

# The mixing time of the swap (switch) Markov chains: a unified approach

Péter L. Erdős<sup>a,1</sup>, Catherine Greenhill<sup>b,3</sup>, Tamás Róbert Mezei<sup>a,1</sup>, István Miklós<sup>a,1,2</sup>, Dániel Soltész<sup>a,1</sup>, Lajos Soukup<sup>a,4</sup>

<sup>a</sup>*Alfréd Rényi Institute of Mathematics, Reáltanoda u 13–15 Budapest, 1053 Hungary*  
`<erdos.peter,mezei.tamas.robert,miklos.istvan,soukup.lajos>@renyi.mta.hu`

<sup>b</sup>*School of Mathematics and Statistics, UNSW Sydney, NSW 2052, Australia*  
`c.greenhill@unsw.edu.au`

---

## Abstract

Since 1997 a considerable effort has been spent to study the mixing time of **swap** (switch) Markov chains on the realizations of graphic degree sequences of simple graphs. Several results were proved on rapidly mixing Markov chains on unconstrained, bipartite, and directed sequences, using different mechanisms. The aim of this paper is to unify these approaches. We will illustrate the strength of the unified method by showing that on any  $P$ -stable family of unconstrained/bipartite/directed degree sequences the swap Markov chain is rapidly mixing. This is a common generalization of every known result that shows the rapid mixing nature of the swap Markov chain on a region of degree sequences. Two applications of this general result will be presented. One is an almost uniform sampler for power-law degree sequences with exponent  $\gamma > 2$ . The other one shows that the swap Markov chain on the degree sequence of an Erdős-Rényi random graph  $G(n, p)$  is asymptotically almost surely rapidly mixing.

*Keywords:* rapidly mixing MCMC, Sinclair’s multicommodity flow method, restricted degree sequences

---

## 1. Introduction

An important problem in network science is to algorithmically construct typical instances of networks with predefined properties. In particular, special attention has

---

<sup>1</sup>The authors were supported in part by the National Research, Development and Innovation Office — NKFIH grant K 116769 and KH 126853

<sup>2</sup>IM was supported in part by the National Research, Development and Innovation Office — NKFIH grant SNN 116095.

<sup>3</sup>Research supported by the Australian Research Council, Discovery Project DP140101519.

<sup>4</sup>SL was supported in part by the National Research, Development and Innovation Office — NKFIH grants K 113047 and K 129211.

been devoted to sampling simple graphs with a given degree sequence. In this paper only graphs without parallel edges and loops are considered and we restrict our study to degree sequences which have at least one realization (**graphic**). We study the three most common degree sequence types: bipartite degree sequences, directed degree sequences and the usual degree sequences which we call **unconstrained**, or **UC degree sequence** for short.

In 1997 Kannan, Tetali, and Vempala ([22]) proposed the use of the so-called switch or **swap Markov chain** approach for uniformly sampling realizations of a degree sequence. For all three degree sequence types, the swap Markov chain can be thought of as the Markov chain of smallest possible modifications. To illustrate this, we give an informal description of the swap Markov chain on UC degree sequences. If  $G_1, G_2$  are two realizations of the same UC degree sequence, it is easy to see that the minimal size of the symmetric difference  $E(G_1) \Delta E(G_2)$  is four. We say that  $G_1$  and  $G_2$  differ by a swap if this symmetric difference is exactly four. The states of the swap Markov chain are the realizations of the degree sequence and the probability of going from realization  $G_1$  to  $G_2$  is nonzero if and only if they differ by a swap (and this nonzero quantity is independent of  $G_1$  and  $G_2$ ). For the precise definition of this chain, and for the definition of the chains for other degree sequence types, we refer the reader to Section 3 (unconstrained and bipartite) and Section 5 (directed).

The following conjecture has been named after Kannan, Tetali, and Vempala, in recognition of their pioneering work.

**Conjecture 1.1** (the KTV conjecture). *The swap Markov chain is rapidly mixing for any bipartite, directed, or UC degree sequence.*

To give some context to the Conjecture, we say that a Markov chain is rapidly mixing if the distribution on the state space is close in  $\ell_1$  norm to the unique stationary distribution after  $\text{poly}(\log N)$  steps, where  $N$  is the size of the state space. This property means that sampling the state space with the stationary distribution is a more or less tractable problem, even if the state space has exponential size.

It is not uncommon that uniformly randomly applied, small local modifications of combinatorial objects result in rapid mixing. This is the case for solutions of the 0–1 knapsack problem [26], for the union of perfect and almost perfect matchings of a graph [19], and two-rowed contingency tables [4], for example. In both of these cases, applying the smallest possible modifications of the respective combinatorial objects randomly, result in rapid mixing of the corresponding Markov chain.

Although Conjecture 1.1 is still open, there is a series of results that prove the rapid mixing of the swap Markov chain on various special degree sequence classes. We summarize these results in a very compact way in Table 1 without presenting the sometimes lengthy definitions of the special classes, which can be found in the references provided. Most rapid mixing results on directed degree sequences can be reduced to the case of bipartite degree sequences, as shown in [7]. Since this is also

the case in the present paper, we will not discuss directed degree sequences until Section 5.

UC degree sequences	bipartite deg. seq.	directed deg. seq.
regular [3]	(half-)regular [25]	regular [14]
	almost half regular [7]	
$\Delta \leq \frac{1}{3}\sqrt{2m}$ [16]	$\Delta \leq \frac{1}{\sqrt{2}}\sqrt{m}$ [10]	$\Delta < \frac{1}{\sqrt{2}}\sqrt{m-4}$ [10]
Power-law density-bound, $\gamma > 2.5$ [16]		
$(\Delta - \delta + 1)^2 \leq 4 \cdot \delta(n - \Delta - 1)$ [1]	$(\Delta - \delta)^2 \leq \delta(\frac{n}{2} - \Delta)$ [9, 10] (or Corollary 18 in [1])	similar to the bip. case [9, 10]
	Bip. E.R. with edge prob. $p, 1-p \geq 4\sqrt{\frac{2\log n}{n}}$ [9, 10]	similar to the bip. case [9, 10]
strongly stable degree sequence classes [1]		

Table 1: Some classes of degree sequences for which the swap Markov chain is rapidly mixing. Here  $\Delta$  and  $\delta$  denote the maximum and minimum degrees, respectively. Half of the sum of the degrees is  $m$ , and  $n$  is the number of vertices. The notation is similar for bipartite and directed degree sequences. Some technical conditions have been omitted.

Notice, that some, but not all of the results came in pairs for unconstrained and bipartite (directed) degree sequences. The reason for this discrepancy is the following: while both set of results are based on Sinclair’s multicommodity flow method, one of them has to deal with special circuits (one of the vertices may be visited at most twice) instead of just cycles in the decomposition of symmetric differences of two realizations of a degree sequence. The main goal of this paper to remedy this discrepancy between the machineries used for the bipartite and unconstrained degree sequences by decomposing into circuits where each vertex is visited at most twice. Along the line we also give new, much more transparent proofs for the main results in [25].

Greenhill and Sfragara suggested exploring the connection between the mixing rate of the swap Markov chain and stable degree sequences [16, Subsection 1.1]. The first such result is due to Amanatidis and Kleer [1], who showed rapid mixing of the swap Markov chain on *strongly stable* unconstrained and bipartite degree sequences. We will return to this notion at the beginning of Section 8.

Let us now give an informal definition of the notion of stability that we study in

this paper:  $P$ -stability. We will only focus on UC degree degree sequences here, for a complete formal definition, see Section 7. We say that a class of UC degree sequences is  $P$ -stable if there is a fixed polynomial  $p(x)$  and the number of realizations of any degree sequence in the class can not grow by more than a  $p(n)$  factor (where  $n$  is the number of vertices) by moving to another degree sequence within  $\ell_1$  distance at most 2. An even more informal description is that when we slightly perturb the degree sequence from a  $P$ -stable class, the number of realization can not grow too much.

To the best of our knowledge, all of the previously known cases where Conjecture 1.1 holds are  $P$ -stable classes. The main contribution of the paper is developing a unified machinery that proves all of the previous results as a special case.

**Theorem 1.2** (proved in Section 7). *The swap Markov chain is rapidly mixing on  $P$ -stable unconstrained, bipartite, and directed degree sequence classes.*

The unified framework in which we prove Conjecture 1.1 for  $P$ -stable unconstrained, bipartite, and directed degree sequences allows us to prove Theorem 1.2 with minimal branching.

There are two interesting direct consequences of Theorem 1.2 concerning popular unconstrained random graph models. It turns out that asymptotically almost surely, the degree sequence of an Erdős-Rényi random graph  $G(n, p)$  is  $P$ -stable, see Corollary 8.6.

In [13], the authors claim that every power-law distribution-bounded degree sequence with parameter  $\gamma > 2$  is  $P$ -stable. Consequently, Theorem 1.2 implies that the swap Markov chain is rapidly mixing on all real world power-law like degree sequences, solving a decades old open problem (see Section 8.2).

The proof of Theorem 1.2 relies on Sinclair’s multicommodity flow method (Section 3), which can be described informally in the case of the swap Markov chain as follows. Suppose that the chain has  $N$  states, and let  $\mathbb{G}$  be the graph on them as vertices where two states are adjacent in  $\mathbb{G}$  if the transition probability between them is non-zero in the chain. Sinclair’s multicommodity flow method ensures the rapid mixing of the chain if we can design a multicommodity flow on  $\mathbb{G}$  which transfers a unit amount of commodities between each pair of vertices (different commodities for different pairs), and no more than  $N \cdot \text{poly}(\log(N))$  amount of commodities go through every vertex (no vertex is overloaded). Hence most of the present paper is devoted to the design of a flow and the proof that it does not overload any vertex.

The paper is structured as follows. In Section 2, we give a slightly more detailed description of the flow and we present the necessary graph theoretic tools with which we will use to construct paths that will form the flow. In Section 3 we give the formal definition of Sinclair’s multicommodity flow method and its simplified version that is tailored to our needs. In Section 4 we finally describe the mulicommodity flow. In Section 5 we define the swap Markov chain for directed degree sequences. Before

turning into the home straight, an auxiliary structure that tracks the defined flow is analyzed in Section 6. In Section 7 we finally prove Theorem 1.2 and we also provide the necessary modifications to deal with bipartite and directed degree sequences. Lastly, we describe the known  $P$ -stable regions of degree sequences in Section 8 and present the connections between Theorem 1.2 and the aforementioned popular graph models.

## 2. Definitions and preliminaries, the structure of the sets of realizations.

Let us recall some well known notions and notations. Let  $\mathbf{d} = (d(v_1), \dots, d(v_n))$  denote a UC degree sequence and let

$$\mathbf{D} = (\mathbf{d}(U), \mathbf{d}(V)) = ((d(u_1), \dots, d(u_n)), (d(v_1), \dots, d(v_m)))$$

denote a bipartite degree sequence on the bipartition  $(U, V)$ . (For convenience we assume that  $n \geq m$ .) As it was mentioned earlier we assume that all degree sequences are **graphic**. We will use the notations  $\mathbb{G}(\mathbf{d})$  and  $\mathbb{G}(\mathbf{D})$  for the sets of all realizations of the corresponding degree sequences.

The **swap operation** exchanges two disjoint edges  $ac$  and  $bd$  in the realization  $G$  with  $ad$  and  $bc$  if the resulting configuration  $G'$  is again a simple graph (we denote the operation by  $ac, bd \Rightarrow ad, bc$ ). In our terminology, a switch is a matrix operation, which will be introduced in Section 6.

For an  $ac, bd \Rightarrow ad, bc$  swap operation to be valid, it is necessary but not always sufficient that both  $ac, bd \in E(G)$  and  $ad, bc \notin E(G)$  hold. We will use the name **chord** for any vertex pair  $u, v$  where  $uv$  is allowed to be an edge in a realization of some degree sequences (even when we do not know or do not care whether it is an edge or a non-edge in the current realization). Consequently, if a pair of vertices is not a chord, which we call a **non-chord**, then the two vertices are forbidden to form an edge. We emphasize that whether or not a pair of vertices form a **chord** is entirely arbitrary in the sense that it only depends on the class of graphs to which the model should be restricted to.

We reformulate the definition of the swap operation to avoid touching non-chords: an  $ac, bd \Rightarrow ad, bc$  swap operation can be applied if  $ac, bd \in E(G)$ ,  $ad, bc \notin E(G)$ , and  $ad, bc$  are both chords. We now define the set of chords and non-chords in the case of unconstrained and bipartite graph models.

**Definition 2.1.** For simple graphs, the non-chords are exactly the pairs of the form  $(v, v)$ , as loops are forbidden. Because no further constraints have to be set, we call their degree sequences **unconstrained**. In **bipartite** graphs,  $(u, v)$  is a chord if and only if  $u$  and  $v$  are in different vertex classes.

In the case of directed graphs (Section 5), we further restrict the set of **chords**.

It is a well-known fact that the set of all possible realizations of a **graphic** UC degree sequence is connected under the swap operation. See for example [18] or [17].

It is interesting to know, however, that the first known proof is from 1891 [27]. For bipartite graphs the equivalent results were proved in 1957 in [11] and [28]. The “classical” proofs work through so called “canonical” realizations. However, the paths between different realizations, created in this way, are often very far from shortest possible. Therefore in this paper we will use another way to design these paths. To that end, let us consider two realizations of the same (bipartite or UC degree) degree sequence. To any alternating circuit decomposition of the symmetric difference of the realizations, we are going to assign a sequence of swaps that transform the first realization into the second (this is described right after the proof of Lemma 2.5). If the given decomposition contains a maximum number of elementary circuits, then the sequence of swaps will be the shortest possible, see [5, Theorem 3.6].

A graph  $H$ , with edges colored by either red or blue, will be called a **red-blue graph**. For vertex  $v$  let  $d_r(v)$  and  $d_b(v)$  be the degree of vertex  $v$  in red and blue edges, respectively. This red-blue graph is **balanced** if for each  $v \in V(H)$  equality  $d_r(v) = d_b(v)$  holds.

Let  $G, G'$  both be realizations (on the same vertex set) of an unconstrained degree sequence  $\mathbf{d}$  or a bipartite degree sequence  $\mathbf{D}$ . Let the symmetric difference of the edges be

$$\nabla = E(G) \triangle E(G').$$

Color the edges of  $\nabla$  according to which graph they come from: the  $E(G)$  edges are colored red and the  $E(G')$  edges are colored blue. Equipped with this coloring,  $\nabla$  is a balanced red-blue graph.

A **circuit** in a graph  $H$  is a closed trail (also known as closed walk). As the graph is simple, a circuit is determined by the sequence of the vertices  $v_0, \dots, v_t$ , where  $v_0 = v_t$ . Note that there can also be other indices  $i < j$  such that  $v_i = v_j$ . A circuit is called a **cycle**, if its simple, i.e., for any  $i < j$ ,  $v_i = v_j$  only if  $i = 0$  and  $j = t$ .

A circuit (or, in particular, a cycle) in a balanced red-blue graph is called **alternating**, if the color of its edges alternates. In other words, the color of the edge from  $v_i$  to  $v_{i+1}$  differs from the color of the edge from  $v_{i+1}$  to  $v_{i+2}$ , and also edges  $v_0v_1$  and  $v_{t-1}v_t$  have different colors. Consequently, alternating circuits have even length. The following observations are easy to see.

**Lemma 2.2** (adapted from [5]).

- (i) *If  $H$  is a balanced red-blue graph then the edge set can be decomposed into alternating circuits.*
- (ii) *Let  $C = v_0, v_1, \dots, v_{2t} = v_0$  be an alternating circuit in a balanced red-blue graph  $H$ , in which for some  $i < j < 2t$ ,  $j - i$  is even and  $v_i = v_j$ . Then the circuit can be decomposed into two shorter alternating circuits.*

- (iii) *If  $B$  is a bipartite balanced red-blue graph then the edge set can be decomposed into alternating cycles.*

It is clearly possible that a vertex occurs twice in an alternating circuit without the possibility to divide it into two, smaller alternating circuits. The smallest example is a “bow-tie” circuit:  $v_1, v_2, v_3, v_1, v_4, v_5, v_1$  with an alternating edge coloring. (The very first and very last occurrences of  $v_1$  shows the closing of the alternating circuit.) Recalling our earlier discussion, these two copies of the vertex  $v_1$  form a **non-chord**.

**Definition 2.3.** An alternating circuit is **elementary**, if it cannot be decomposed into shorter alternating circuits.

From Lemma 2.1 it follows, that in an elementary alternating circuit, no vertex can appear more than twice, moreover, the distance of two copies of the same vertex must be odd. This definition is slightly weaker than the original one in [5].

**Lemma 2.4.** *Let  $C$  be an alternating circuit of length 6 in  $\nabla$ . If loops are non-chords, then there is at most one vertex which is visited more than once by  $C$ .*

*Proof.* If  $v$  is a vertex that is visited at least twice by  $C$ , it has at least four other neighbors that are pairwise distinct from each other and  $v$ . Since  $v$  is counted twice in the length of  $C$ , every visit of  $C$  is accounted for, and the claim holds (and  $C$  is a bow-tie).  $\square$

We will use Sinclair’s multi-commodity flow method (Theorem 3.2) to bound the mixing time of the swap Markov chain. The multi-commodity flow is given by a set of swap sequences between any two realizations of the degree sequence. The main idea behind the definition of the flow can be described roughly as follows:

We will decompose  $\nabla$  into alternating circuits in every possible way, and subsequently each decomposition is further refined into elementary alternating circuits in a canonical way (Section 4). The flow is built by concatenating the paths that we obtain for the elementary alternating circuits via Algorithm 2.2 as described in Lemma 2.5. This way we obtain a path for each alternating circuit decomposition, and we spread out the flow of the commodity evenly over them.

The SWEEP procedure in Algorithm 2.2 will be used to construct the swap sequence between two realizations whose symmetric difference cannot be decomposed any further. It calls two subroutines, SWAP and DOUBLE STEP, which are described by Algorithm 2.1. Addition and subtraction of edges naturally means that we add the edge to or remove the edge from the edge set of the first operand.

**Lemma 2.5.** *Suppose that  $E(\nabla)$  is an elementary alternating circuit  $C$  of length  $2\ell$ , and  $(x_1, x_2, \dots, x_{2\ell})$  is a list of its vertices such that  $x_1 x_2 \notin E(G)$ . Then Algorithm 2.2 provides a valid swap sequence of length  $\ell - 1$  between  $G$  and  $G'$  in the case of unconstrained and bipartite graph models. In the bipartite case, DOUBLE STEP is never called.*

---

**Algorithm 2.1** Swap and Double step (which is composed of two swaps)

---

**function** SWAP( $G, x_1, [x_{2t}, x_{2t+1}, x_{2t+2}]$ )  
     **return**  $G + (x_1x_{2t} - x_1x_{2t+2}) + (-x_{2t}x_{2t+1} + x_{2t+1}x_{2t+2})$   
**end function**

**Ensure:**  $x_{2t-2}x_{2t+1}$  is a chord

**function** DOUBLE STEP( $G_0, x_1, [x_{2t-2}, x_{2t-1}, x_{2t}, x_{2t+1}, x_{2t+2}]$ )  
     **if**  $x_{2t-2}x_{2t+1} \in E(G)$  **then**  
          $G_1 \leftarrow G_0 + (-x_{2t-2}x_{2t+1} + x_{2t+1}x_{2t+2} - x_1x_{2t+2} + x_1x_{2t-2})$   
          $G_2 \leftarrow G_1 + (+x_{2t-2}x_{2t+1} - x_{2t-2}x_{2t-1} + x_{2t-1}x_{2t} - x_{2t}x_{2t+1})$   
     **else if**  $x_{2t-2}x_{2t+1} \notin E(G)$  **then**  
          $G_1 \leftarrow G_0 + (+x_{2t-2}x_{2t+1} - x_{2t-2}x_{2t-1} + x_{2t-1}x_{2t} - x_{2t}x_{2t+1})$   
          $G_2 \leftarrow G_1 + (-x_{2t-2}x_{2t+1} + x_{2t+1}x_{2t+2} - x_1x_{2t+2} + x_1x_{2t-2})$   
     **end if**  
     **return**  $G_1, G_2$   
**end function**

---



---

**Algorithm 2.2** Sweeping an elementary circuit

---

**Ensure:**  $x_1x_2 \notin E(G)$  and  $x_1, x_2, \dots, x_{2\ell}$  is an alternating elementary circuit

**procedure** SWEEP( $G, [x_1, x_2, \dots, x_{2\ell}] \rightarrow [Z_1, Z_2, \dots, Z_{\ell-1}]$ )  
      $Z_0 \leftarrow G$   
      $q \leftarrow 1$   
      $endChord \leftarrow 2$   
     **while**  $endChord < 2\ell$  **do**  
          $startChord \leftarrow \min \{2i \in 2\mathbb{N} : 2i > endChord \text{ and } x_1x_{2i} \in E(G)\}$   
          $2t \leftarrow startChord - 2$   
         **while**  $2t \geq endChord$  **do**  
             **if**  $x_1x_{2t}$  is a chord **then**  
                  $Z_q \leftarrow \text{SWAP}(Z_{q-1}, x_1, [x_{2t}, x_{2t+1}, x_{2t+2}])$   
                  $q \leftarrow q + 1$   
                  $2t \leftarrow 2t - 2$   
             **else if**  $x_1x_{2t}$  is a non-chord **then**  
                  $Z_q, Z_{q+1} \leftarrow \text{DOUBLE STEP}(Z_{q-1}, x_1, [x_{2t-2}, \dots, x_{2t+2}])$   
                  $q \leftarrow q + 2$   
                  $2t \leftarrow 2t - 4$   
             **end if**  
         **end while**  
          $endChord \leftarrow startChord$   
     **end while**  
**end procedure**

---



*Proof.* The processing done by Algorithm 2.2 is governed by two nested loops. The outer loop iterates the variable *startChord* through  $\{2i : 4 \leq 2i \leq 2\ell, x_1x_{2i} \in E(G)\}$  in increasing order. Since  $x_1x_{2\ell} \in E(G)$ , the set is not empty. In the first iteration, *endChord* = 2, and in the successive iterations *endChord* takes the value taken by *startChord* in the previous iteration.

The inner loop performs a series of swaps that changes the status of the edges and non-edges induced by consecutive vertices in the interval of vertices between  $x_{startChord}, \dots, x_{endChord}$ . As a side effect, it also changes chords induced by  $x_1$  and one of the vertices from the list. We have to check that each time the functions SWAP and DOUBLE STEP (both are in Algorithm 2.1) are called by SWEEP, they indeed perform one and two valid swaps, respectively.

Because  $C$  is alternating, in the bipartite case,  $x_1$  and  $x_{2\ell}$  are in different vertex classes, so DOUBLE STEP is never called. In the case of unconstrained degree sequences, Lemma 2.4 implies that DOUBLE STEP does not use a non-chord:  $x_{2t-2}x_{2t+1}$  must be a chord when  $x_1x_{2t}$  is a non-chord.

We now make two important observations.

- (1) If  $x_i = x_j$ , then either  $i = j$  or  $i \not\equiv j \pmod{2}$ , because of Lemma 2.2(ii).
- (2) If  $x_ix_j \in \nabla$  such that  $i \not\equiv j \pmod{2}$  and  $i \not\equiv j \pm 1 \pmod{2\ell}$ , then there is a shorter alternating cycle through  $x_ix_j$ , because  $x_i$  and  $x_j$  cuts the original circuit into two alternating paths of odd length.

Notice, that SWAP and DOUBLE STEP only add or remove chords whose endpoints have indices of different parity. The two observations guarantee that if both  $x_{2i}x_{2j+1}$  and  $x_{2k}x_{2l+1}$  **takes part** in a swap during SWEEP, then the two chords are only equal if  $i = k$  and  $j = l$  (this is not trivial, because a vertex may have two copies  $x_{2i} = x_{2l+1}$  in the circuit). Therefore elements of  $E(C)$  only take part in exactly one swap during a SWEEP operation, while other chords take part in exactly zero or two swaps. The rest of the proof that Algorithm 2.2 provides a valid swap sequence is now an easy exercise.  $\square$

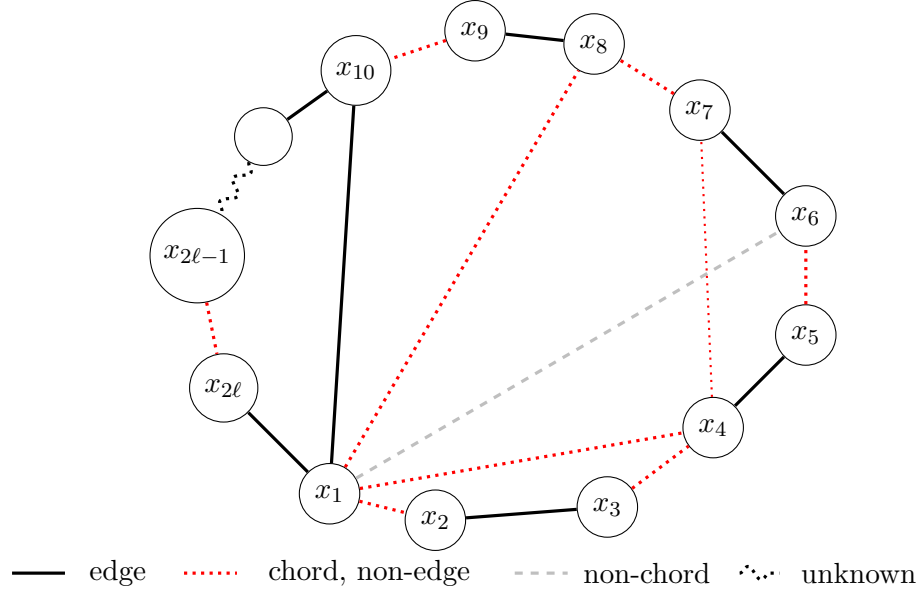


Figure 1: Sweeping a cycle

We will demonstrate the algorithm on Figure 1. In the first iteration of the outer loop, *startChord* takes 10 as its value. We call  $x_1x_{10}$  the **start-chord** and  $x_1x_2$  the **end-chord**. The algorithm **sweeps** the alternating chords along the circuit between  $x_2$  and  $x_{10}$ , and vertex  $x_1$  will be the **cornerstone** of this procedure.

The inner loop works from the start-chord  $x_1x_{10}$  (edge) towards the end-chord  $x_1x_2$  (non-edge). The first value taken by  $2t$  is 8. Since  $x_1x_8$  is a chord,  $Z_1$  is obtained by swapping along  $x_1, x_8, x_9, x_{10}$ . In the next step,  $2t = 6$ . However,  $x_1x_6$  is a non-chord, therefore SWEEP calls DOUBLE STEP instead of SWAP. Because  $x_4x_7$  is not an edge,  $Z_2$  is obtained by swapping along  $x_4, x_7, x_6, x_5$ , and subsequently  $Z_3$  is obtained by swapping along  $x_1, x_4, x_7, x_8$ . The last iteration of the inner loop swaps along  $x_1, x_2, x_3, x_4$  and produces  $Z_4$ . Notice, that all of the chords on the circuit from  $x_2$  to  $x_{10}$  changed their status and  $x_1x_{10}$  is no longer an edge (that is, in  $Z_4$ ), but the rest of the chords have the same status in  $Z_4$  as they had in  $G$ .

For the second iteration of the outer loop, *endChord* = 10 and *startChord* is assigned a new value too. Eventually, *startChord* =  $2\ell$ , which marks the last iteration of the outer loop, at the end of the algorithm produces  $Z_{\ell-1} = G \triangle \nabla = G'$ .

The demonstration shows that some chords that are not in  $\nabla$  change from being an edge to a non-edge and vica versa during this procedure. However, there are strict patterns that these irregularities must abide, as shown by the next lemma.

**Lemma 2.6.** *Suppose  $Z_q$  is an intermediate realization produced by Algorithm 2.2 on the swap-sequence between  $G$  and  $G'$ , when  $\nabla = E(G) \triangle E(G') = (x_1, x_2, \dots, x_{2\ell})$*

is an elementary alternating circuit. Let

$$R = (Z_q \triangle G) \setminus \nabla. \quad (2.1)$$

The following statements hold at the moment when  $Z_q$  is assigned a graph in SWEEP.

- (a)  $R$  is a set of chords induced by vertices of  $\nabla$ ,
- (b)  $R = \emptyset$  for  $Z_0$  and  $Z_{\ell-1}$ ,
- (c)  $R = \{x_1 x_{startChord}, x_1 x_{endChord}\} \triangle \{x_1 x_{2t}\}$  when  $Z_q$  gets its value from SWAP,
- (d)  $R = \{x_1 x_{startChord}, x_1 x_{endChord}\} \triangle \{x_1 x_{2t-2}\}$  when  $Z_q$  is the second graph returned by DOUBLE STEP and  $x_1 x_{2t}$  is a non-chord,
- (e)  $R = \{x_1 x_{startChord}, x_1 x_{endChord}\} \triangle \{x_1 x_{2t-2}, x_{2t-2} x_{2t+1}\}$  if  $Z_q$  is the first graph returned by DOUBLE STEP,  $x_{2t-2} x_{2t+1} \in E(G)$ , and  $x_1 x_{2t}$  is a non-chord.
- (f)  $R = \{x_1 x_{startChord}, x_1 x_{endChord}\} \triangle \{x_1 x_{2t+2}, x_{2t-2} x_{2t+1}\}$  if  $Z_q$  is the first graph returned by DOUBLE STEP,  $x_{2t-2} x_{2t+1} \notin E(G)$ , and  $x_1 x_{2t}$  is a non-chord.
- (g)  $(Z \triangle R) \triangle G$  is a subset of  $\nabla$ , composed of walks (of at most 3 for unconstrained, and at most 2 for bipartite graphs) starting and ending at endpoints of chords in  $R$ .

*Proof.* Each statement is easy to show via induction over the iterations of the outer and inner loops of SWEEP.  $\square$

Let  $G$  and  $G'$  be two realizations and assume that we can decompose the symmetric difference  $\nabla$  into  $k$  elementary circuits (cycles). Then SWEEP can process all elementary alternating circuits one by one, therefore it can transform  $G$  into  $G'$  with  $\frac{|\nabla|}{2} - k$  swap operations. The process only changes the status of a chord induced by vertices of the current circuit.

Here we reached a very important point: Algorithm 2.2 does not require an order for processing the elementary circuits; in principle it can be done arbitrarily. One of the novelties of this paper leading to the unified proof is constructing a very delicate order of processing the circuits. We will return to this point in Section 4.

### 3. Sinclair's multicommodity flow method

For UC degree sequences we define our Markov chain  $(\mathbb{G}_d, P_d)$  as follows: in the Markov graph  $\mathbb{G}_d(\mathbf{V}_d, \mathbb{E}_d)$  the pair  $(G, G')$  is an edge if these two realizations differ in exactly one swap. To make a move, choose an unordered pair of two distinct, non-adjacent edges uniformly at random from  $G$ , say  $F = \{(x, y), (z, w)\}$  and choose a perfect matching  $F'$  from the other two perfect matchings on the same four vertices.

If  $F \subseteq E(G)$  and  $F' \cap E(G) = \emptyset$ , then perform the swap (so  $E(G') = E(G) \cup F' \setminus F$ ). Assuming that  $P(G, G') \neq 0$  and  $G \neq G'$ , we have

$$\text{Prob}(G \rightarrow G') = P(G, G') := \frac{1}{2 \binom{n}{2} \binom{n-2}{2}}. \quad (3.1)$$

Equation (3.1) immediately gives that the Markov chain is symmetric. Notice, that if  $(F, F')$  corresponds to a feasible swap, then  $(F', F)$  does not, therefore  $P(G, G) \geq \frac{1}{2}$  for any realization  $G$ . A Markov chain possessing this property of staying in the current state with probability at least  $\frac{1}{2}$  is called **lazy**. Laziness implies that the eigenvalues of the transition matrix of the Markov chain are non-negative, and it also implies that the chain is aperiodic.

For bipartite degree sequences we define our Markov chain  $(\mathbb{G}_{\mathbf{D}}, P_{\mathbf{D}})$  as follows: in the Markov graph  $\mathbb{G}_{\mathbf{D}}(\mathbf{V}_{\mathbf{D}}, \mathbb{E}_{\mathbf{D}})$  the pair  $(G, G')$  is an edge, if these two realizations differ in exactly one swap. The transition matrix  $P$  is defined as follows: we choose uniformly two-two vertices  $u_1, u_2; v_1, v_2$  from classes  $U$  and  $V$ , respectively, and uniformly randomly choose one of the two matchings between the two pairs. If it preserves the degree sequence, we remove the chosen matching, and add the other. The swap moving from  $G$  to  $G'$  is unique, therefore the probability of this transformation (the *jumping probability* from  $G$  to  $G' \neq G$ ) is:

$$\text{Prob}(G \rightarrow G') = P(G, G') := \frac{1}{2 \binom{n}{2} \binom{m}{2}}. \quad (3.2)$$

The transition probabilities are time- and edge-independent, and symmetric. Also, the entries in the main diagonal are at least  $\frac{1}{2}$ , so the chain is lazy.

To start with we recall some definitions and notations from the literature. Since the stationary distribution of the swap Markov chain is the uniform distribution, we will not state the results we use in full generality. Let  $P^t$  denote the  $t^{\text{th}}$  power of the transition probability matrix and define

$$\Delta_X(t) := \frac{1}{2} \sum_{Y \in V(\mathbb{G})} |P^t(X, Y) - 1/N|,$$

where  $X$  is an element of the state space of the Markov chain and  $N$  is the size of the state space. We define the *mixing time* as

$$\tau_X(\varepsilon) := \min_t \{ \Delta_X(t') \leq \varepsilon \text{ for all } t' \geq t \}.$$

Our Markov chain is said to be *rapidly mixing* iff

$$\tau_X(\varepsilon) \leq O\left(\text{poly}(\log(N/\varepsilon))\right)$$

for any  $X$  in the state space. In this case the swap Markov chain method provides is a *fully polynomial almost uniform sampler* (FPAUS) of the realizations of the given degree sequences. Note, that Jerrum and Sinclair have shown that the perfect matchings of realizations of  $P$ -stable degree sequences have a fully polynomial almost uniform sampler [20]

Consider the different eigenvalues of  $P$  in non-increasing order:

$$1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_N \geq -1.$$

The *relaxation time*  $\tau_{rel}$  is defined as

$$\tau_{rel} = \frac{1}{1 - \lambda^*}$$

where  $\lambda^*$  is the *second largest eigenvalue modulus*,

$$\lambda^* := \max\{\lambda_2, |\lambda_N|\}.$$

However, the eigenvalues of any lazy Markov chain are non-negative, so we do know that  $\lambda^* = \lambda_2$  for our Markov chain. The following result was proved implicitly by Diaconis and Strook in 1991, and explicitly stated by Sinclair [29, Proposition 1]:

**Theorem 3.1** (Sinclair).  $\tau_X(\varepsilon) \leq \tau_{rel} \cdot \log(N/\varepsilon)$ .  $\square$

So one way to prove that our Markov chain is rapidly mixing is to find a  $\text{poly}(\log N)$  upper bound on  $\tau_{rel}$ . We need rapid convergence of the process to the stationary distribution otherwise the method cannot be used in practice.

There are several different methods to prove fast convergence, here we use — similarly to [22] — Sinclair’s *multicommodity flow method* ([29, Theorem 5’]).

**Theorem 3.2** (Sinclair). *Let  $\mathbb{H}$  be a graph whose vertices represent the possible states of a time reversible finite state Markov chain  $\mathcal{M}$ , and where  $(U, V) \in E(\mathbb{H})$  iff the transition probabilities of  $\mathcal{M}$  satisfy  $P(U, V)P(V, U) \neq 0$ . For all  $X \neq Y \in V(\mathbb{H})$  let  $\Gamma_{X,Y}$  be a set of paths in  $\mathbb{H}$  connecting  $X$  and  $Y$  and let  $\pi_{X,Y}$  be a probability distribution on  $\Gamma_{X,Y}$ . Furthermore let*

$$\Gamma := \bigcup_{X \neq Y \in V(\mathbb{H})} \Gamma_{X,Y}$$

where the elements of  $\Gamma$  are called **canonical paths**. We also assume that there is a stationary distribution  $\pi$  on the vertices  $V(\mathbb{H})$ . We define the capacity of an edge  $e = (W, Z)$  as

$$Q(e) := \pi(W)P(W, Z)$$

and we denote the length of a path  $\gamma$  by  $|\gamma|$ . Finally let

$$\kappa_\Gamma := \max_{e \in E(\mathbb{H})} \frac{1}{Q(e)} \sum_{\substack{X, Y \in V(\mathbb{H}) \\ \gamma \in \Gamma_{X,Y} : e \in \gamma}} \pi(X) \pi(Y) \pi_{X,Y}(\gamma) |\gamma|. \quad (3.3)$$

Then

$$\tau_{\text{rel}}(\mathcal{M}) \leq \kappa_\Gamma \quad (3.4)$$

holds.  $\square$

We are going to apply Theorem 3.2 for  $(\mathbb{G}, P)$ , which is either the unconstrained  $(\mathbb{G}, P) = (\mathbb{G}(\mathbf{d}), P_{\mathbf{d}})$  or the bipartite  $(\mathbb{G}, P) = (\mathbb{G}(\mathbf{D}), P_{\mathbf{D}})$  swap Markov chain. Using the notation  $N := |V(\mathbb{G})|$ , the (uniform) stationary distribution has the value  $\pi(X) = 1/N$  for each vertex  $X \in V(\mathbb{G})$ . Furthermore each transition probability has the property  $P(X, Y) \geq 1/n^4$  (see (3.1) and (3.2)). So if we can design a multicommodity flow such that each path is shorter then an appropriate  $\text{poly}(n)$  function, then simplifying inequality (3.3) we can turn inequality (3.4) to the form:

$$\tau_{\text{rel}} \leq \frac{\text{poly}(n)}{N} \left( \max_{e \in E(\mathbb{H})} \sum_{\substack{X, Y \in V(\mathbb{H}) \\ \gamma \in \Gamma_{X,Y} : e \in \gamma}} \pi_{X,Y}(\gamma) \right). \quad (3.5)$$

If  $Z \in e$ , then

$$\sum_{\substack{X, Y \in V(\mathbb{H}) \\ \gamma \in \Gamma_{X,Y} : e \in \gamma}} \pi_{X,Y}(\gamma) \leq \sum_{\substack{X, Y \in V(\mathbb{H}) \\ \gamma \in \Gamma_{X,Y} : Z \in \gamma}} \pi_{X,Y}(\gamma), \quad (3.6)$$

so we have

$$\tau_{\text{rel}} \leq \frac{\text{poly}(n)}{N} \left( \max_{Z \in V(\mathbb{H})} \sum_{\substack{X, Y \in V(\mathbb{H}) \\ \gamma \in \Gamma_{X,Y} : Z \in \gamma}} \pi_{X,Y}(\gamma) \right). \quad (3.7)$$

We make one more assumption. Namely, that for each pair of realizations  $X, Y \in V(\mathbb{G})$  there is a non-empty finite set  $S_{X,Y}$  (which draws its elements from a pool of symbols) and for each  $s \in S_{X,Y}$  there is a path  $\Upsilon(X, Y, s)$  from  $X$  to  $Y$  such that

$$\Gamma_{X,Y} = \{\Upsilon(X, Y, s) : s \in S_{X,Y}\}. \quad (3.8)$$

It can happen that  $\Upsilon(X, Y, s) = \Upsilon(X, Y, s')$  for  $s \neq s'$ , so we consider  $\Gamma_{X,Y}$  as a “multiset” and so we should take

$$\pi_{X,Y}(\gamma) = \frac{|\{s \in S_{X,Y} : \gamma = \Upsilon(X, Y, s)\}|}{|S_{X,Y}|}$$

for  $\gamma \in \Gamma_{X,Y}$ .

Putting together the observations and simplifications above we obtain the

**Simplified Sinclair’s method:**

For each  $X \neq Y \in V(\mathbb{G})$  find a non-empty finite set  $S_{X,Y}$  and for each  $s \in S_{X,Y}$  find a path  $\mathcal{T}(X,Y,s)$  from  $X$  to  $Y$  such that

- each path is shorter than an appropriate  $\text{poly}(n)$  function,
- for each  $Z \in V(\mathbb{G})$

$$\sum_{X,Y \in V(\mathbb{G})} \frac{|\{s \in S_{X,Y} : Z \in \mathcal{T}(X,Y,s)\}|}{|S_{X,Y}|} \leq \text{poly}(n) \cdot N. \quad (3.9)$$

Then our Markov chain  $(\mathbb{G}, P)$  is rapidly mixing.

#### 4. Multicommodity flow — general considerations

Let  $X$  and  $Y$  be two realizations of the same (unconstrained or bipartite) degree sequence, by notation they are  $\in \mathbb{G}$ . The high level description of the definition of multicommodity flow from  $X$  to  $Y$  can go like this:

- (Step 1) We decompose the symmetric difference  $\nabla = E(X) \triangle E(Y)$  into alternating circuits:  $W_1, W_2 \dots, W_{k_s}$ .
- (Step 2) We decompose every alternating circuit  $W_i$  into smaller, “simple” alternating circuits  $C_1^i, C_2^i \dots, C_{k_i}^i$ .
- (Step 3) We will construct the canonical path from  $X$  to  $Y$  along these “simple” alternating circuits, step by step, using Algorithm 2.2 iteratively.

Typically, the successful application of the Sinclair’s method requires decomposing  $\nabla$  into alternating circuits (Step 1) in very many ways, and each decomposition requires one canonical path. These different decompositions will be parameterized by the set  $S_{X,Y}$  (see (3.8)). This parametrization (described in details in Lemma 4.1) and its application to (Step 1) was introduced in [22]. Now we arrived at the most sensitive part of the construction: (Step 2). Here we need the following ability:

Let  $Z$  denote an arbitrary vertex along a canonical path. To apply Sinclair’s method we will need that the elements of  $S_{X,Y}$  can be reconstructed from elements of  $S_{\nabla \cap E(Z), \nabla \setminus E(Z)}$  (using another small parameter set). In case of UC degree sequences (at the current setting, see for example [3] or [16]) these simpler alternating circuits have the following property: in each “simple” circuit  $C$  there is one predefined vertex (actually, the smallest vertex in a predefined full order), which occurs at most twice in  $C$ . This makes the reconstruction above relatively simple, but

makes (Step 3) more complicated. In the case of bipartite degree sequences (see, for example, [25] or [7]) the decomposition in (Step 2) became finer: there each vertex occurs at most once. That made the reconstruction method more complicated but it makes (Step 3) much more simple and capable. To adapt this method to the UC degree sequences we cannot expect to be able to decompose into alternating cycles (bow tie!). Instead we use the following idea.

(Step  $\widehat{2}$ ) We decompose every alternating circuit  $W_i$  into elementary alternating circuits  $C_1^i, C_2^i, \dots, C_{k_i}^i$ .

This makes the reconstruction process much more demanding but it provides more verifying power (as it is shown by the new results). In case of bipartite realizations the elementary circuits were also cycles, but here we will extend this approach for UC degree sequences and will provide the analogue results.

#### 4.1. (Step 1): parameterizing the circuit decomposition

Now let  $K = (W, F \cup F')$  be a simple graph where  $F \cap F' = \emptyset$  and assume that for each vertex  $w \in W$  the  $F$ -degree and  $F'$ -degree of  $w$  are the same:  $d(w) = d'(w)$  for all  $w \in W$ . An *alternating circuit decomposition* of  $(F, F')$  is a circuit decomposition such that no two consecutive edges of any circuit are in  $F$  or in  $F'$ . By definition, that means that each circuit is of even length. Next we are going to parameterize the alternating circuit decompositions.

The set of all edges in  $F$  (in  $F'$ ) which are incident to a vertex  $w$  is denoted by  $F(w)$  (by  $F'(w)$ , respectively).

If  $A$  and  $B$  are sets, denote by  $[A, B]$  the complete bipartite graph with classes  $A$  and  $B$ . Let

$$\mathbb{S}(F, F') = \left\{ s : s \text{ is a function, } \text{dom}(s) = W, \text{ and for all } w \in W \right. \\ \left. s(w) \text{ is a 1-factor of the complete bipartite graph } [F(w), F'(w)] \right\}. \quad (4.1)$$

**Lemma 4.1.** *There is a natural one-to-one correspondence between the family of all alternating circuit decompositions of  $(F, F')$  and the elements of  $\mathbb{S}(F, F')$ .*

*Proof.* If  $\mathcal{C} = \{C_1, C_2, \dots, C_n\}$  is an alternating circuit decomposition of  $(F, F')$ , then define  $s_{\mathcal{C}} \in \mathbb{S}(F, F')$  as follows:

$$s_{\mathcal{C}}(w) := \left\{ ((w, u), (w, u')) \in [F(w), F'(w)] : \right. \\ \left. (w, u) \text{ and } (w, u') \text{ are consecutive edges in some } C_i \in \mathcal{C} \right\}. \quad (4.2)$$

On the other hand, to each  $s \in \mathbb{S}(F, F')$  assign an alternating circuit decomposition

$$\mathcal{C}_s = \{W_1^s, W_2^s, \dots, W_{k_s}^s\}$$



of  $(F, F')$  as follows: Consider the bipartite graph  $\mathcal{F} = (F, F'; R(s))$ , where

$$R(s) = \{((u, w), (u', w)) : w \in W \text{ and } ((u, w), (u', w)) \in s(w)\}.$$

$\mathcal{F}$  is a 2-regular graph because for each edge  $(u, v) \in F \cup F'$  there is exactly one  $(u, w) \in F \cup F'$  with  $((u, w), (u, w)) \in s(u)$ , there is exactly one  $(t, v) \in F \cup F'$  with  $((u, v), (t, v)) \in s(v)$ , therefore the  $\mathcal{F}$ -neighbors of  $(u, v)$  are  $(u, w)$  and  $(t, v)$ .

$\mathcal{F}$  is a 2-regular, so it is the union of vertex disjoint cycles  $\{W_i^s : i \in I\}$ . Now  $W_i^s$  can also be viewed as a sequence of edges in  $F \cup F'$ , which is an alternating circuit in  $\langle W, F \cup F' \rangle$ , so  $\{W_i^s : i \in I\}$  is an alternating circuit decomposition of  $(F, F')$ . Since

$$s_{C_s} = s,$$

we proved the Lemma.  $\square$

If the  $F$ -degree sequence (and therefore the  $F'$ -degree sequence) is  $d_1, \dots, d_k$ , then write

$$t_{F, F'} = \prod_{i=1}^k (d_i!).$$

Clearly

$$|\mathbb{S}(F, F')| = t_{F, F'}.$$

#### 4.2. Preparing for (Step 2): the $T$ -operator

For bipartite degree sequences (Step 2) simply required a cycle decomposition which was provided by the  $T$ -operator defined in Section 5.2 of [25]. For UC degree sequences the best we may hope is a decomposition into elementary circuits (See [5]). In this subsection we generalize the  $T$ -operator so that it can process any balanced red-blue graph. When this balanced red-blue graph is bipartite, the generalized and the original  $T$ -operator produce the same decomposition of alternating cycles. The new proof described in this section is simpler than that of [25] because it is described on a higher level of abstraction.

Let  $[m]$  be a base set, denote  $S_{[m]}$  the symmetric group on  $[m]$  and let  $Pos$  denote the set of **Positions**, where  $Pos = \{(1)^+, (2)^+, \dots, (m-1)^+\}$ . For convenience we also allow the alternating naming  $Pos = \{(2)^-, \dots, (m)^-\}$ . So we consider  $(i)^+ = (i+1)^-$ . Let  $f$  be a two-coloring on  $Pos$  with  $f \in \{green, red\}^{Pos}$ . We will describe the state of our system with the pair

$$(\pi, f) \quad : \quad \pi \in S_{[m]}, \quad f \in \{green, red\}^{Pos}.$$

Let  $\mathcal{E} \subset \binom{[m]}{2}$  be a fixed subset which we call the set of **eligible reversals**. Assume that

$$\text{the connected components of } G = ([m], \mathcal{E}) \text{ are cliques.} \quad (4.3)$$

It is important to recognize, that the eligible reversals belong to the elements of the base set, and they do not depend on the actual  $\pi$  value of those elements. Accordingly, to make the definitions more readable, let us define

$$\pi^{-1}(\mathcal{E}) = \left\{ \{ \pi^{-1}(x), \pi^{-1}(y) \} : \{x, y\} \in \mathcal{E} \right\}.$$

We now define an operator  $T_{\mathcal{E}}$ , or  $T$  for short, as the index  $\mathcal{E}$  is fixed anyway. This  $T$  is a function mapping  $S_{[m]} \times \{green, red\}^{Pos}$  into itself. To determine the image of  $(\pi, f)$  under  $T$ , an interval will be selected first. For that end let

$$j_{(\pi, f)} := \min \left\{ j' \mid \exists i' < j' : f((i')^+) = f((j')^-) = green, \{i', j'\} \in \pi^{-1}(\mathcal{E}) \right\},$$

then let

$$i_{(\pi, f)} := \max \left\{ i' < j_{(\pi, f)} \mid f((i')^+) = green, \{i', j_{(\pi, f)}\} \in \pi^{-1}(\mathcal{E}) \right\}.$$

We define  $\max \emptyset = -\infty$  and  $\min \emptyset = +\infty$ . For an integer  $k : 1 \leq k \leq m$  we define two positions from  $Pos$ . Let  $a_{(\pi, f)}(k) := k$  if  $f((k)^-) = green$  and

$$a_{(\pi, f)}(k) := \min \left\{ i' \leq k \mid \forall i'' \text{ s.t. } i' \leq i'' < k : f((i'')^+) = red \right\}$$

otherwise. Furthermore let  $b_{(\pi, f)}(k) := k$  if  $f((k)^+) = green$  and

$$b_{(\pi, f)}(k) := \max \left\{ j' \geq k \mid \forall j'' \text{ s.t. } k < j'' \leq j' : f((j'')^-) = red \right\}$$

otherwise.

We are ready now to define the image of the pair  $(\pi, f)$  under the operator  $T$ . If  $j_{(\pi, f)} = +\infty$ , then let  $(\pi, f)$  be a fixed point of the  $T$  operator, so its image is itself. Otherwise define  $T : (\pi, f) \mapsto (\pi', f')$  as follows. For any  $k \in [1, m]$ , let  $\overleftarrow{k}^{(\pi, f)} := a_{(\pi, f)}(k) + b_{(\pi, f)}(k) - k$ . We omit the index  $(\pi, f)$  in the following.

$$\begin{aligned} \pi'(k) &= \begin{cases} \pi(k) & \text{if } k \notin [a(i), b(j)], \\ \overleftarrow{k}^{(\pi, f)} & \text{if } k \in [a(i), b(j)], \end{cases} \\ f'((k)^+) &= \begin{cases} f((k)^+) & \text{if } 1 \leq k < a(i) \text{ or } b(j) \leq k < m, \\ red & \text{if } a(i) \leq k < b(j). \end{cases} \end{aligned}$$

We can expand the definition of  $\pi'(k)$  to a more accessible form:

$$\pi'(k) = \begin{cases} \pi(k) & \text{if } k \notin [a(i), b(j)], \\ \pi(a(j) + (k - i)) & \text{if } a(i) \leq k < i, \\ \pi(a(i) + (k - j)) & \text{if } j < k \leq b(j), \\ \pi(a(i) + b(j) - k) & \text{if } i \leq k \leq j. \end{cases}$$

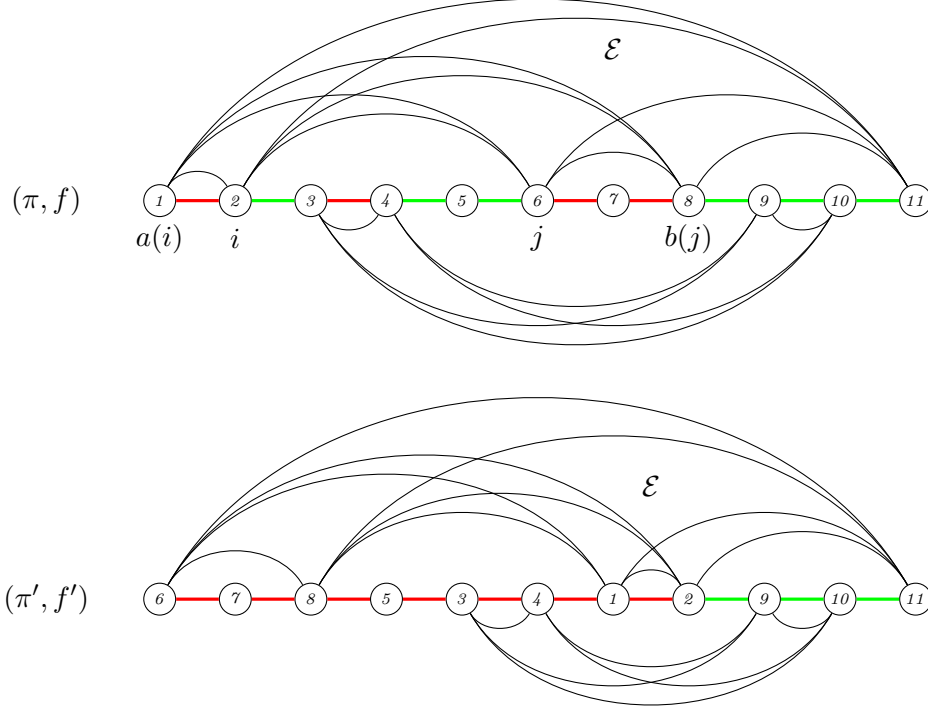


Figure 2: An example for  $T(\pi, f) = (\pi', f')$ . The curved arcs represent the pairs in  $\mathcal{E}$ . The encircled numbers are  $\pi(x)$  and  $\pi'(x)$ , respectively, where  $x$  is the first coordinate of the center of the circle.

In yet another form, we extend the interval  $\{(i_{(\pi,f)})^+, \dots, (j_{(\pi,f)})^-\}$  to the maximal containing integer interval  $\{(a_{(\pi,f)}(i_{(\pi,f)}))^+, \dots, (b_{(\pi,f)}(j_{(\pi,f)}))^-\}$  such that the  $f$ -image of the new positions of the extended interval are *red*. Take a look at Figure 2. To construct  $\pi'$  from  $\pi$ , every maximal red interval  $\{x^+, \dots, y^-\}$  in  $\{a^+, \dots, b^-\}$  is shifted to  $\{(a + b - y)^+, \dots, (a + b - x)^-\}$ , and the green positions are taken in reverse order in the remaining positions between the shifted red intervals.

**Lemma 4.2.** *For any  $(\pi, f)$  which is not a fixed point of the  $T$  operator, we have  $j_{T(\pi,f)} > j_{(\pi,f)}$ .*

*Proof.* Since

$$(f')^{-1}(\text{green}) \subsetneq f^{-1}(\text{green}),$$

we have  $f((i_{(\pi',f')})^+) = f((j_{(\pi',f')})^-) = \text{green}$ . Clearly,  $\{\pi'(i_{(\pi',f')}), \pi'(j_{(\pi',f')})\} \in \mathcal{E}$ , so

$$\{\pi^{-1}(\pi'(i_{(\pi',f')})), \pi^{-1}(\pi'(j_{(\pi',f')}))\} \in \pi^{-1}(\mathcal{E}). \quad (4.4)$$

Since  $f'((j_{(\pi', f')})^-) = \text{green}$ , we must have  $j_{(\pi', f')} > b_{(\pi, f)}(j_{(\pi, f)})$  or  $j_{(\pi', f')} \leq a_{(\pi, f)}$ . In the first case, the proof is trivial, since  $b_{(\pi, f)}(j_{(\pi, f)}) \geq j_{(\pi, f)}$ .

If  $j_{(\pi', f')} < a_{(\pi, f)}$ , then  $\pi'(j_{(\pi', f')}) = \pi(j_{(\pi', f')})$  and  $\pi'(i_{(\pi', f')}) = \pi(i_{(\pi', f')})$ . Plugged into Equation (4.4), the definition of  $j_{(\pi, f)}$  implies that  $j_{(\pi', f')} \geq j_{(\pi, f)} > a_{(\pi, f)}$ , a contradiction.

The only remaining case to deal with is  $j_{(\pi', f')} = a_{(\pi, f)}$ . Then

$$\{\pi^{-1}(\pi'(i_{(\pi', f')})) , \pi^{-1}(\pi'(j_{(\pi', f')}))\} = \{i_{(\pi', f')}, b_{(\pi, f)}(j_{(\pi, f)})\}.$$

Using property (4.3) and Equation (4.4), we have

$$\{i_{(\pi', f')}, a_{(\pi, f)}(i_{(\pi, f)})\} \in \pi^{-1}(\mathcal{E}),$$

which is in contradiction with the choice of  $j_{(\pi, f)}$ .  $\square$

Let  $(\pi_r, f_r) = T^r(\pi_0, \mathbf{green})$ . In indices, we shorten  $(\pi_r, f_r)$  by writing  $r$  instead. For example,  $j_r = j_{(\pi_r, f_r)}$ , etc. Using this notation, the previous statement implies the following claim.

**Corollary 4.3.** *For any  $r \geq 0$ , we have  $b_r(j_r) = j_r$ . Moreover, for any  $j_{r-1} \leq k < m$  we have  $f_r((k)^+) = \text{green}$ .*

**Lemma 4.4.** *For arbitrary  $\pi_0$ ,  $r \geq 0$ , and  $k \in [1, m]$ , we have*

$$\pi_r(k) = \pi_0(a_{(\pi_r, f_r)}(k) + b_{(\pi_r, f_r)}(k) - k).$$

*Proof.* If  $r = 1$ , the statement immediately follows from the definition. Suppose the statement holds for  $r - 1$ . If  $k \notin [a_{r-1}(i_{r-1}), b_{r-1}(j_{r-1})]$  then  $a_{r-1}(k) = a_r(k)$  and  $b_{r-1}(k) = b_r(k)$ , so

$$\pi_r(k) = \pi_{r-1}(k) = \pi_0(a_{r-1}(k) + b_{r-1}(k) - k) = \pi_0(a_r(k) + b_r(k) - k),$$

as we wished.

Suppose from now on that  $k \in [a_{r-1}(i_{r-1}), b_{r-1}(j_{r-1})]$ . We have  $a_r(k) = a_{r-1}(i_{r-1})$  and  $b_r(k) = b_{r-1}(j_{r-1})$ . Let

$$l = a_{r-1}(i_{r-1}) + b_{r-1}(j_{r-1}) - k.$$

Since the edges in  $[a_{r-1}(i_{r-1}), b_{r-1}(j_{r-1})]$  are all red in  $f_r$ , we have

$$a_r(k) = a_{r-1}(i_{r-1}), \quad b_r(k) = b_{r-1}(j_{r-1}).$$

Again, by induction

$$\pi_r(k) = \pi_{r-1}(\overleftarrow{l}^{(r-1)}) = \pi_0(a_{r-1}(\overleftarrow{l}^{(r-1)}) + b_{r-1}(\overleftarrow{l}^{(r-1)}) - \overleftarrow{l}^{(r-1)}).$$

Since  $a_{r-1}(\overleftarrow{l}^{(r-1)}) = a_{r-1}(l)$  and  $b_{r-1}(\overleftarrow{l}^{(r-1)}) = b_{r-1}(l)$ , the right hand side is equal to  $\pi_0(l)$ . Expanding it, we get

$$\pi_0(l) = \pi_0(a_{r-1}(i_{r-1}) + b_{r-1}(j_{r-1}) - k) = \pi_0(a_r(k) + b_r(k) - k),$$

which is what we intended to prove.  $\square$

**Lemma 4.5.** *The pair of endpoints of a maximal path formed by elements of  $f_r^{-1}(\text{red})$  is an element of  $\pi_r^{-1}(\mathcal{E})$ .*

*Proof.* Use property (4.3) and the fact that  $f((i)^+) = f((j)^-) = \text{green}$ . By induction, either  $a_{(\pi,f)}(i_{(\pi,f)}) = i_{(\pi,f)}$  or  $\{a_{(\pi,f)}(i_{(\pi,f)}), i_{(\pi,f)}\} \in \pi^{-1}(\mathcal{E})$ . By definition,  $\{i_{(\pi,f)}, j_{(\pi,f)}\} \in \pi^{-1}(\mathcal{E})$ , thus we also have  $\{a_{(\pi,f)}(i_{(\pi,f)}), j_{(\pi,f)}\} \in \pi^{-1}(\mathcal{E})$ . The same argument goes through for  $j$  and  $b(j)$ .  $\square$

**Lemma 4.6.** *If  $\{1, m\} \in \mathcal{E}$ , then  $\exists s \in \mathbb{N}$  such that  $f_s^{-1}(\text{red}) = \text{Pos}$ .*

*Proof.* Lemma 4.5 implies that  $\{1, \min\{t : f_r((t)^+) = \text{green}\}\} \in \pi_r^{-1}(\mathcal{E})$ , therefore  $\{\min\{t : f_r((t)^+) = \text{green}\}, m\} \in \pi_r^{-1}(\mathcal{E})$ , except if  $f_r^{-1}(\text{red}) = \text{Pos}$  already.  $\square$

The following claim shows that we cannot have such eligible reversals  $\{x, y\}$  that  $x < y < j_r$  and  $f((x)^+) = \text{green}$  and  $f((y)^-) = \text{red}$ .

**Lemma 4.7.** *Given  $r \geq 0$  and any  $\{x, y\} \in \pi_r^{-1}(\mathcal{E})$  such that  $x < y$ , either  $\{x, y\} = \{i_r, j_r\}$ , or  $f((x)^+) = \text{red}$ , or  $y \geq j_r + 1$ .*

*Proof.* The lemma trivially holds for  $r = 0$ . Suppose now, that  $r \geq 1$ .

If  $f_r((x)^+) = f_r((y)^-) = \text{green}$ , then by definition  $y \geq j_r$ . If  $y \geq j_r + 1$ , the lemma holds. If  $y = j_r$ , then definition of  $i_r$  implies that  $x \leq i_r$ . If  $y = j_r$  and  $x < i_r$ , then  $x < a_r(i_r)$ . By property (4.3),  $\{x, a_r(i_r)\} \in \pi_r^{-1}(\mathcal{E})$  holds. Since  $f_r((a_r(i_r))^-) = \text{green}$ , we have a contradiction with the definition of  $j_r$ .

Suppose, that  $f_r((x)^+) = \text{green}$ ,  $f_r((y)^-) = \text{red}$ , and the lemma does not hold. By Corollary 4.3, we have  $y \leq j_{r-1}$ . Then we must also have  $x < a_{r-1}(i_{r-1})$  (otherwise  $f_r((x)^+) = \text{red}$ , a contradiction). Thus  $\pi_{r-1}^{-1}(\pi_r(x)) = x$ ,  $x < \pi_{r-1}^{-1}(\pi_r(y)) \leq j_{r-1}$ , and so  $\{x, \pi_{r-1}^{-1}(\pi_r(y))\} \in \pi_{r-1}^{-1}(\mathcal{E})$ . By induction, we should have  $f_{r-1}((x)^+) = \text{red}$ , which implies  $f_r((x)^+) = \text{red}$ , a contradiction.

We have checked and eliminated every possible case where the statement of the lemma is not satisfied.  $\square$

Let us define  $\text{greenify}(\pi, f) := (\pi, \mathbf{green})$ , i.e., the operator replaces the coloring  $f$  of  $\text{Pos}$  with the identical  $\text{green}$  coloring.

**Theorem 4.8.**  $\forall r \in \mathbb{N} \exists w \in \mathbb{N}$  and  $\exists g \in \{\text{green}, \text{red}\}^{\text{Pos}}$  such that

$$T^w \circ \text{greenify} \circ T^r(\pi_0, \mathbf{green}) = (\pi_0, g).$$

*Proof.* If  $f_r(Pos) \equiv \text{red}$ , then Lemma 4.5 implies that  $\{1, m\} \in \pi_r^{-1}(\mathcal{E})$ , therefore Lemma 4.6 provides the existence of an appropriate  $s$ .

If  $f_r^{-1}(\text{red})$  is composed of multiple components, then  $T^s$  will successively work in these components. Lemma 4.4 say that the order of elements in these components have been reversed in  $\pi_r$  compared to  $\pi_0$ . Outside these intervals, however,  $\pi_r$  is identical to  $\pi_0$ .

Because of Lemma 4.5, we see that Lemma 4.6 implies that the maximal *red* intervals will be completely processed after a certain number of steps. Lemmas 4.2 and 4.7 together imply that if the  $T$  operator starts working inside a component of  $f_r^{-1}(\text{red})$  then the next selected interval  $[i, j]$  will also be inside until the whole component becomes red again.  $\square$

#### 4.3. (Step $\widehat{2}$ ) — decomposing a circuit into elementary circuits

Given  $X, Y \in \mathbb{G}$  (we do not specify whether the degree sequence is unconstrained or bipartite), and  $s \in S_{X,Y}$ , we construct a path between  $X$  and  $Y$  in  $\mathbb{G}(d)$  as follows. The matching  $s$  decomposes  $\nabla = X \triangle Y$  into alternating circuits

$$W_1, \dots, W_{k_s}. \quad (4.5)$$

We will use our operator  $T$  to decompose the circuits above into elementary circuits with the ability to reconstruct (with some extra parameters) the realization  $X, Y$  and the original matching function  $s$  at any given moment.

Let  $W_k$  be an arbitrary but fixed circuit from the collection (4.5). Let  $v_1 v_2$  be the lexicographically first edge of  $W_k$  (with  $v_1$  precedes  $v_2$ ). Let the Eulerian trail induced by  $s$  starting on the edge  $v_1 v_2$  be  $v_1 v_2 v_3 v_4 \dots v_{|E(W_k)|} v_{|E(W_k)|+1}$ , where  $v_{|E(W_k)|+1} = v_1$ . Let  $m = |E(W_k)| + 1$ ,  $\pi_0 = \text{id}_{|E(W_k)|+1}$ ,  $f_0 \equiv \mathbf{green}$ , and

$$\mathcal{E} = \left\{ \{x, y\} \in \binom{[m]}{2} : v_x = v_y \text{ and } x \equiv y \pmod{2} \right\}. \quad (4.6)$$

By transitivity, this set possesses property (4.3), so we can apply the  $T$ -operator on  $\pi_0$  with  $\mathcal{E}$  as the set of eligible reversals. Let  $(\pi_r, f_r) = T^r(\pi_0, f_0)$ .

**Lemma 4.9.** *Visiting the vertices  $v_{\pi_r(1)} v_{\pi_r(2)} \dots v_{\pi_r(|E(W_k)|+1)}$  in this order is an Eulerian trail of  $W_k$  in the graph  $(V, \nabla)$ .*

*Proof.* Easily seen by induction on  $r$ . Lemma 4.5 and the definition of the  $T$  operator implies that we get  $\pi_r$  by reversing such intervals of the trail defined by  $\pi_{r-1}$  whose first and last vertices are identical. Consequently, every edge is visited by the new trail too.  $\square$

It is also clear, by definition, that  $(x)^+$  in the set  $Pos$  coincides with the  $x$ th edge along the Eulerian trail  $\pi_r$  in the graph  $(V, \nabla)$ . We will denote this edge with  $e_x^{(r)} := v_{\pi_r(x)} v_{\pi_r(x+1)}$

**Lemma 4.10.** *Let*

$$\begin{aligned} E_r^k &= E(X) \triangle \left( \bigcup_{i=1}^{k-1} E(W_i) \right) \triangle \left\{ e_x^{(r)} : (x)^+ \in f_r^{-1}(\text{red}) \right\}, \\ Z_r^k &= (V, E_r^k). \end{aligned}$$

*Then the graph  $Z_r^k$  is a realization from  $\mathbb{G}$  and the closed vertex sequence  $\pi_r$  designates an alternating chord circuit in it.*

In words, we get  $Z_r^k$  from  $X$  by exchanging edges with non-edges (and vice versa) in the following subsets of edges: any  $W_i$  for  $1 \leq i \leq k-1$  and any  $C_j^k$  for  $1 \leq j \leq r$ . For any  $k, r$  we call  $Z_r^k$  a **milestone**. Milestones are special realizations: both  $Z_r^k \triangle X$  and  $Z_r^k \triangle Y$  are subgraphs of  $X \triangle Y$ . Milestones are uniquely determined by  $(X, Y, s)$  and the fixed lexicographical order.

**Lemma 4.11.** *For any  $r \in \mathbb{N}$ , we can describe  $\pi_r^{-1}(\mathcal{E})$  as the set of endpoints of even circuits formed by subintervals of the Eulerian trail in  $(V, E_r^k)$ , defined by  $\pi_r$ :*

$$\pi_r^{-1}(\mathcal{E}) = \left\{ \{x, y\} \in \binom{[m]}{2} : v_{\pi_r(x)} = v_{\pi_r(y)} \text{ and } x \equiv y \pmod{2} \right\}.$$

*Proof.* From (4.6), we have

$$\begin{aligned} \pi_r^{-1}(\mathcal{E}) &= \left\{ \{\pi_r^{-1}(x), \pi_r^{-1}(y)\} : \{x, y\} \in \binom{[m]}{2}, v_x = v_y \text{ and } x \equiv y \pmod{2} \right\} \\ &= \left\{ \{x, y\} \in \binom{[m]}{2} : v_{\pi_r(x)} = v_{\pi_r(y)} \text{ and } \pi_r(x) \equiv \pi_r(y) \pmod{2} \right\}, \end{aligned}$$

because  $\pi_r$  is a permutation. It is enough to show that  $\pi_r$  preserves parity, i.e.,  $\pi_r(k) \equiv k \pmod{2}$  for any  $k$ . For  $r = 0$  this is trivial. Suppose  $\pi_{r-1}$  preserves parity. For  $k \notin [a_r(i_r), b_r(j_r)]$ , we have  $\pi_r(k) \equiv \pi_{r-1}(k) \pmod{2}$ , so parity is preserved.

For  $k \in [a_r(i_r), b_r(j_r)]$ , first observe that Lemma 4.5 implies  $a_r(x) \equiv b_r(x) \pmod{2}$  for arbitrary  $x$ . Moreover,  $a_r(i_r) \equiv i_r \equiv j_r \equiv b_r(i_r) \pmod{2}$ , thus

$$\begin{aligned} \pi_r(k) &\equiv \\ &\equiv \pi_{r-1} \left( a_r(a_r(i_r) + b_r(j_r) - k) + b_r(a_r(i_r) + b_r(j_r) - k) - (a_r(i_r) + b_r(j_r) - k) \right) \\ &\equiv \pi_{r-1}(a_r(i_r) + b_r(j_r) - k) \equiv \pi_{r-1}(k) \equiv k, \pmod{2} \end{aligned}$$

which is what we desired.  $\square$

Let  $\ell_k$  be the maximum for which  $\pi_{\ell_k-1} \neq \pi_{\ell_k}$ . For any  $1 \leq r \leq \ell_k$ , let

$$E(C_r^k) = \left\{ e_x^{(r)} : (x)^+ \in f_r^{-1}(\text{red}) \setminus f_{r-1}^{-1}(\text{red}) \right\} = \quad (4.7)$$

$$= \left\{ e_x^{(r)} : (x \in [i_{r-1}, j_{r-1} - 1]) \wedge ((x)^+ \in f_{r-1}^{-1}(\text{green})) \right\}. \quad (4.8)$$

Take the list of edges of  $W_k$  starting with  $v_1v_2$  in the order defined by the Eulerian trail  $\pi_0 = \text{id}_{[m]}$ . This order can be restricted to the edges of  $C_r^k$ , so there is a natural Eulerian trail on  $C_r^k$ , too.

**Lemma 4.12.**  $\bigcup_{r=1}^{\ell_k} C_r^k$  is an alternating elementary circuit decomposition of  $W_k$ .

*Proof.* Observe that  $f_r$  defines a coloring of the edges of  $W_k$ : the edge  $v_lv_{l+1}$  has color  $f_r((l)^+)$ . Moreover, as  $r$  increases, red edges stay red. As  $W_k$  is an alternating circuit,  $|E(W_k)|$  is divisible by two, so  $\{1, |E(W_k)| + 1\} \in \mathcal{E}$ . Lemma 4.6 implies that  $\bigcup_{r=1}^{\ell_k} C_r^k$  is indeed an edge disjoint partition of  $W_k$ . Furthermore,  $C_r^k$  is an alternating circuit because of the definition of  $\mathcal{E}$  and Lemma 4.5.

Suppose  $C_r^k$  visits some vertex three times, that is

$$\exists x < y < z \text{ s.t. } i_r \leq x, y, z < j_r \text{ and } v_{\pi_r(x)} = v_{\pi_r(y)} = v_{\pi_r(z)}.$$

The proof of Lemma 4.11 shows that  $\pi_r$  preserves parity. If  $x \equiv y \pmod{2}$ , then  $y \geq j_{r-1}$ , a contradiction. Similarly, we must have  $y \not\equiv z \pmod{2}$  and  $x \not\equiv z \pmod{2}$ . In any case, we have a contradiction.

Similarly, if an even number of steps lead from one copy of a vertex to another copy of it on  $C_r^k$ , then  $j_{r-1}$  is not minimal, contradiction.  $\square$

Observe, that an alternating elementary circuit on a bipartite graph is a cycle. The produced decomposition is identical to the one described in [25].

#### 4.4. (Step 3) — Describing the swap sequence along an elementary alternating circuit

In this subsection we will construct swap sequences to transform one milestone realization into the next one, using Algorithm 2.2. Recursive application of this procedure will provide the entire swap sequence between realizations  $X$  and  $Y$ .

From Lemma 4.10 and Lemma 4.12, it follows that for any  $1 \leq k \leq k_s$  and  $1 \leq r \leq \ell_k$  we have

$$E_r^k = E_{r-1}^k \triangle E(C_r^k).$$

Clearly,  $X = Z_0^1$  and  $Y = Z_{\ell_k}^{k_s}$ . Also,  $Z_{\ell_k}^k = Z_0^{k+1}$  for  $1 \leq k < k_s$ .

We have

$$E(C_r^k) = \{v_x v_{x+1} : i_{r-1} \leq x < j_{r-1} \text{ and } f_{r-1}((x)^+) = \text{green}\},$$

so there is a natural Eulerian circuit on the edges (and vertices) of  $C_r^k$ .

$$\text{Let } x_1 \in V(C_r^k) \text{ be a vertex which has minimum degree in } Z^{k,r-1}[V(C_r^k)]. \quad (4.9)$$

We take swap sequence between  $Z_{r-1}^k$  and  $Z_r^k$  which is produced by Algorithm 2.2 as described by Lemma 2.5. The symmetric difference of  $Z_{r-1}^k$  and  $Z_r^k$  is  $E(C_r^k)$ .



Let  $q = |E(C_r^k)|$ , which is even. To apply Lemma 2.5, we label the vertices of  $C_r^k$  by  $x_1, x_2, \dots, x_{q-1}, x_q = x_1$ . The natural Eulerian circuit on  $C_r^k$  which starts on a non-edge  $x_1 x_2 \notin E_{r-1}^k$  uniquely determines the Eulerian trail  $x_1, x_2, \dots, x_q$ .

Recall Lemma 2.6 and Equation (2.1), which in our setting translates to

$$R = (Z \triangle Z_{r-1}^k) \setminus E(C_r^k) \quad (4.10)$$

We use the  $T$ -operator in the previous subsection to decompose  $W_k$  into alternating elementary circuits  $C_r^k$ , a process which is tracked by the pair  $(\pi_r, f_r)$ , an Eulerian trail on  $W_k$  (Lemma 4.9) and a red-green coloring of  $E(W_k)$ . Let  $w$  be given by Theorem 4.8 for  $\pi_{r-1}$ . Unfortunately,  $R$  may intersect  $W_i$  for some  $i \neq k$ , so it is favorable to work on a graph  $Z'$  which is close to  $Z$ :

$$Z' = Z \triangle R. \quad (4.11)$$

Although generally  $Z'$  is not a realization of  $\mathbf{d}$ , it is well-behaved with respect to alternating circuits other than  $W_k$ . From the definitions and Lemma 2.6(g) we can read off that

$$\begin{aligned} Z' \triangle X, Z' \triangle Y &\subseteq E(C_r^k) \subseteq X \triangle Y, \\ E(Z') \cap E(W_i) &= E(X) \cap E(W_i) \text{ for } i > k, \\ E(Z') \cap E(W_i) &= E(Y) \cap E(W_i) \text{ for } i < k, \\ E(Z') \cap E(C_j^k) &= E(X) \cap E(C_j^k) \text{ for } j > r, \\ E(Z') \cap E(C_j^k) &= E(Y) \cap E(C_j^k) \text{ for } j < r. \end{aligned} \quad (4.12)$$

If  $Z$  is a **milestone**, ie.,  $Z = Z_r^k$  for some  $k, r$ , then  $R = \emptyset$  and  $E(Z') \cap E(C_r^k) = E(Y) \cap E(C_r^k)$  holds too.

Let  $s(X, Y, Z) \in S_{\nabla \cap Z', \nabla \setminus Z'}$  be the matching we get by modifying the original  $s \in S_{X, Y}$  in such a way that on  $W_k$  it induces the Eulerian-trail  $\pi_{r-1}$ . However, this trail does not alternate at a constant number of positions in  $W_k$ , so we delete any matches between two non-edges or two edges of  $Z$ .

Thus knowing only  $Z'$ ,  $\nabla$ , and  $s^*$ , generally we cannot reassemble  $\pi_{r-1}$  nor  $W_k$ . However:

**Lemma 4.13.** *If  $Z$  is an intermediate realization produced on the swap-sequence between  $Z_{r-1}^k$  and  $Z_r^k$ , then  $Z' \triangle Z_0^k \subseteq W_k$ . Moreover, the Eulerian trail defined by  $\pi_{r-1}$  on  $W_k$  alternates on  $Z'$  with the exception of at most 8 pairs of chords in the unconstrained case (6 in the bipartite case).*

*Proof.* From Lemma 2.6(g) it follows that the natural Eulerian trail on  $C_r^k$  alternates on  $Z'$ , except at the two-two endpoints of the at most 3 walks (2 in the bipartite case) formed by  $Z' \triangle Z_{r-1}^k = Z' \triangle Z_0^k$ . Beyond these at most 6 (at most 4) sites of non-alternation, two extra irregularities are created at the points where the trail defined by  $\pi_{r-1}$  enters and leaves  $C_r^k$ .  $\square$

Since  $W_k$  is split into at most 8 walks (6 in the bipartite case), the parts can be reassembled in a constant number of ways. Let  $\sigma$  be the list of assembly instructions, in which we also indicate in which direction  $\pi_{r-1}$  starts from  $v_0$ . Trivially,  $\sigma$  can be encoded in a constant number of bits. Let us define a list of additional parameters associated to some quadruplet where  $Z \in \mathcal{R}(X, Y, s)$ :

$$B(X, Y, Z, s) := (x_1, \sigma, R, w). \quad (4.13)$$

#### 4.5. Reconstructing the swap sequence

Let the auxiliary structure  $\widehat{M}$  be defined by the equation  $\widehat{M} = A_X + A_Y - A_Z$ , where  $A_X, A_Y, A_Z$  are the adjacency matrices of  $X, Y, Z$ , respectively. (In the adjacency matrices in the columns the vertices are enumerated from left to right, while in the rows from top to bottom. Therefore the matrix is symmetric with identically zero main diagonal.) In this subsection we focus on the role of  $\widehat{M}$  in the reconstruction process. We will discuss further its properties in Section 6.

**Lemma 4.14.** *There is a function  $\Psi$  and a parameter set  $\mathbb{B}$  such that  $|\mathbb{B}| \leq \mathcal{O}(n^8)$  (and  $|\mathbb{B}| \leq \mathcal{O}(n^6)$  in the bipartite case), and for each  $(X, Y, s) \in \mathfrak{X}(Z, \widehat{M})$ , there exists a  $B \in \mathbb{B}$  for which*

$$\Psi(Z, \widehat{M}, s^*, B) = (X, Y, s),$$

where  $s^* = s(X, Y, Z)$ .

*Proof.* The heavy lifting in the proof is reconstructing the  $s \in S_{X,Y}$  for which  $s^* = s(X, Y, Z)$  holds. The graph  $X \triangle Y$  is determined by  $Z$  and  $\widehat{M}$  (but these data alone do not separate the  $X$ -edges and  $Y$ -edges). Let  $B$  be the set of additional parameters we defined in Equation (4.13). Clearly,  $0 \leq w \leq n^2$ , because the generalized  $T$ -operator decreases the number of green positions in each iteration. Take a look at Lemma 2.6: given  $x_1$   $R$  has only  $n^5$  possible values ( $n^3$  in the bipartite case). Since  $\sigma$  has a description of constant size, we have  $|\mathbb{B}| \leq \mathcal{O}(n^8)$  (and  $\mathcal{O}(n^6)$  in the bipartite case).

Using  $B$ , construct  $Z' = Z \triangle R$ . On  $X \triangle Y$ ,  $Z'$  and  $s^*$  determine the alternating cycles  $W_i$  for  $i \neq k$ . Though  $W_k$  may be cut into at most 8 components (see Lemma 4.13),  $\sigma$  allows us to reassemble  $W_k$  and  $\pi_{r-1}$ . Using  $w$ , the last coordinate of  $B$ , it is possible to restore the original Eulerian trail on  $W_k$ , namely  $\pi_0$  (see Theorem 4.8). Consequently, the elementary cycle decomposition  $\cup_{y=1}^{\ell_s} C_y^k$  can be determined (along with the value of  $r$ , too).

Recall Equation (4.12). It determines which edges of  $C_j^i$  belong to  $X$  (and  $Y$ ), except when  $i = k$  and  $j = r$ . On  $E(C_r^k)$ , we know the identity of  $x_1$  and  $x_2$ , and that  $x_1 x_2$  is a non-edge in  $X$ . Because  $C_r^k$  alternates between edges and non-edges in  $X$ , the Eulerian-trail  $\pi_0$  on  $W_k$  determines which edges of  $C_r^k$  belong to  $X$ , and which belong to  $Y$ .  $\square$

## 5. Directed degree sequences

Now it is time to extend our management for directed degree sequences. This short description goes more or less in parallel with [10].

Let  $\vec{G}$  be a simple directed graph (parallel edges and loops are forbidden, but oppositely directed edges between two vertices are allowed) with vertex set  $X(\vec{G}) = \{x_1, x_2, \dots, x_n\}$  and edge set  $E(\vec{G})$ . For every vertex  $x_i \in X$  we associate two numbers: the *in-degree* and the *out-degree* of  $x_i$ . These numbers form the directed degree bi-sequence  $\vec{\mathbf{d}}$ .

We introduce the following *bipartite representation* of  $\vec{G}$ : let  $B(\vec{G}) = (U, V; E)$  be a bipartite graph where each class consists of one copy of every vertex from  $X(\vec{G})$ . The edges adjacent to a vertex  $u_x$  in class  $U$  represent the out-edges from  $x$ , while the edges adjacent to a vertex  $v_x$  in class  $V$  represent the in-edges to  $x$  (so a directed edge  $xy$  corresponds the edge  $u_x v_y$ ). If a vertex has zero in- (respectively out-) degree in  $\vec{G}$ , then we delete the corresponding vertex from  $B(\vec{G})$ . (This representation was used by Gale [11], but one can find it already in [27].) The directed degree bi-sequence  $\vec{\mathbf{d}}$  gives rise to a bipartite degree sequence  $\vec{\mathbf{D}}$ .

Since there are no loops in our directed graph, there cannot be any  $(u_x, v_x)$  edge in its bipartite representation — these vertex pairs are **non-chords**. It is easy to see that these forbidden edges form a forbidden (partial) matching  $\mathcal{F}$  in the bipartite graph  $B(\vec{G})$ , or in more general terms, in  $B(\vec{\mathbf{D}})$ , and we call this a **restricted** bipartite degree sequence.

**Definition 5.1.** For restricted bipartite degree sequences, the set of chords is the vertex pairs of the form  $u_x v_y$  where  $x \neq y$ .

We consider all bipartite realizations  $\mathbb{G}(\vec{\mathbf{D}})$  which avoid the non-chords from  $\mathcal{F}$ . Now it is easy to see that the bipartite graphs in  $\mathbb{G}(\vec{\mathbf{D}})$  are in one-to-one correspondence with the possible realizations of the directed degree bi-sequence.

Consider now two oppositely oriented triangles,  $\vec{C}_3$  and  $\overleftarrow{C}_3$ . Consider the bipartite representations  $B(\vec{C}_3)$  and  $B(\overleftarrow{C}_3)$ , and take their symmetric difference  $\nabla$ . It contains exactly one alternating cycle (the edges come alternately from  $B(\vec{C}_3)$  and  $B(\overleftarrow{C}_3)$ ), s.t. each vertex pair of distance 3 along the cycle in  $\nabla$  forms a non-chord. In this alternating cycle no “classical” swap can be performed. To address this issue, we introduce a new swap operation: we exchange all edges coming from  $B(\vec{C}_3)$  with all edges coming from  $B(\overleftarrow{C}_3)$  in one operation.

In general, a **triple-swap** is defined as follows: take a length-6 alternating cycle  $C$  in  $\nabla$ , and if all three vertex pairs of distance 3 in  $C$  form non-chords, we exchange all edges of  $C$  to non-edges and vica-versa. The swaps and the restricted triple-swaps together are called the  **$\mathcal{F}$ -swaps**. It is a well known fact ([5], [7]) that the set  $\mathbb{G}(B(\vec{\mathbf{D}}))$  of all realizations is irreducible under  **$\mathcal{F}$ -swaps**.

The example of  $\vec{C}_3$  and  $\overleftarrow{C}_3$  demonstrates why the triple-swap operation is necessary. However, as long as some steps of the Markov-chain require choosing 6

vertices, it seems wasteful to not perform the triple-swap simply because some of the the vertex pairs of distance 3 are chords.

In this paper, we relax the restrictions on triple-swaps: given a length-6 alternating cycle  $C$  in  $\nabla$ , a triple swap is valid if and only if at least one of the three vertex pairs of distance 3 in  $C$  is a non-chord. This relaxation allows us to shave off a factor of  $n^6$  from the mixing time of the Markov chain. To see this, compare the proofs of Theorem 7.3 and Theorem 7.7.

The inner loop of Algorithm 2.2 has to be modified, because the conclusion of Lemma 2.4 does not necessarily hold in the directed case. Instead of calling DOUBLE STEP in Algorithm 2.2 when  $x_1x_{2t}$  is a non-chord, the procedure should call TRIPLE-SWAP of Algorithm 5.1. If  $Z_q$  gets its value from TRIPLE-SWAP, then

---

**Algorithm 5.1** The (relaxed) triple-swap operation

---

**Ensure:**  $x_1x_{2t}$  is a non-chord

**function** TRIPLE-SWAP( $G, x_1, [x_{2t-2}, x_{2t-1}, x_{2t}, x_{2t+1}, x_{2t+2}]$ )

**return**  $G + (x_1x_{2t-2} - x_{2t-2}x_{2t-1} + x_{2t-1}x_{2t} - x_{2t}x_{2t+1} + x_{2t+1}x_{2t+2} - x_{2t+2}x_1)$

**end function**

---

Lemma 2.6(d) will apply to  $R$  (see Equation (2.1)). Because of this, the statements of Lemma 4.13 and 4.14 about the bipartite case apply to the directed case as well.

We are ready to define our swap Markov chain on  $(\mathbb{G}(\vec{\mathbf{D}}), P)$  for the restricted bipartite degree sequence  $\vec{\mathbf{D}}$ .

The transition (probability) matrix  $P$  of the Markov chain is defined as follows: let the current realization be  $G$ . Then

- (a) with probability  $1/2$  we uniformly choose two-two vertices  $u, u'; v, v'$  from classes  $U$  and  $V$ , respectively. There are two matchings  $\{\{uv, u'v'\}, \{uv', u'v\}\}$  between the two-two vertices, let  $F$  be chosen randomly, and let  $F'$  be the other matching. If both  $F$  and  $F'$  consist of chords only and  $F \subseteq E(G)$  and  $F' \cap E(G) = \emptyset$ , then perform the swap (so  $E(G') = E(G) \cup F' \setminus F$ ), otherwise  $G' = G$ .
- (b) With probability  $1/2$  we choose three-three vertices from  $U$  and  $V$ . Let  $F$  and  $F'$  be a uniformly randomly selected pair of disjoint perfect matchings between the three-three vertices. If both  $F$  and  $F'$  consist of chords only, and the remaining matching between the three-three vertices contains a non-chord, and  $F \subseteq E(G)$  and  $F' \cap E(G) = \emptyset$ , then perform the triple-swap (so  $E(G') = E(G) \cup F' \setminus F$ ), otherwise  $G' = G$ .

The (triple-)swap moving from  $G$  to  $G'$  is unique, therefore the probability of this transformation (the *jumping probability* from  $G$  to  $G' \neq G$ ) is:

$$\text{Prob}(G \rightarrow_{(a)} G') := P(G, G') = \frac{1}{4} \cdot \frac{1}{\binom{|U|}{2} \binom{|V|}{2}}, \quad (5.1)$$

and

$$\text{Prob}(G \rightarrow_{(b)} G') := P(G, G') = \frac{1}{24} \cdot \frac{1}{\binom{|U|}{3} \binom{|V|}{3}}. \quad (5.2)$$

The probability of transforming  $G$  to  $G'$  (or vice versa) is time-independent and symmetric. Therefore,  $P$  is a symmetric matrix, where the entries in the main diagonal are non-zero, but (possibly) distinct values. Again,  $P(G, G) \geq \frac{1}{2}$ , because if  $(F, F')$  corresponds to a feasible (triple-)swap, then  $(F', F)$  does not. Therefore the chain is aperiodic and the eigenvalues of its transition matrix are non-negative.

However it is important to recognize that in papers [14] and [16] there was a slightly different Markov chain studied, where it is assumed that the degree sequences under study are irreducible using swaps only. One notable example is the regular directed degree sequence. Papers [2] and [24] provide a full characterization of directed degree sequences with this property.

## 6. The auxiliary matrix $\widehat{M}$

The auxiliary matrix  $\widehat{M} = A_X + A_Y - A_Z$  is a linear combination of three adjacency matrices. The row and columns sums are equal to the corresponding degrees prescribed by  $\mathbf{d}$ . If  $Z = Z_r^k$ , then  $Z_r^k \triangle X \subseteq X \triangle Y$  implies that  $\widehat{M}$  is a 0–1 matrix. If  $Z$  is an intermediate realization,  $\widehat{M}$  is still a 0–1 matrix except on the entries associated to edges in  $R$ , since  $(Z \triangle R) \triangle X \subseteq X \triangle Y$ . These +2 and –1 entries will be called **bad entries**, and the chords to which they correspond to are called type-(2) and type-(–1) chords, respectively.

**Lemma 6.1.** *If  $R$  falls under case (c) or (d) of Lemma 2.6, then  $R$  contains at most two type-(2) and at most one type-(–1) chords.*

*Proof.* Lemma 2.6(c) or (d) claims that  $R$  has at most three elements. Of these,  $x_1 x_{\text{startChord}}$  and  $x_1 x_{\text{endChord}}$  are edges in  $X$ , so the entries associated to them in  $\widehat{M}$  are two-two +2 or +1 entries. In case (c), if  $R$  contains the third chord,  $x_1 x_{2t}$ , and it is an edge in  $X$ , then we must have  $\text{endChord} = 2t$ , so  $R$  actually does not contain  $x_1 x_{2t}$ . Thus  $x_1 x_{2t} \in R \implies x_1 x_{2t} \notin E(X)$ , so the entries associated to  $x_1 x_{2t}$  in  $\widehat{M}$  are –1’s or 0’s. Case (d) is similar to case (c).  $\square$

A **switch** on a symmetric matrix is defined as follows. Suppose  $M \in \mathbb{Z}^{[k] \times [k]}$ . For any  $x, y \in [k]$  we define the one-edge graph  $G^{x,y} = ([k]; \{xy\})$  with the adjacency matrix  $A_{xy}$ . Clearly,  $A_{xy}$  is a symmetric matrix with two 1’s. Let  $(x, y; z, w)$  be a list of four **pairwise distinct** elements of  $[k]$ . Switching along these four vertices produces the symmetric matrix

$$M' = M + A_{xz} - A_{zy} + A_{yw} - A_{wx}. \quad (6.1)$$

Clearly, the row and column sums of  $M'$  are identical to that of  $M$ . Notice, that a swap in  $Z$  translates into a switch on  $\widehat{M}$ .

Notice, that for bipartite degree sequences, the “top-right” submatrix of this  $\widehat{M}$  is equal to the auxiliary matrix used in [25] (the bipartite adjacency matrix).

**Lemma 6.2.** *Suppose  $M \in \mathbb{Z}^{[k] \times [k]}$  is a symmetric matrix with 0’s in the diagonal, such that each row and column sum is in  $[1, k-2]$ . Also, suppose that the row sum of the first row is minimal. If the entries of  $M$  are 0 and 1, except for at most 2 pairs of entries of +2 in the first row and in the first column, and at most two, symmetric -1’s anywhere in the matrix, then there exist at most 2 switches that transform  $M$  into a 0-1 matrix except for at most two symmetric -1.*

*Proof.* Suppose  $M_{1,j} = 2$ . We must have  $j \neq 1$ , which means that the maximum of an entry in the rest of the column of  $j$  is 1. Because the column sum is at most  $k-2$  and there is at most one -1 in the column, there exist  $i, i' \in [k] \setminus \{1, j\}$  such that  $i \neq i'$  and  $M_{i,j}, M_{i',j} \in \{-1, 0\}$ . We have two cases.

1. There  $\exists l \in [k] \setminus \{1, i\}$  such that  $M_{i,l} > M_{1,l}$ : since  $1 \neq i, l$ , we assumed that  $M_{i,l} < 2$ , therefore  $M_{1,l} \in \{0, -1\}$ . Switch along  $(1, i; l, j)$  in  $M$ . The operation decreases  $M_{1,j}$  to 1. If  $M_{i,l} = 0$ , then  $M_{1,l} = -1$ , so when the switch creates a symmetric pair of -1’s, it also eliminates another pair. The matrix resulting from the switch operation satisfies the assumptions of this lemma and contains two fewer +2 entries.
2. If  $\forall l \in [k] \setminus \{1, i\}$   $M_{i,l} \leq M_{1,l}$ : since the row sum of the first row is minimal, we have

$$\sum_{l=1}^k M_{1,l} \leq \sum_{l=1}^k M_{i,l}$$

$$0 \leq M_{i,1} + M_{i,j} - M_{1,i} - M_{1,j} = M_{i,j} - 2,$$

which implies that  $i = 1$ , a contradiction.

By recursion a second pair of +2 can also be eliminated. □

## 7. An application of the unified method

In this Section we harvest some fruits of our unified machinery proving a rather general result for all typical degree sequence types.

In 1990 Jerrum and Sinclair published a very influential paper ([20]) about fast uniform generation of regular graphs and about realizations of degree sequences where no degree exceeds  $\sqrt{n/2}$ . To achieve this goal they applied the Markov chain they have developed in [19]. Informally it is known as **JS chain**, and it is sampling the perfect and near-perfect 1-factors on the corresponding Tutte gadget. The fast mixing nature of the JS chain depends on the ratio of the number of perfect and the number near-perfect 1-factors. As they proved it is applicable if and only if the degree sequence  $\mathbf{d}$  belongs to a **P-stable** class.

Roughly speaking, a degree sequence is  $P$ -stable, if small perturbations to the sequence (like considering a graphic degree sequence of at most distance two in  $\ell_1$ -norm) do not change the numbers of all possible realizations greatly. More precisely we use the following definition:

**Definition 7.1.** Let  $\mathcal{D}$  be a set of (unconstrained, bipartite or directed) 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  $\mathbf{d} \in \mathcal{D}$  on  $n$  vertices we have

$$\left| \mathbb{G}(\mathbf{d}) \cup \left( \bigcup_{x,y \in [n], x \neq y} \mathbb{G}(\mathbf{d} + \mathbf{1}_x + \mathbf{1}_y) \right) \right| \leq p(n) \cdot |\mathbb{G}(\mathbf{d})|,$$

where  $\mathbf{1}_x$  is the  $x^{\text{th}}$  unit vector.

Without proof we state, that the notion of  $P$ -stability does not change even if we require  $|\{G \mid G \in \mathbb{G}(\mathbf{d}'), \mathbf{d}' \in \mathbb{N}^n, \ell_1(\mathbf{d}, \mathbf{d}') \leq 2\}| \leq p(n) \cdot |\mathbb{G}(\mathbf{d})|$ . In a forthcoming paper we will study  $P$ -stability in details, but here we confine ourselves to the following observation, some statements and their direct consequences.

Careful examination of the known results about rapidly mixing swap Markov chains revealed the fact that all known “good” degree sequence classes (for UC degree bipartite or directed degree sequences) are  $P$ -stable. It raises the conjecture that the swap Markov chains on  $P$ -stable degree classes are rapidly mixing. We resolve this conjecture affirmatively in this section.

Let the set of auxiliary structures be

$$\mathcal{M} = \left\{ \widehat{M} : \exists X, Y, Z \in \mathbb{G}(\mathbf{d}) \text{ s.t. } \widehat{M} = A_X + A_Y - A_Z \right\}.$$

For a fixed  $B \in \mathbb{B}$ , let the set of compatible auxiliary structures be

$$\begin{aligned} \mathcal{M}_B = \left\{ \widehat{M} : \exists X, Y, Z \in \mathbb{G}(\mathbf{d}), s \in S_{X,Y} \right. \\ \left. \text{s.t. } \widehat{M} = A_X + A_Y - A_Z, B = B(X, Y, Z, s) \right\}. \end{aligned}$$

At this point, the proofs for unconstrained, bipartite, and directed degree sequences slightly diverge. The most general of these is the case of unconstrained degree sequences. First, we discuss this case. Having understood the argument, it is relatively simple to fit it to the cases of the bipartite and directed degree sequence cases. Moreover, the tools required for proving our results on the latter two classes have already been published in [25], so their proofs will be less verbose than the next section on unconstrained degree sequences.

### 7.1. Unconstrained degree sequences

First, let us bound the number of auxiliary structures compatible with a given parameter set.

**Lemma 7.2.** *If the stability of an unconstrained degree sequence  $\mathbf{d}$  is bounded by the polynomial  $p(n)$ , then*

$$|\mathcal{M}_B| \leq n^6 \cdot p(n)$$

*holds for any  $B \in \mathcal{B}$ .*

*Proof.* Equation (3.8) defines  $B = (x_1, \sigma, R, w)$ . Recall, that  $(Z \triangle R) \triangle X \subseteq X \triangle Y$ , so the bad entries (+2 and -1 values) in  $\widehat{M}$  correspond to positions assigned to chords in  $R$ . Let  $M$  be the symmetric submatrix of  $\widehat{M}$  induced by the vertices of  $C_r^k$  as rows and columns. We have two cases.

*Case 1:  $\mathbf{R}$  falls under case (b), (c), or (d) of Lemma 2.6*

All of the non 0-1 entries of  $\widehat{M}$  are contained in  $M$ , in the rows and columns associated to  $x_1$ . Lemma 6.1 and Assumption (4.9) implies that we can use Lemma 6.2 to remove the +2's from  $\widehat{M}$  with at most two switches. For each switch, the type-(2) chord determines two vertices of the switch, thus there are  $n^4$  ways to choose the at most two switches that eliminate the +2 entries.

Let  $\widehat{M}'$  be the matrix we get after applying the switches defined by Lemma 6.2. Either  $\widehat{M}'$  is an adjacency matrix of a realization of  $\mathbf{d}$ , or  $\widehat{M}'$  contains -1 entries at positions associated to the chord  $xy$ . In the former case  $\widehat{M}' \in \mathbb{G}(\mathbf{d})$ , and in the latter  $\widehat{M}' + A_{xy} + A_{yx} \in \mathbb{G}(\mathbf{d} + \delta_x + \delta_y)$ .

*Case 2:  $\mathbf{R}$  falls under case (e) or (f) of Lemma 2.6*

Let  $f \in R$  be the unique edge which is not incident to  $x_1$ . The auxiliary structure belonging to the intermediate realization before or after  $Z$  on the swap sequence is one switch away from  $\widehat{M}$ , moreover the switch touches  $f$ . As in the previous case,  $f$  determines two vertices of this switch, so there are  $n^2$  possibilities to choose the other two vertices. After performing the switch, Case 1 applies.  $\square$

We are ready to prove one of the main results of this paper.

**Theorem 7.3.** *The swap Markov chain is rapidly mixing on  $P$ -stable unconstrained degree sequence classes.*

*Proof.* Apply the simplified Sinclair's method. It is sufficient to show that Equation (3.9) holds. For any  $Z \in \mathbb{G}(\mathbf{d})$  we have to estimate from above the value of

$$\sum_{X, Y \in V(\mathbb{G})} \frac{|\{s \in S_{X, Y} : Z \in \mathcal{R}(X, Y, s)\}|}{|S_{X, Y}|}. \quad (7.1)$$



The set in the nominator can be rewritten as follows:

$$\left\{ s \in S_{X,Y} : \exists B \in \mathbb{B}, \widehat{M} \in \mathcal{M}, s^* \in S_{\nabla \cap Z', \nabla \setminus Z'} \text{ s.t. } \Phi(Z, \widehat{M}, s^*, B) = (X, Y, s) \right\}$$

where  $\nabla = X \triangle Y$ . Observe, that  $|S_{X,Y}|$  is already determined by  $\nabla$ , which in turn is determined by  $Z$  and  $\widehat{M}$ . Let  $t_\nabla := |S_{X,Y}|$ . Furthermore,  $Z'$  is determined by  $B$  and  $Z$  (see Equations (4.11) and (4.13)). Lastly, observe that

$$|S_{\nabla \cap Z', \nabla \setminus Z'}| \leq (\max \mathbf{d})^8 \cdot t_\nabla$$

because half of the  $\ell_1$  distance of the degree sequences of  $Z'$  and  $\nabla \setminus Z'$  is at most 8 (follows from Lemma 4.13). Let

$$\mathbb{B}_{\widehat{M}} = \left\{ B(X, Y, Z, s) : \widehat{M} = A_X + A_Y - A_Z, s \in S_{X,Y} \right\}.$$

We can continue writing (7.1) as follows:

$$\begin{aligned} &= \sum_{\widehat{M} \in \mathcal{M}} \frac{\left| \{ \Phi(Z, \widehat{M}, s^*, B) : \exists B \in \mathbb{B}, s^* \in S_{\nabla \cap Z', \nabla \setminus Z'} \} \right|}{t_\nabla} \leq \\ &\leq \sum_{\widehat{M} \in \mathcal{M}} \frac{|\mathbb{B}_{\widehat{M}}| \cdot |S_{\nabla \cap Z', \nabla \setminus Z'}|}{t_\nabla} \leq (\max \mathbf{d})^6 \cdot \sum_{\widehat{M} \in \mathcal{M}} |\mathbb{B}_{\widehat{M}}|. \end{aligned}$$

Continue by applying Lemma 7.2.

$$\begin{aligned} (\max \mathbf{d})^8 \cdot \sum_{\widehat{M} \in \mathcal{M}} |\mathbb{B}_{\widehat{M}}| &\leq (\max \mathbf{d})^8 \cdot \sum_{B \in \mathbb{B}} |\widehat{M}_B| \leq \\ &\leq (\max \mathbf{d})^8 \cdot |\mathbb{B}| \cdot n^6 \cdot p(n) \cdot |\mathbb{G}(\mathbf{d})| \leq \mathcal{O}(n^{22}) \cdot p(n) \cdot |\mathbb{G}(\mathbf{d})| \end{aligned}$$

In the last step, we used Lemma 4.14.  $\square$

## 7.2. Bipartite degree sequences

Let  $\mathbf{D}$  denote a bipartite degree sequence on  $N = n + m$  vertices. Recall that an alternating elementary circuit on a bipartite graph is a cycle. Also, for any  $X, Y, Z \in \mathbb{G}(\mathbf{D})$ , the auxiliary structure  $\widehat{M} = A_X + A_Y - A_Z$  is determined by the submatrix spanned by  $U \times V \subset (U \uplus V)^2$ . This is the “top-right” submatrix, often called the bipartite adjacency matrix. The “top-left” and the “bottom-right” submatrices are zero.

**Lemma 7.4.** *If the stability of a bipartite degree sequence  $\mathbf{D}$  is bounded by the polynomial  $p(N)$ , then*

$$|\mathcal{M}_B| \leq (nm)^2 \cdot p(N)$$

*holds for any  $B \in \mathcal{B}$ .*

*Proof.* The proof is simpler and slightly different than that of Lemma 7.2. DOUBLE STEP is never called in the bipartite case (Lemma 2.5), so either  $R$  is empty or Lemma 6.1 applies to it. Hence, there are  $(nm)^2$  possibilities to choose the two-two other vertices of the switches that eliminate the type-(2) bad chords.

Secondly, we have to make sure that the switches produced by Lemma 6.2 respect the bipartition. As before, let  $M$  be the submatrix of  $\widehat{M}$  induced by the vertices of  $C_r^k$ . Let  $H = K_{U(C_r^k)} \uplus K_{V(C_r^k)}$  be the disjoint union of the two cliques within the classes. Instead of applying Lemma 6.2 on  $M$ , apply it on  $M + A_H$ . Each row and column sum increased by the same number, therefore assumptions of the lemma are still satisfied. Any swap which eliminates a +2 from this matrix which is valid in the unconstrained sense also respects the bipartition.  $\square$

**Theorem 7.5.** *The swap Markov chain is rapidly mixing on  $P$ -stable bipartite degree sequence classes.*

*Proof.* Instead of Lemma 7.2 we use Lemma 7.4. The bound on the size of  $\mathbb{B}$  is  $\mathcal{O}(n^6)$  according to Lemma 4.14. Furthermore, the constant 8 improves to 6 in Lemma 4.13, thus half of the  $\ell_1$  distance of the degree sequences of  $Z'$  and  $\nabla \setminus Z'$  is at most 6.

Other than the mentioned differences, the proof is identical to that of Theorem 7.3:

$$\begin{aligned} \sum_{X,Y \in V(\mathbb{G})} \frac{|\{s \in S_{X,Y} : Z \in \mathcal{Y}(X,Y,s)\}|}{|S_{X,Y}|} &\leq \\ &\leq (\max \mathbf{D})^6 \cdot |\mathbb{B}| \cdot (nm)^2 \cdot p(N) \cdot |\mathbb{G}(\mathbf{D})| \leq \mathcal{O}(N^{16}) \cdot p(N) \cdot |\mathbb{G}(\mathbf{D})| \end{aligned}$$

$\square$

### 7.3. Directed degree sequences

Recall from Section 5 that instead of directly manipulating directed graphs, we work on their bipartite representations. Formally, the degree sequence of the directed graph is identical to that of its bipartite representation. Let  $\vec{\mathbf{D}}$  denote a directed degree sequence on  $N = n + n$  vertices (so the bipartite representation has  $N$  vertices).

**Lemma 7.6.** *If the stability of a directed degree sequence  $\vec{\mathbf{D}}$  is bounded by the polynomial  $p(n)$ , then*

$$|\mathcal{M}_B| \leq n^4 \cdot p(n)$$

*holds for any  $B \in \mathcal{B}$ .*

*Proof.* The proof of Lemma 7.4 applies to the bipartite representation, but we have to check that applying Lemma 6.2 on  $M + A_H$  produces switches that avoid the non-chords. Indeed, this is the case, because the non-chords of the form  $u_x v_x$  correspond to the main diagonal in  $M$ , which the switches chosen by the lemma avoid.  $\square$

**Theorem 7.7.** *The swap Markov chain is rapidly mixing on  $P$ -stable directed degree sequence classes.*

*Proof.* By Lemma 4.14, the bound on the size of  $\mathbb{B}$  is  $\mathcal{O}(n^6)$ , as in the proof of Theorem 7.3. Similarly, half of the  $\ell_1$  distance of the degree sequences of  $Z'$  and  $\nabla \setminus Z'$  is at most 8 (Lemma 4.13).

Other than the mentioned differences, the proof is identical to that of Theorem 7.5:

$$\begin{aligned} \sum_{X,Y \in V(\mathbb{G})} \frac{|\{s \in S_{X,Y} : Z \in \mathcal{R}(X,Y,s)\}|}{|S_{X,Y}|} &\leq \\ &\leq (\max \vec{\mathbf{D}})^6 \cdot |\mathbb{B}| \cdot n^4 \cdot p(n) \cdot |\mathbb{G}(\vec{\mathbf{D}})| \leq \mathcal{O}(n^{16}) \cdot p(n) \cdot |\mathbb{G}(\vec{\mathbf{D}})|. \end{aligned}$$

□

## 8. $P$ -stable degree sequence classes

In the proof of almost every previous result on rapid mixing of the swap Markov chain, it turns out there is a short hidden proof that the degree sequences under study are  $P$ -stable. The unified proof contains most of the technical difficulty of proving rapid mixing of the swap Markov chain.

There have already been successful attempts at unifying some of the proofs, most notably [1], which studies the notion of strong stability:

**Definition 8.1** (adapted from [1]). Let  $\mathcal{D}$  be a set of degree sequences. Let

$$\mathbb{G}'(\mathbf{d}) = \bigcup_{x,y \in [n]} \mathbb{G}(\mathbf{d} - \mathbf{1}_x - \mathbf{1}_y).$$

We say that  $\mathcal{D}$  is strongly stable if there exists a constant  $\ell$  such that for any  $\mathbf{d} \in \mathcal{D}$  and any  $G' \in \mathbb{G}'(\mathbf{d})$  there exists a set of chords  $E_{G'}$  of cardinality at most  $\ell$  such that  $G' \triangle E_{G'} \in \mathbb{G}(\mathbf{d})$ .

The statement without proof after Definition 7.1 hints at the fact that strongly stable degree sequence classes are also  $P$ -stable, because  $E_{G'}$  comes from a set of size at most  $\binom{n^2}{\ell}$ , which is a polynomial of  $n$ .

**Theorem 8.2** (Amanatidis and Kleer [1]). *The swap Markov chain is rapidly mixing on strongly stable unconstrained and bipartite degree sequence classes.*

In the following subsections of this section we discuss all known  $P$ -stable degree sequence regions. It is an intriguing problem to discover other  $P$ -stable regions.

### 8.1. Unconstrained degree sequences

For the sake of having more readable and compact formulas, let  $\Delta = \max \mathbf{d}$ ,  $\delta = \min \mathbf{d}$ , and  $m = \frac{1}{2} \sum_{v \in V} \mathbf{d}(v)$  be functions of  $\mathbf{d}$ .

Recently, Greenhill and Sfragara [16] published a breakthrough result on the rapid mixing of the swap Markov chain.

**Theorem 8.3** ([16]). *The swap Markov chain is rapidly mixing on the following family of unconstrained degree sequences:*

$$\mathcal{D}_{G-S} := \left\{ \mathbf{d} \in \mathbb{Z}^+ : \delta \geq 1, 3 \leq \max \mathbf{d} \leq \frac{1}{3} \sqrt{2m} \right\} \quad (8.1)$$

It turns out that the authors implicitly prove on page 10 of [16] that  $\mathcal{D}_{G-S}$  is a  $P$ -stable class. However, this implicit result is actually not new: Jerrum, McKay, and Sinclair extensively studied the notion of  $P$ -stability in their seminal work [21].

**Theorem 8.4** (Jerrum, McKay, Sinclair — Theorem 8.1 in [21]). *The family of unconstrained degree sequences*

$$\mathcal{D}_{JMS} := \left\{ \mathbf{d} \in \mathbb{N}^n : (\max \mathbf{d} - \min \mathbf{d} + 1)^2 \leq 4 \cdot \min \mathbf{d} \cdot (n - \max \mathbf{d} - 1) \right\}$$

*is  $P$ -stable.*

**Theorem 8.5** (Jerrum, McKay, Sinclair — Theorem 8.3 in [21]). *The family of unconstrained degree sequences*

$$\begin{aligned} \mathcal{D}_{JMS+} := \left\{ \mathbf{d} \in \mathbb{N}^n : (2m - n\delta)(n\Delta - 2m) \leq \right. \\ \left. \leq (\Delta - \delta)((2m - n\delta)(n - \Delta - 1) + (n\Delta - 2m)\delta) \right\}, \end{aligned}$$

*is  $P$ -stable.*

Theorem 7.3 implies that the swap Markov chain is rapidly mixing on elements of  $\mathcal{D}_{JMS}$  and  $\mathcal{D}_{JMS+}$ . Moreover, it is easy to see that  $\mathcal{D}_{G-S} \subset \mathcal{D}_{JMS+}$ . However, the proofs of Theorems 8.4 and 8.5 actually prove a bit more than just  $P$ -stability.

In [21] it is also shown that  $\mathcal{D}_{JMS}$  and  $\mathcal{D}_{JMS+}$  are **strongly stable** regions with  $\ell \leq 10$ , so Theorem 8.2 already applies to them.

The following corollary is a consequence of the fact that the degrees in an Erdős-Rényi random graph are tightly concentrated around their expected value.

**Corollary 8.6.** *If  $G(n, p)$  is an Erdős-Rényi random graph of order  $n \geq 100$  with edge probability  $p$  then  $\Pr(\mathbf{d}(G(n, p)) \in \mathcal{D}_{JMS+}) \geq 1 - \frac{3}{n}$ .*

*Proof.* We may suppose that  $p \leq \frac{1}{2}$  by taking the complement of  $G$  if necessary. Let  $p = p(n)$ ,  $\varepsilon_1 = \sqrt{\frac{\log n}{(n-1)}}$  and  $m = \sum_{v \in V} \mathbf{d}(v) = \binom{n}{2}(p + \varepsilon_2)$ . By Hoeffding's inequality, we have

$$\begin{aligned} \Pr(\Delta(G) > (p + \varepsilon_1) \cdot (n-1)) &\leq \\ &\leq \sum_{v \in V(G)} \Pr(\mathbf{d}(v) < (p + \varepsilon_1) \cdot (n-1)) \leq n \cdot e^{-2\varepsilon_1^2(n-1)} \leq \frac{1}{n} \\ \Pr(\delta(G) < (p - \varepsilon_1) \cdot (n-1)) &\leq \frac{1}{n} \end{aligned}$$

If  $\Delta(G) \leq (p + \varepsilon_1) \cdot (n-1)$  and  $\delta(G) \geq (p - \varepsilon_1) \cdot (n-1)$ , then  $\mathbf{d}(G(n, p)) \in \mathcal{D}_{\text{JMS}+}$  holds if

$$(2m - n\delta)(n\Delta - 2m) \leq (\Delta - \delta)((2m - n\delta)(n - \Delta - 1) + (n\Delta - 2m)\delta) \quad (8.2)$$

Because increasing  $\Delta$  or decreasing  $\delta$  makes the inequality stricter, without loss of generality, we may substitute  $\Delta = (p + \varepsilon_1) \cdot (n-1)$  and  $\delta = (p - \varepsilon_1) \cdot (n-1)$  into inequality.

$$\begin{aligned} (2m - n\delta) &= \left(2\binom{n}{2}(p + \varepsilon_2) - n(p + \varepsilon_1)(n-1)\right) = n(n-1)(\varepsilon_1 + \varepsilon_2) \\ (n\Delta - 2m) &= n(n-1)(\varepsilon_1 + \varepsilon_2) \\ (2m - n\delta)(n - \Delta - 1) &= n(n-1)(\varepsilon_1 + \varepsilon_2) \cdot (1 - p - \varepsilon_1)(n-1) \\ (n\Delta - 2m)\delta &= n(n-1)(\varepsilon_1 + \varepsilon_2) \cdot (p - \varepsilon_1)(n-1) \end{aligned}$$

Therefore (8.2) becomes

$$\begin{aligned} (n(n-1)(\varepsilon_1 + \varepsilon_2))^2 &\leq 2\varepsilon_1(n-1) \cdot n(n-1)(\varepsilon_1 + \varepsilon_2) \cdot (1 - 2\varepsilon_1)(n-1) \\ n(\varepsilon_1 + \varepsilon_2) &\leq 2\varepsilon_1(1 - 2\varepsilon_1)(n-1) \\ n\varepsilon_2 &\leq (n-2)\varepsilon_1 - 4\varepsilon_1^2(n-1) \end{aligned}$$

If  $\varepsilon_2 \leq \frac{\sqrt{\log n}}{n}$ , then last inequality is satisfied, if

$$\begin{aligned} \sqrt{\log n} &\leq (n-2)\sqrt{\frac{\log n}{n-1}} - 4\log n \\ 4\sqrt{\log n} + 1 &\leq (n-2)\sqrt{\frac{1}{n-1}} \end{aligned}$$

Clearly, the right hand side grows orders of magnitude faster than the left hand side as  $n \rightarrow \infty$ , and the inequality already holds for  $n = 101$ . Lastly, it is very likely that  $\varepsilon_2 \leq \frac{\sqrt{\log n}}{n}$ , thus

$$\Pr\left(\frac{1}{2} \sum_{v \in V} d(v) > (p + \varepsilon_2) \binom{n}{2}\right) \leq e^{-2\varepsilon_2^2 \binom{n}{2}} \leq \frac{1}{n}.$$

All in all,  $\Pr(\mathbf{d}(G(n, p)) \notin \mathcal{D}_{\text{JMS}+}) \leq \frac{1}{n} + \frac{1}{n} + \frac{1}{n}$ .  $\square$

Similar result has already been shown about bipartite Erdős-Rényi graphs [9, 10], with the requirement that  $p$  is bounded away from 0 and 1 by at least  $4\sqrt{\frac{2 \log n}{n}}$ .

### 8.2. Unconstrained power-law bounded degree sequences

The swap Markov chain is not the only way to exactly sample the uniform distribution on the realizations of a degree sequence. Recently, Gao and Wormald presented the first “provably practical” sampler for power-law distribution-bounded degree sequence where  $\gamma$  is allowed to be less than 3; in fact they can go as low as 2.8811. For such degree sequences, they provide a *linear time* approximate sampler and a polynomial time exact sampler.

In degree distributions of empirical networks following a power-law, the parameter  $\gamma$  is usually between 2 and 3.

Gao and Wormald also enumerate the number of realizations of several heavy-tailed degree sequences [12]. In particular, they calculate the number of realizations of degree sequences that are

- (1) power-law **density-bounded** with  $\gamma > 2.5$ ,
- (2) power-law **distribution-bounded** with  $\gamma > 1 + \sqrt{3} \approx 2.732$ .

Their analysis of degree sequences obeying (1) shows that they are contained by  $\mathcal{D}_{\text{G-S}}$ . The formula in [12] enumerating the realizations of degree sequences obeying (2) directly implies  $P$ -stability, thus our Theorem 7.3 applies. Gao and Wormald also claim that degree sequences obeying a power-law distribution-bound with  $\gamma > 2$  are  $P$ -stable [13]. Therefore Theorem 7.3 provides a fully polynomial time almost uniform sampler:

**Theorem 8.7** (Follows from Theorem 7.3 and [13]). *The swap Markov chain is rapidly mixing on unconstrained degree sequences satisfying a power-law distribution-bound for any  $\gamma > 2$ .*

### 8.3. Bipartite degree sequences

Let  $\mathbf{D}$  be a bipartite degree sequence on  $U$  and  $V$  as color classes. We use the following shorthands in this sub-section:

$$\begin{aligned} \delta_U &= \min_{u \in U} \mathbf{D}(u), & \delta_V &= \min_{v \in V} \mathbf{D}(v), \\ \Delta_U &= \max_{u \in U} \mathbf{D}(u), & \Delta_V &= \max_{v \in V} \mathbf{D}(v), \end{aligned}$$

$$\text{and } m = \sum_{u \in U} \mathbf{D}(u) = \sum_{v \in V} \mathbf{D}(v).$$

**Theorem 8.8** (implicitly proved in Theorem 2 in [10]). *The set of bipartite degree sequences  $\mathbf{D}$  that satisfy*

$$2 \leq \Delta \leq \sqrt{\frac{1}{2}m}, \quad (8.3)$$

*is  $P$ -stable.*

Clearly, Theorems 8.3 and 8.8 are closely related, the difference in constants is caused by the different structural constraints only.

**Theorem 8.9** (implicitly proved in Theorem 3 in [10]). *The set of bipartite degree sequences  $\mathbf{D}$  that satisfy*

$$(\Delta_U - \delta_U - 1)(\Delta_V - \delta_V - 1) \leq \max\left(\delta_U(|U| - \Delta_V + 1), \delta_V(|V| - \Delta_U + 1)\right)$$

*is  $P$ -stable.*

The previous theorem is also present in **v3** (6 Oct 2018) of [1] as Corollary 18, but it does not cite paper [10], where the theorem was first published, even though the paper is present in [1] as reference 17.

**Theorem 8.10** (Corollary 19 in [1]). *The set of bipartite degree sequences that satisfy*

$$(\Delta_U - \delta_U) \cdot (\Delta_V - \delta_V) \leq 4 \cdot \min\left(\delta_U(|U| - \Delta_V), \delta_V(|V| - \Delta_U)\right) \quad (8.4)$$

*is  $P$ -stable.*

#### 8.4. Directed degree sequences

Let  $\vec{\mathbf{D}}$  be a directed degree sequence on  $X$  as vertices. Let  $\vec{\mathbf{D}}_{\text{out}}$  be the out-degree sequence and  $\vec{\mathbf{D}}_{\text{in}}$  be the in-degree sequence. We use the following abbreviations in this sub-section:

$$\begin{aligned} \delta_{\text{out}} &= \min_{x \in X} \vec{\mathbf{D}}_{\text{out}}(x), & \delta_{\text{in}} &= \min_{x \in X} \vec{\mathbf{D}}_{\text{in}}(x), \\ \Delta_{\text{out}} &= \max_{x \in X} \vec{\mathbf{D}}_{\text{out}}(x), & \Delta_{\text{in}} &= \max_{x \in X} \vec{\mathbf{D}}_{\text{in}}(x), \end{aligned}$$

and  $m = \sum_{x \in X} \vec{\mathbf{D}}_{\text{out}}(x) = \sum_{x \in X} \vec{\mathbf{D}}_{\text{in}}(x)$ .

**Theorem 8.11** (implicitly proved in [16]). *The set of bipartite degree sequences  $\vec{\mathbf{D}}$  that satisfy*

$$2 \leq \max(\Delta_{\text{out}}, \Delta_{\text{in}}) \leq \frac{1}{4}\sqrt{m}, \quad (8.5)$$

*is  $P$ -stable.*

**Theorem 8.12** (implicitly proved in Theorem 4 in [10]). *The set of directed degree sequences  $\vec{\mathbf{D}}$  satisfying*

$$2 \leq \max(\Delta_{\text{out}}, \Delta_{\text{in}}) < \frac{1}{\sqrt{2}}\sqrt{m-4},$$

*is  $P$ -stable.*

**Theorem 8.13** (implicitly proved in Theorem 5 in [10]). *The set of directed degree sequences  $\vec{\mathbf{D}}$  satisfying*

$$\begin{aligned} (\Delta_{\text{out}} - \delta_{\text{out}}) \cdot (\Delta_{\text{in}} - \delta_{\text{in}}) &\leq 2 - n + \\ &+ \max\left(\delta_{\text{out}}(n - \Delta_{\text{in}} - 1) + \delta_{\text{in}} + \Delta_{\text{out}}, \delta_{\text{in}}(n - \Delta_{\text{out}} - 1) + \delta_{\text{out}} + \Delta_{\text{in}}\right) \end{aligned}$$

*is  $P$ -stable.*

## 9. Summary

To summarize the new results of the paper, let us present Table 2, an updated version of Table 1 which contains both entirely new and improved results. A strongly stable class is, as the name suggests, naturally  $P$ -stable, see the recent paper of Amanatidis and Klier [1]. Their results already provide a unified framework for proving all previously known bipartite and UC degree sequence results.

The flexibility of our unified method allowed us to extend the rapid mixing results of the swap Markov chain in two directions (in the table):

- vertically (power of machinery) to  $P$ -stable degree sequence classes, and
- horizontally (applicability of machinery) to directed degree sequences.

Theorem 8.7 extends the class of power-law like degree sequences having an almost uniform sampler from degree sequences obeying a power-law density-bound with  $\gamma > 2.5$  to ones conforming to a power-law distribution-bound with  $\gamma > 2$ . Empirical evidence suggests that this latter class contains every real-world network following a power-law [12].

We have also shown that the degree sequence of the Erdős-Rényi random graph  $G(n, p)$  is rapidly mixing with high probability as  $n \rightarrow \infty$ , for any (including non-constant) edge probability.

The notion of  $P$ -stability is a natural obstacle on the rapid mixing of the swap Markov chain [20, 21], and it would be really intriguing to find even a small rapidly mixing degree sequence class which is not  $P$ -stable.

Finding the bipartite and directed analogues of Theorem 8.5 seems to be a relatively easy and moderately rewarding open problem. For example, such a theorem would probably be sufficient to prove that the bipartite and directed Erdős-Rényi graphs have  $P$ -stable degree sequences asymptotically almost surely.



UC degree sequences	bipartite deg. seq.	directed deg. seq.
regular [3]	(half-)regular [25]	regular [14]
	almost half regular [7]	
$\Delta \leq \frac{1}{3}\sqrt{2m}$ [16]	$\Delta \leq \frac{1}{\sqrt{2}}\sqrt{m}$ [10]	$\Delta < \frac{1}{\sqrt{2}}\sqrt{m-4}$ [10]
Power-law distribution-bound, $\gamma > 2$		
$(\Delta - \delta + 1)^2 \leq 4 \cdot \delta(n - \Delta - 1)$ [1]	$(\Delta - \delta)^2 \leq \delta(\frac{n}{2} - \Delta)$ [9, 10]	similar to the bip. case [9, 10]
Erdős-Rényi $G(n, p)$ with high prob.	Bip. E.R. with edge prob. $p, 1 - p \geq 4\sqrt{\frac{2 \log n}{n}}$ [9, 10]	similar to the bip. case [9, 10]
strongly stable degree sequence classes [1]		
<b><i>P</i>-stable degree sequence classes</b>		

Table 2: Updated version of Table 1 with the new results in this paper. Here  $\Delta$  and  $\delta$  denote the maximum and minimum degrees, respectively. Half of the sum of the degrees is  $m$ , and  $n$  is the number of vertices. The notation is similar for bipartite and directed degree sequences. Some technical conditions have been omitted. Gray text is used for cells which have not been updated.

## References

- [1] Amanatidis, G. – Kleer, P.: Rapid Mixing of the Switch Markov Chain for Strongly Stable Degree Sequences, *arXiv preprint* **1803.01338** (2018), 1–21.
- [2] Berger, A. – Müller-Hannemann, M.: Uniform sampling of digraphs with a fixed degree sequence, in *Graph Theoretic Concepts in Computer Science*, **LNCS 6410** Springer, Berlin, (2010), 220–231. DOI:10.1007/978-3-642-16926-7\_21
- [3] Cooper, C. – Dyer, M. – Greenhill, C.: Sampling regular graphs and a peer-to-peer network, *Comb. Prob. Comp.* **16** (4) (2007), 557–593. DOI:10.1017/S0963548306007978
- [4] Dyer, M. – Greenhill, C.: Polynomial-time counting and sampling of two-rowed contingency tables, *Theoretical Computer Science* **246** (1) (2000), 265–278.
- [5] Erdős, P.L. – Király, Z. – Miklós, I.: On graphical degree sequences and realizations, *Combinatorics, Probability and Computing* (2013), 1–22. DOI:10.1017/S0963548313000096

- [6] Erdős, P.L. – Miklós, I. – Toroczkai, Z.: A decomposition based proof for fast mixing of a Markov chain over balanced realizations of a joint degree matrix, *SIAM J. Disc. Math.* **29** (1) (2015), pp. 481–499. DOI:10.1137/130929874
- [7] Erdős, P.L. – Kiss, Z.S. – Miklós, I. – Soukup, L.: Approximate Counting of Graphical Realizations, *PLOS ONE* (2015), pp 20. #e0131300. DOI:10.1371/journal.pone.0131300
- [8] Erdős, P.L. – Miklós, I. – Toroczkai, Z.: New classes of degree sequences with fast mixing swap Markov chain sampling, *Combinatorics Probability and Computing* **27** (2) (2018), 186–207. DOI:10.1017/S0963548317000499
- [9] Erdős, P.L. – Mezei, T.R. – Miklós, I.: Efficiently sampling the realizations of irregular, but linearly bounded bipartite and directed degree sequences, *arXiv* **1712.01709** (2017), 1–21.
- [10] Erdős, P.L. – Mezei, T.R. – Miklós, I. – Soltész, D.: Efficiently sampling the realizations of bounded, irregular degree sequences of bipartite and directed graphs, *PLOS One* **2018** (2018), # e0201995, 1–19. DOI:10.1371/journal.pone.0201995
- [11] Gale, D.: A theorem on flows in networks, *Pacific J. Math.* **7** (2) (1957), 1073–1082. DOI:10.2140/pjm.1957.7-2
- [12] Gao, P. – Wormald, N.: Enumeration of graphs with a heavy-tailed degree sequence, *Advances in Mathematics* **287** (2016), pp. 412–450. DOI:10.1016/j.aim.2015.09.002
- [13] Gao, P. – Wormald, N.: Uniform generation of random graphs with power-law degree sequences, in *SIAM Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms* (2018), pp. 1741–1758. DOI:10.1137/1.9781611975031.114
- [14] Greenhill, C.: A polynomial bound on the mixing time of a Markov chain for sampling regular directed graphs, *Elec. J. Combinatorics* **18** (2011), #P234.
- [15] Greenhill, C.: The switch Markov chain for sampling irregular graphs, in *Proc. 26th ACM-SIAM Symposium on Discrete Algorithms*, New York-Philadelphia (2015), pp. 1564–1572.
- [16] Greenhill, C. – Sfragara, M.: The switch Markov chain for sampling irregular graphs, *Theoretical Comp. Sci* **719** (2018), 1–20. DOI:10.1016/j.tcs.2017.11.010
- [17] Hakimi, S.L.: On the realizability of a set of integers as degrees of the vertices of a simple graph. *J. SIAM Appl. Math.* **10** (1962), 496–506.

- [18] Havel, V.: A remark on the existence of finite graphs. (in Czech), *Časopis Pěst. Mat.* **80** (1955), 477–480.
- [19] Jerrum, M. – Sinclair, A.: Approximating the Permanent, *SIAM J. Comput.*, **18** (6) (1989), 1149–1178. DOI:10.1137/0218077
- [20] Jerrum, M. – Sinclair, A.: Fast uniform generation of regular graphs, *Theoretical Computer Science* **73** (1) (1990), 91–100. DOI:10.1016/0304-3975(90)90164-D
- [21] Jerrum, M. – McKay, B.D. – Sinclair, A.: When is a graphical sequence stable?, in *Random Graphs Vol 2* (eds. A. Frieze and T. Łuczak; Wiley-Interscience) (1992), 101–116. ISBN: 978-0-471-57292-3
- [22] Kannan, R. – Tetali, P. – Vempala, S.: Simple Markov-chain algorithms for generating bipartite graphs and tournaments, Extended abstract, in *Proc. SODA '97* (1997), 193–200. ISBN:0-89871-390-0
- [23] Kleitman, D.J. – Wang, D.L.: Algorithms for constructing graphs and digraphs with given valences and factors, *Discrete Math.* **6** (1973), 79–88.
- [24] LaMar, M.D.: Directed 3-Cycle Anchored Digraphs And Their Application In The Uniform Sampling Of Realizations From A Fixed Degree Sequence, in *ACM & IEEE & SCS Proc. of 2011 Winter Simulation Conference* (Eds. S. Jain, R.R. Creasey et. al.) (2011), 1–12.
- [25] Miklós, I. – Erdős, P.L. – Soukup, L.: Towards random uniform sampling of bipartite graphs with given degree sequence, *Electronic J. Combinatorics* **20** (1) (2013), #P16, 1–49.
- [26] Morris, B. – Sinclair, A.: Random Walks on Truncated Cubes and Sampling 0-1 Knapsack Solutions, *SIAM J. Comput.* **34** (1) (2004), 195–226.
- [27] Petersen, J.: Die Theorie der regulären Graphen, *Acta Math.* **15** (1891), 193–220.
- [28] Ryser, H. J.: Combinatorial properties of matrices of zeros and ones, *Canad. J. Math.* **9** (1957), 371–377.
- [29] Sinclair, A.: Improved bounds for mixing rates of Markov chains and multicommodity flow, *Combin. Probab. Comput.* **1** (1992), 351–370. DOI:10.1017/S0963548300000390