}[i+1, i+z]$ with the following properties.*

- *The induced subgraphs of $T_{X,Y}[i+1, i+z]$ and Y on the vertices U_i are identical.*
- *The induced subgraphs of $T_{X,Y}[i+1, i+z]$ and X on the vertices $(A \cup B) \setminus U_{i+z}$ are identical.*
- *$T_{X,Y}[i+1, i+z]$ contains a buffer starting at $i+1$ and ending at $i+z$.*

6.2. **Constructing the canonical path $\gamma_{X,Y}$.** We will explicitly construct $2(n-2k)$ intermediate realizations along the swap sequence $\gamma_{X,Y}$. These realizations and the associated k-path systems are called **milestones**. Now we proceed with the definition of the milestones of the swap sequences between any pair of different k-path systems. Let X and Y be the two different k-path systems which we intend to connect. The swap sequence will go through the k-path systems $T_{X,Y}[i, i+2k-1]$, $T_{X,Y}[i, i+2k]$, $T_{X,Y}[i+1, i+2k]$ for each $i = 1, \ldots, n-2k$ in increasing order. This is shown on Figure 7.

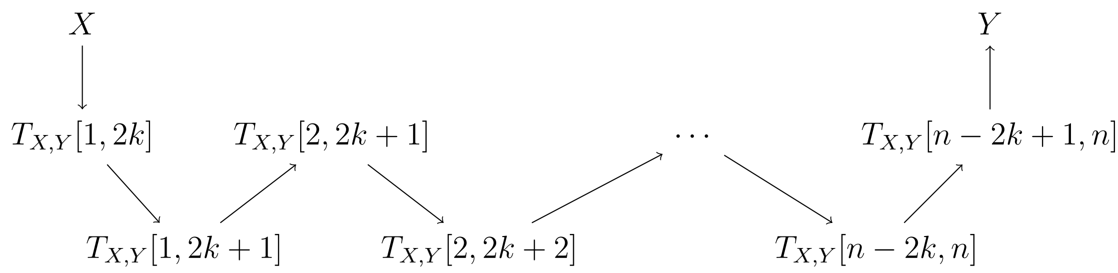

FIGURE 7. The swap sequence between X and Y with the milestones of the sequence, where each arrow represents a swap sequence whose existence is guaranteed by part (4) of Lemma 6.3.

To complete the proof of the existence the proposed swap sequence, observe, that

- $T_{X,Y}[1, 2k] = \widetilde{X}[1, 2k]$.
- Similarly, $T_{X,Y}[n - 2k + 1, n] = \widetilde{Y}[n - 2k + 1, n]$
- For $1 \leq i \leq n - 2k$, observe, that $T_{X,Y}[i, i + 2k] = \widetilde{T}[i, i + 2k]$ holds if $T = T_{X,Y}[i, i + 2k - 1]$ or $T = T_{X,Y}[i + 1, i + 2k]$.

By applying statement (4) of Lemma 6.3, the arrows in Figure 7 can be substituted with swap sequences of length at most c. Concatenating these short swap sequences and pruning the circuits from the resulting trail (so that any k-path system is visited at most once by it) produces the swap sequence $\gamma_{X,Y}$ connecting X to Y in the Markov graph.

6.3. **Assigning the encodings.** Each k-path system visited by $\gamma_{X,Y}$ receives an encoding that will be an ordered 4-tuple consisting of another k-path system, two graphs of constant size, and an integer in $\{1, \ldots, n\}$. For the two graphs of constant size, we need the following definition.

Definition 6.5 (left-compressed induced subgraph)**.** *Let X be a k-path system on the vertices $A \cup B$ and let $R \subset A \cup B$. Let us group the vertices $A \cup B$ into pairs: $\{(a_i, b_{n-i+1})\}_{i=1}^{n}$. Remove the pairs that do not intersect R, and let the remaining pairs be $\{(a_{i_j}, b_{n-i_j+1})\}_{j=1}^{r}$ for some $i_1 < \ldots < i_r$. For each edge of $E(X[R])$, map $a_{i_j} \mapsto a_j$ and $b_{n-i_j+1}) \mapsto b_{n-j+1}$ for all j simultaneously. This changes the embedding of the vertices of $X[R]$ in the plane, and we call this new graph the left-compressed copy of $X[R]$.*

To any realization on the swap sequence between the k-path systems X and Y we will assign an encoding

$$L_i(X, Y) := \Big(T_{Y,X}[i, i + 2k], G_X[i], G_Y[i], i\Big)$$

for some $1 \leq i \leq n - 2k$, where $G_X[i]$ and $G_Y[i]$ are defined as follows. Let $G_X[i]$ be the left-compressed subgraph of X induced by $N_X[U_{i+2k} \setminus U_{i-1}]$. Similarly, $G_Y[i]$ is the left-compressed copy of subgraph of Y induced $N_Y[U_{i+2k} \setminus U_{i-1}]$. An encoding is assigned to each realization along the swap sequence $\gamma_{X,Y}$ as follows:

- The encoding $L_1(X, Y)$ is used from the beginning X of the swap sequence until it reaches $T_{X,Y}[2, 2k + 1]$ (including this realization).
- For $2 \leq i \leq n - 2k - 1$, the encoding L_i is used between $T_{X,Y}[i, i + 2k - 1]$ (not included) and $T_{X,Y}[i + 1, i + 2k]$ (included).
- The encoding L_{n-2k} is used between $T_{X,Y}[n - 2k + 1, n]$ (not included) and Y.

6.4. **Estimating the load $\rho(\Gamma)$.** The total number of possible encodings is at most

$$\mathcal{O}_k(|\mathcal{G}(\boldsymbol{h}_k(n))| \cdot n)$$

(where the index k warns that this expression may depend on k), since the number of left-compressed graphs on at most $6k+1$ vertices is a constant depending only on k.

Lemma 6.6. *There are no two pairs (X_1, Y_1) and (X_2, Y_2) which assign the same encoding L to the same k-path system Z.*

Proof. We will prove that given the encoding L and k-path system Z, we can determine which (X, Y) pair put this encoding on Z. The first coordinate of L is a k-path system, of the form $T_{Y,X}[i, i+2k]$ for an unknown X, Y and i. From the fourth coordinate of L we get the value i. First we recover X. By the structure of $T_{Y,X}[i, i+2k]$, and the knowledge of i, we also know the induced subgraph of X on the vertices U_{i-1}. Similarly, the induced subgraph of Z on the vertices $(A \cup B) \setminus U_{i+2k}$ is identical to the induced subgraph of X on the same vertices. Hence the only unknown part of X is its induced subgraph on $N_X[U_{i+2k} \setminus U_{i-1}]$. The subgraph in the second coordinate of $L_i(X, Y)$ is the left-compressed copy of $X[N_X[U_{i+2k+1} \setminus U_{i-1}]]$. Since left-compression preserves the order of the x-coordinates, we can fully recover X.

It is easy to see that Y can be recovered by a parallel, symmetric procedure. □

Proof of Theorem 1.6. We have shown that $\rho(\Gamma) = \mathcal{O}(n \cdot n^4)$ and $\ell(\Gamma) = \mathcal{O}(n)$, thus $\tau(\varepsilon) \leq \mathcal{O}(n^8 \log \varepsilon^{-1})$, verifying that the swap Markov chain is rapidly mixing on $\mathcal{H}_k$. □

7. Concluding remarks

7.1. **Relationship to prior results.** Although the sets of degree sequences $\mathcal{H}_k$ (for some k) are certainly not diverse compared to the class of P-stable degree sequences, they are more numerous than, say, the regular degree sequences, for which rapid mixing of the swap Markov chain were first proven in [2, 12, 8]. Because $\mathcal{H}_k$ is not P-stable, the Jerrum-Sinclair chain [10] cannot produce a sample in expected polynomial time. Although in principle, the proof of rapid mixing on P-stable degree sequences [5] may apply to $\mathcal{H}_k$, we do not expect that it can be easily tweaked to accommodate $\mathcal{H}_k$, for the following reasoning:

Let $\mathcal{T}$ be the set of (X, Y) pairs of realizations of $\boldsymbol{h}_1(n)$ such that the paths $H_0(n) \triangle X$ and $H_0(n) \triangle Y$ are edge disjoint. It is simple to show that $|\mathcal{T}| \geq \exp(cn) \cdot |\mathcal{G}(\boldsymbol{h}_1(n))|$, because for almost every realization X we have $|E(H_0(n) \triangle X) \approx \frac{2n}{\sqrt{5}}$. For a pair $(X, Y) \in \mathcal{T}$, the edges $E(X) \triangle E(Y)$ form a cycle which traverses both a_n and b_1. From this structure it follows that the multicommodity flow Γ described in [5] between a pair of realizations $(X, Y) \in \mathcal{T}$ is a single swap sequence that passes through $H_0(n) - a_n b_1 \in \mathcal{G}(\boldsymbol{h}_1(n))$. Consequently, the load $\rho(\Gamma) \geq |\mathcal{T}|/|\mathcal{G}(\boldsymbol{h}_1(n))| \geq \exp(cn)$ is exponential in n.

7.2. **Previous attempts.** The proof of Theorem 1.6 is actually the fifth proof that the authors found where the swap Markov chain is rapidly mixing on a non-P-stable family of degree sequences. The first four unpublished proofs were based on the following ideas.

(1) The first proof only verifies rapid mixing on $\mathcal{H}_1$. The description is relatively long and the heavy lifting is done by Theorem 4.3 of [4] and a coupling argument.
(2) The second proof is also applies only to $\mathcal{H}_1$, but it is a pure coupling argument.
(3) The third proof uses Sinclair's multicommodity flow method to show rapid mixing on $\mathcal{H}_1$.

(4) The fourth proof verifies rapid mixing on $\mathcal{H}_k$ for all fixed k, via the canonical paths method. The main difference from the presented proof of Theorem 1.6 is that instead of moving a small buffer to "rewrite" X to Y, that proof moved the entirety of X to the right, and created Y on the left.

The first three proofs resisted our attempts to generalize them even to $\mathcal{H}_2$. In contrast, the proof of Theorem 1.6 presented in Section 6 works almost verbatim up to $k = \Theta(\sqrt{\log n})$ (one just has to check the dependence on k in Section 6.4). We have not proved nor refuted P-stability of $\mathcal{H}_k$ when $k = \Theta(\sqrt{\log n})$.

Question 7.1. *Is the swap Markov chain rapidly mixing on $\mathcal{H}_k$ when $k = \Theta(\sqrt{n})$?*

7.3. **Possible generalizations.** Let $\mathcal{H}_k'$ be the family of those degree sequences whose realizations have geometric representations which can be decomposed into exactly k edge-disjoint x-monotone paths (the difference from $\mathcal{H}_k$ being that the endpoints of the paths are not necessarily the bottom left and top right vertex). It is a simple exercise to generalize the proof of Theorem 1.6 to yield the rapid mixing of the swap Markov chain for degree sequences in $\mathcal{H}_k'$ for any fixed k. The authors conjecture that the proof can be extended to a much more general family of degree sequences.

A defining property of $\boldsymbol{h}_k(n)$ is that in any geometric representation of a realization of it, exactly k edges cross the vertical line S_i with $x = i/2 - 1/4$, where $i = 1, \ldots, 2n-1$. Suppose that instead of exactly k, we require that exactly $t_i \leq k$ edges cross S_i for every $i = 1, \ldots, 2n-1$, and we define $\mathcal{H}_{\leq k}$ to be the set of degree sequences satisfying these constraints.

Conjecture 7.2. *For any fixed k, the swap Markov chain is rapidly mixing for the degree sequences in $\mathcal{H}_{\leq k}$.*

References

[1] G. Amanatidis and P. Kleer. "Rapid Mixing of the Switch Markov Chain for Strongly Stable Degree Sequences and 2-Class Joint Degree Matrices". In: *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*. 0 vols. Proceedings. Society for Industrial and Applied Mathematics, Jan. 1, 2019, pp. 966–985. DOI: 10.1137/1.9781611975482.60.

[2] Colin Cooper, Martin Dyer, and Catherine Greenhill. "Sampling Regular Graphs and a Peer-to-Peer Network". In: *Combinatorics, Probability and Computing* 16.4 (July 2007), pp. 557–593. ISSN: 1469-2163, 0963-5483. DOI: 10.1017/S0963548306007978.

[3] Rick Durrett. *Probability: Theory and Examples*. Google-Books-ID: evbGTPhuvSoC. Cambridge University Press, Aug. 30, 2010. 441 pp. ISBN: 978-1-139-49113-6.

[4] P. Erdös, I. Miklós, and Z. Toroczkai. "A Decomposition Based Proof for Fast Mixing of a Markov Chain over Balanced Realizations of a Joint Degree Matrix". In: *SIAM Journal on Discrete Mathematics* 29.1 (Jan. 1, 2015), pp. 481–499. ISSN: 0895-4801. DOI: 10.1137/130929874.

[5] Péter L. Erdős, Catherine Greenhill, Tamás Róbert Mezei, István Miklós, Dániel Soltész, and Lajos Soukup. "The mixing time of the swap (switch) Markov chains: a unified approach". In: *arXiv:1903.06600 [cs, math]* (Mar. 15, 2019). arXiv: 1903.06600.

[6] Péter L. Erdős, Sándor Z. Kiss, István Miklós, and Lajos Soukup. "Approximate Counting of Graphical Realizations". In: *PLOS ONE* 10.7 (July 10, 2015), e0131300. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0131300.

[7] Péter L. Erdős, Tamás Róbert Mezei, István Miklós, and Dániel Soltész. "Efficiently sampling the realizations of bounded, irregular degree sequences of bipartite and directed graphs". In: *PLOS ONE* 13.8 (Aug. 13, 2018), e0201995. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0201995.

[8] Catherine S. Greenhill. “A Polynomial Bound on the Mixing Time of a Markov Chain for Sampling Regular Directed Graphs”. In: *Electr. J. Comb.* 18 (2011).

[9] Catherine Greenhill and Matteo Sfragara. “The switch Markov chain for sampling irregular graphs and digraphs”. In: *Theoretical Computer Science* 719 (Apr. 6, 2018), pp. 1–20. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2017.11.010.

[10] Mark Jerrum and Alistair Sinclair. “Fast uniform generation of regular graphs”. In: *Theoretical Computer Science* 73.1 (June 8, 1990), pp. 91–100. ISSN: 0304-3975. DOI: 10.1016/0304-3975(90)90164-D.

[11] Ravi Kannan, Prasad Tetali, and Santosh Vempala. “Simple Markov-chain algorithms for generating bipartite graphs and tournaments”. In: *Random Structures & Algorithms* 14.4 (1999), pp. 293–308. ISSN: 1098-2418. DOI: 10.1002/(SICI)1098-2418(199907)14:4<293::AID-RSA1>3.0.CO;2-G.

[12] István Miklós, Péter L. Erdős, and Lajos Soukup. “Towards random uniform sampling of bipartite graphs with given degree sequence”. In: *Electronic Journal of Combinatorics* 20.1 (2013), Paper 16, 51. ISSN: 1077-8926.

[13] A. Sinclair. *Algorithms for Random Generation and Counting: A Markov Chain Approach.* Progress in Theoretical Computer Science. Birkhäuser Basel, 1993. ISBN: 978-0-8176-3658-6.

[14] Alistair Sinclair. “Improved Bounds for Mixing Rates of Markov Chains and Multicommodity Flow”. In: *Combinatorics, Probability and Computing* 1.4 (Dec. 1992), pp. 351–370. ISSN: 1469-2163, 0963-5483. DOI: 10.1017/S0963548300000390.