Approximately Sampling and Counting Graphs with Near-Regular Degree Intervals

APPROXIMATE SAMPLING AND COUNTING OF GRAPHS WITH NEAR-REGULAR DEGREE INTERVALS*

GEORGIOS AMANATIDIS† AND PIETER KLEER‡

Abstract. The approximate uniform sampling of graphs with a given degree sequence is a well-known, extensively studied problem in theoretical computer science and has significant applications, e.g., in the analysis of social networks. In this work we study a generalization of the problem, where *degree intervals* are specified instead of a single degree sequence. We are interested in sampling and counting graphs whose degree sequences satisfy the corresponding degree interval constraints. A natural scenario where this problem arises is in hypothesis testing on networks that are only partially observed. We provide the first *fully polynomial almost uniform sampler (FPAUS)* as well as the first *fully polynomial randomized approximation scheme (FPRAS)* for sampling and counting, respectively, graphs with near-regular degree intervals, i.e., graphs in which every node has a degree from an interval not too far away from a given $r \in \mathbb{N}$. In order to design our FPAUS, we rely on various state-of-the-art tools from Markov chain theory and combinatorics. In particular, by carefully using Markov chain decomposition and comparison arguments, we reduce part of our problem to the recent breakthrough of Anari, Liu, Oveis Gharan, and Vinzant (2019) on sampling a base of a matroid under a strongly log-concave probability distribution, and we provide the first non-trivial algorithmic application of a breakthrough asymptotic enumeration formula of Liebenau and Wormald (2017). As a more direct approach, we also study a natural Markov chain recently introduced by Rechner, Strowick and Müller-Hannemann (2018), based on three local operations—switches, hinge flips, and additions/deletions of an edge. We obtain the first theoretical results for this Markov chain, showing it is rapidly mixing for the case of near-regular degree intervals of size at most one.

Key words. graph sampling, switch Markov chain, degree intervals

AMS subject classifications. 68Q25, 05C80, 68R10

1. Introduction.

The (approximate) uniform sampling and counting of graphs with given degrees has received a lot of attention during the last few decades, see, e.g., [1, 4, 5, 7, 9, 13–17, 20, 22, 24–28, 30, 37, 41, 42, 44, 47–49, 56]. Given a degree sequence $\boldsymbol{d} = (d_1, \ldots, d_n)$, the goal of approximate uniform sampling is to design a randomized algorithm that outputs a labelled simple undirected graph G with degree sequence $\boldsymbol{d}$, according to a distribution that is close to the uniform distribution over the set of all graphs with this degree sequence. Such an algorithm is called an *approximate (uniform) sampler*. Approximate samplers find applications in fields such as complex network analysis, where they serve as null models for hypothesis testing. Consider, e.g., a social network with edges representing friendships or relationships. One might see a very high number of edges between a certain group of nodes and, based on this, conjecture that these nodes form a *community* of friends or colleagues. In order to test this hypothesis, one would like to be able to generate graphs with *similar characteristics* as the observed network and, based on these generated samples, decide how likely it is that there is a high number of edges between that particular group of nodes by chance alone. Here the characteristic of interest is the degree sequence of the observed network [52]. For determining how many samples are sufficient in order to test the hypothesis, we also need to be able to count the

*A preliminary version of this work appeared in the Proceedings of the 40th International Symposium on Theoretical Aspects of Computer Science (STACS 2023).

Funding: This work has been partially supported by project MIS 5154714 of the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGenerationEU Program.

†School of Mathematics, Statistics and Actuarial Science, University of Essex, United Kingdom, and Archimedes/Athena RC, Greece; georgios.amanatidis@essex.ac.uk.

‡Tilburg School of Economics and Management, Tilburg University, The Netherlands; p.s.kleer@tilburguniversity.edu.

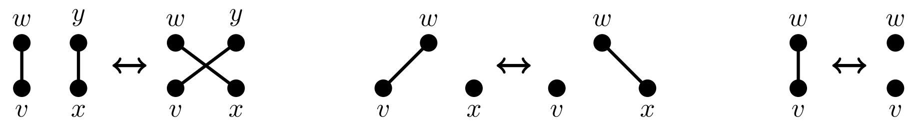

FIG. 1. *Left to right: switch on v, w, x, y; hinge flip on v, w, x; edge addition/deletion on v, w.*

number of graphs with the given degree sequence.

In practice, it is not always possible to have exact knowledge of the degree sequence of an observed network, due to erroneous measurements. In order to overcome this, there is a need for more robust null models. One such model was proposed by Rechner, Strowick and Müller-Hannemann [54]. Instead of a given degree sequence $\boldsymbol{d}$, the null models now consist of all graphs with given *degree interval* constraints $[\ell_i, u_i]$, for $i \in [n] = \{1, \ldots, n\}$. In this case we say that a graph G has degrees in the interval $[\boldsymbol{\ell}, \boldsymbol{u}]$ with $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_n)$ and $\boldsymbol{u} = (u_1, \ldots, u_n)$. The algorithmic task at hand then becomes to develop algorithms for sampling and counting graphs from the set $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ of all graphs satisfying the interval constraints. An intuitive two-step approach for solving this problem is to first sample *according to the correct proportional distribution* a degree sequence $\boldsymbol{d} = (d_1, \ldots, d_n)$ from the set of all degree sequences satisfying the interval constraints $\ell_i \leq d_i \leq u_i$, for $i \in [n]$, and then sample uniformly at random a graphical realization from the set $\mathcal{G}(\boldsymbol{d})$ of all graphs with degree sequence $\boldsymbol{d}$. A crucial difficulty that arises here is that the probability with which each degree sequence $\boldsymbol{d}$ needs to be sampled in the first step is not obvious. This probability should be proportional to the number $|\mathcal{G}(\boldsymbol{d})|$ which is not known in general.

To make the problem more concrete, we give a brief example in the context of the social network application that we started out with. Suppose we have a partially observed network. For a given node i, we let ℓ_i be the number of observed edges adjacent to i, δ_i the number of missing observations and, thus, $u_i = n - 1 - (\ell_i + \delta_i)$ the number of observed non-edges (i.e., pairs $\{i, j\}$ for which we know there is no edge between nodes i and j). There are now two extreme cases: either all missing observations are non-edges, meaning that node i has degree ℓ_i, or all missing observations are indeed edges, meaning that node i has degree u_i. Hence, we are interested in sampling (and counting) graphs for which each node i has a degree in the interval $[\ell_i, u_i]$, for every $i \in [n]$. In this and other similar settings, these problems seem to be natural and elegant generalizations of the classic graph sampling and counting problems.

Towards sampling graphs with given degree intervals, Rechner et al. [54] introduced a Markov chain based on three simple operations: *switches*, *hinge flips* and *additions/deletions*. The chain in each step selects one of these operations uniformly at random and performs it, if possible. We call this chain the *degree interval Markov chain*. The operations are shown in Figure 1 and a formal definition is given in Section 2. These three operations are the ones described by Coolen et al. [14] as the most commonly used operations in Markov Chain Monte Carlo algorithms for the generation of simple undirected graphs *in practice*. This serves as additional motivation for rigorously studying Markov chains based on these operations. We will also be interested in the *switch-hinge flip Markov chain* that only uses the switch and hinge flip operations. The hinge flip and switch operations are of particular interest both in theory and in practice as they preserve the number of edges and the degree sequence of a graph, respectively.

1.1. Our contributions. In this work, we give the first efficient approximate sampler and approximate counter for graphs with so-called *near-regular* degree intervals.

Near-regularity here refers to the fact that all graphs have degrees which are close to a common value up to a sublinear margin. To be more precise, we show that there is a *fully polynomial almost uniform sampler (FPAUS)* and a *fully polynomial randomized approximation scheme (FPRAS)* (for formal definitions see Section 2), in case the degree intervals are close to a common value $r = r(n) \in \mathbb{N}$, i.e., if $[\ell_i, u_i] \subseteq [r - \min\{r, n-r-1\}^\alpha, r + \min\{r, n-r-1\}^\alpha]$ for some $0 < \alpha \leq \frac{1}{2}$.[1] We also need a minor technical assumption on the value of r in order to avoid some (arguably not very interesting) boundary cases. The main result of this work is Theorem 1.1 below.

For vectors $\boldsymbol{a} = (a_1, \ldots, a_n), \boldsymbol{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$, we write $\boldsymbol{a} \leq \boldsymbol{b}$ if $a_i \leq b_i$ for all $i \in [n]$. Given $\boldsymbol{\ell}, \boldsymbol{u} \in \mathbb{N}^n$, by $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ we denote the set of all graphs G whose degree sequence $\boldsymbol{d}(G)$ satisfies $\boldsymbol{\ell} \leq \boldsymbol{d}(G) \leq \boldsymbol{u}$.

THEOREM 1.1. *Let $0 < \alpha \leq 1/2$ and $0 < \sigma < 1$ be fixed. Let $r = r(n)$ with $2 \leq r \leq (1-\sigma)n$. If for every node $i \in [n]$ it holds that $[\ell_i, u_i] \subseteq [r - \min\{r, n-r-1\}^\alpha, r + \min\{r, n-r-1\}^\alpha]$, then there is an FPAUS for the approximate uniform sampling of graphs from $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ and an FPRAS for approximating $|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$.*

For given degree intervals $[\boldsymbol{\ell}, \boldsymbol{u}]$ and $m \in \mathbb{N}$, we write $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ for the set of graphs G whose degree sequence $\boldsymbol{d}(G)$ satisfies $\boldsymbol{\ell} \leq \boldsymbol{d}(G) \leq \boldsymbol{u}$ *and* $\sum_i d_i = 2m$. By using reductions between approximate sampling and approximate counting (see Appendix B) we get that to prove Theorem 1.1 it suffices to show the existence of an FPAUS for sampling from $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$. To this end, we show that the switch-hinge flip Markov chain is rapidly mixing under the conditions of Theorem 1.1. This result is summarized in Theorem 1.2.

THEOREM 1.2. *Let α, σ, and r be as in Theorem 1.1. If $[\ell_i, u_i] \subseteq [r - \min\{r, n-r-1\}^\alpha, r + \min\{r, n-r-1\}^\alpha]$, for all $i \in [n]$, and $2m \in \left[\sum_i \ell_i, \sum_i u_i\right]$, then the switch-hinge flip Markov chain is rapidly mixing on $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$.*

A more direct approach for sampling from $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ than the one behind Theorem 1.1 would be to use the degree interval Markov chain. An interesting open question is whether this chain is rapidly mixing under the assumptions in Theorem 1.1 (or under weaker assumptions). As a first step into this direction, we show rapid mixing when all the degree intervals have size at most one, i.e., when $u_i - 1 \leq \ell_i \leq u_i$.

THEOREM 1.3. *Let α, σ, and r be as in Theorem 1.1. If $[\ell_i, u_i] \subseteq [r - \min\{r, n-r-1\}^\alpha, r + \min\{r, n-r-1\}^\alpha]$ and $u_i - 1 \leq \ell_i \leq u_i$, for all $i \in [n]$, then the degree interval Markov chain is rapidly mixing on $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$.*

The technical novelty of our work lies in the highly nontrivial combination of state-of-the-art tools from Markov chain theory and combinatorics. An overview of our proof approach is given in Section 3. It relies on Markov chain decomposition and comparison techniques of Martin and Randall [46], rapid mixing results for the switch Markov chain by Amanatidis and Kleer [1], the breakthrough work of Anari et al. [2] on strongly log-concave probability distributions, and the work of Liebenau and Wormald [44] regarding asymptotic enumeration formulas for the number of near-regular graphs.

Remark 1.4. Our theorems—and all the building blocks used in their proofs—are shown to be true for all $n \geq n_0$, where $n_0 \in \mathbb{N}$ is a constant that depends on the other constant parameters involved. It is straightforward that for $n < n_0$ our results are always

[1] The parameter α models the maximum length of the degree intervals that we allow; this length should be relatively small compared to r. Note that an assumption of this kind is to be expected. Otherwise, we would be also solving the problem of (approximately) uniformly sampling a graph with *any* given degree sequence, which is a long-standing open problem. Furthermore, one could work in an additional polylogarithmic factor, based on the n^ϵ factor in Theorem 2.6, but we leave this to the interested reader.

true.

1.2. Related work. There is an extensive literature on the problem of sampling graphs with a given degree sequence, particularly on Markov Chain Monte Carlo (MCMC) methods. Jerrum and Sinclair [37] provide an approximate uniform sampler and an approximate counter for P*-stable* degree sequences, for which the number of graphical realizations of a given degree sequence does not vary too much under small perturbations of the sequence. A first step beyond P-stability was recently made by Erdős et al. [21]. Jerrum, Sinclair and Vigoda [39] provide an approximate sampler (and counter) for arbitrary bipartite degree sequences by reducing the problem to sampling perfect matchings in an appropriate graph representation of the given instance. The work of Bézakova, Bhatnagar and Vigoda [6] provides a more direct approach. There are also various non-MCMC methods available in the literature, see, e.g., [4, 27, 28, 42, 48, 56]. One MCMC approach that has received considerable attention is the *switch Markov chain*, based on the switch operation in Figure 1. This is a simpler, more direct approach than reducing the problem to sampling a perfect matching from a large auxiliary graph. The chain was first analyzed by Kannan, Tetali and Vempala [41], and has been extensively studied, see, e.g., [1, 15, 22, 49]. The state of the art on its mixing time is the work of Erdős et al. [22], who show that the chain is rapidly mixing for all P-stable degree sequences.

Rechner et al. [54] introduce the degree interval Markov chain for the *bipartite* version of the problem of sampling graphs with given degree intervals and show its irreducibility for arbitrary degree intervals. Very recently, Erdős, Mezei and Miklós [23] generalized our Theorem 1.3 to intervals of length 1 centered around P-stable degree sequences. We consider the fact that their meticulous direct approach does not go beyond length 1 as another indication of the difficulty of directly arguing about the degree interval Markov chain.

The decomposition theorem of Martin and Randall [46] we use (Theorem 2.4), based on the decomposition method of Madras and Randall [45], also appeared in an unpublished manuscript by Caracciolo, Pelissetto and Sokal [12]. Erdős et al. [25] use a related decomposition approach for sampling *balanced joint degree matrix* realizations.

The result of Liebenau and Wormald [44] builds on a long line of work on asymptotic expressions for the number of graphs with given degrees. Indicatively, Bender and Canfield [5] gave a formula for bounded degree sequences and Bollobás [9] for r-regular sequences with $r = O(\sqrt{\log(n)})$. McKay and Wormald gave expressions both for sparse sequences with maximum degree $o(n^{1/2})$ [48] and for a certain dense regime [47].

Anari et al. [2], in a breakthrough recent work, gave the first polynomial time algorithm for approximate sampling a base of a matroid under a strongly log-concave probability distribution. The theory of strongly log-concave (or Lorentzian) polynomials dates back to the work of Gurvits [32], and was further developed by Anari, Oveis Gharan and Vinzant [3] and Brändén and Huh [11]. In another recent work, Kleer [43] made a connection between asymptotic enumeration formulas and strongly log-concave polynomials for a case of sparse bipartite graphs where only the degrees on one side of the bipartition can vary.

1.3. Outline. In Section 2 we provide all the necessary preliminaries. We then continue with a proof overview for Theorems 1.2 and 1.3 in Section 3. For readers with some familiarity regarding Markov chains and some intuition about degree sequence problems, it should be possible to go through (most of) Section 3 without delving into (the admittedly long) Section 2 first. As one of the main building blocks of the proof of Theorem 1.2, we show in Section 4 that the asymptotic formula of Liebenau and Wormald [44], when restricted to the degree interval regime of Theorem 1.1, approximately gives rise to a so-called strongly log-concave polynomial, a result which might be of independent interest. Section 5 contains all of the remaining arguments about the Markov chains used in our proofs.

2. Preliminaries.

We need a variety of preliminaries for this work, that are collected in this section (except for the details on the modified log-Sobolev constant, which are deferred to Appendix A.)

2.1. M-convexity and strongly log-concave polynomials.

We start with the notion of M-convexity for functions [50, 51]. Let $\nu : \mathbb{Z}^n_{\geq 0} \to \mathbb{R} \cup \{\infty\}$ be a function. The *effective domain* of ν is given by $\mathrm{dom}(\nu) = \{\boldsymbol{\alpha} \in \mathbb{Z}^n_{\geq 0} : v(\boldsymbol{\alpha}) < \infty\}$. The function ν is called *$M^\natural$-convex* if it satisfies the *(symmetric) exchange property*: For any $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathrm{dom}(\nu)$ and any $i \in [n]$ satisfying $\alpha_i > \beta_i$, there exists a $j \in [n]$ such that $\alpha_j < \beta_j$ and

$$\nu(\boldsymbol{\alpha}) + \nu(\boldsymbol{\beta}) \geq \nu(\boldsymbol{\alpha} - \boldsymbol{e}_i + \boldsymbol{e}_j) + \nu(\boldsymbol{\beta} + \boldsymbol{e}_i - \boldsymbol{e}_j)\,, \tag{2.1}$$

where $\boldsymbol{e}_k$ is defined as $\boldsymbol{e}_k(\ell) = 1$ if $k = \ell$ and $\boldsymbol{e}_k(\ell) = 0$ otherwise. The function ν is called *M-convex* if it is $M^\natural$-convex *and* there is an $d \in \mathbb{N}$ such that $\mathrm{dom}(\nu) \subseteq \{\boldsymbol{\alpha} : \sum_i \alpha_i = d\}$. A subset $C \subseteq \mathbb{Z}^n_{\geq 0}$ is called M-convex if the indicator function $\nu_C : \mathbb{Z}^n_{\geq 0} \to \mathbb{R} \cup \{\infty\}$, given by $\nu_C(\boldsymbol{\alpha}) = 1$ if $\boldsymbol{\alpha} \in C$ and $\nu_C(\boldsymbol{\alpha}) = 0$ otherwise, is M-convex.

We write $\mathbb{R}[x_1, \ldots, x_n]$ to denote the set of all polynomials in $x_1, \ldots, x_n$ with real coefficients. We consider polynomials $p \in \mathbb{R}[x_1, \ldots, x_n]$ with non-negative coefficients. For a vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{Z}^n_{\geq 0}$, we write $\partial^{\boldsymbol{\beta}} = \prod_{i=1}^n \partial^{\beta_i}_{x_i}$ to denote the partial differential operator that differentiates a function β_i times with respect to x_i for $i = 1, \ldots, n$. For $\boldsymbol{\alpha} \in \mathbb{Z}^n_{\geq 0}$, we write $x^{\boldsymbol{\alpha}}$ to denote $\prod_{i=1}^n x_i^{\alpha_i}$. Furthermore, we write $\boldsymbol{\alpha}! = \prod_i \alpha_i!$, and for $\boldsymbol{\alpha}, \boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$ with $\alpha_i \leq \kappa_i$ for all i, we write $\binom{\boldsymbol{\kappa}}{\boldsymbol{\alpha}} = \prod_{i=1}^n \binom{\kappa_i}{\alpha_i}$. For a constant $c \in \mathbb{N}$ with $c \geq \max_i \alpha_i$, we write $\binom{c}{\boldsymbol{\alpha}} = \prod_{i=1}^n \binom{c}{\alpha_i}$. Let $\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}$ and the Cartesian product $K = \times_i \{0, \ldots, \kappa_i\}$. Let $w : K \to \mathbb{R}_{\geq 0}$ be a weight function. The *generating polynomial* of w is $g_{\boldsymbol{\kappa}}(x) = \sum_{\boldsymbol{\alpha} \in K} w(\boldsymbol{\alpha}) x^{\boldsymbol{\alpha}}$. The *support* of $g_{\boldsymbol{\kappa}}$ is the set $\mathrm{supp}(g_{\boldsymbol{\kappa}}) = \{\boldsymbol{\alpha} \in K : w(\boldsymbol{\alpha}) > 0\}$. The polynomial $g_{\boldsymbol{\kappa}}$ is called d-homogeneous if $|\boldsymbol{\alpha}| = \sum_i \alpha_i = d$ for all $\boldsymbol{\alpha} \in \mathrm{supp}(g_{\boldsymbol{\kappa}})$.

Definition 2.1 (Strong log-concavity [32]). *A polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ with non-negative coefficients is called* log-concave *on a subset $S \subseteq \mathbb{R}^n_{\geq 0}$ if its Hessian $\nabla^2 \log(p)$ is negative semidefinite on S. A polynomial p is called* strongly log-concave (SLC) *on S if for any $\boldsymbol{\beta} \in \mathbb{N}^n$, we have that $\partial^{\boldsymbol{\beta}} p$ is log-concave.*

For convenience, the zero polynomial is defined to be SLC always. Finally, if the generating polynomial $g_{\boldsymbol{\kappa}}$ is SLC, then the probability distribution $\pi(\boldsymbol{\alpha}) \propto w(\boldsymbol{\alpha})$ is called SLC as well. We next state some properties of SLC polynomials that will be used in this work.

Proposition 2.2 ([11]). *If $p \in \mathbb{R}[x_1, \ldots, x_n]$ is SLC and $\gamma \in \mathbb{R}_{\geq 0}$, then γp is SLC.*

Proposition 2.3 (Following from [11]).[2] *Let $\nu : \mathbb{Z}^n_{\geq 0} \to \mathbb{R} \cup \{\infty\}$ with $\mathrm{dom}(\nu) \subseteq \{0, 1, \ldots, n-1\}^n$ and let*

$$f_{\boldsymbol{\kappa}}(x) = \sum_{\boldsymbol{\alpha} \in \mathrm{dom}(\nu)} \frac{1}{\boldsymbol{\alpha}!} e^{-\nu(\boldsymbol{\alpha})} x^{\boldsymbol{\alpha}} \quad \text{and} \quad g_{\boldsymbol{\kappa}}(x) = \sum_{\boldsymbol{\alpha} \in \mathrm{dom}(\nu)} \binom{\gamma}{\boldsymbol{\alpha}} e^{-\nu(\boldsymbol{\alpha})} x^{\boldsymbol{\alpha}}\,, \tag{2.2}$$

be $2m$-homogeneous polynomials, where $\gamma = (n-1, \ldots, n-1)$. If ν is M-convex, then $f_{\boldsymbol{\kappa}}$ and $g_{\boldsymbol{\kappa}}$ are SLC.

[2] Our $g_{\boldsymbol{\kappa}}$ is a slight variant the corresponding function of Theorem 3.14 of [11] with $q = 1/e$. The statement for this $g_{\boldsymbol{\kappa}}$ follows from a simple transformation of $f_{\boldsymbol{\kappa}}$ that preserves strong log-concavity, namely the operator that maps $x^{\boldsymbol{\alpha}}$ to $\boldsymbol{\alpha}!\binom{\gamma}{\boldsymbol{\alpha}} x^{\boldsymbol{\alpha}}$ [10].

2.2. Markov chains and mixing times. Let $\mathcal{M} = (\Omega, P)$ be an ergodic, time-reversible Markov chain with state space Ω, transition matrix P, and stationary distribution π. We write $P^t(x, \cdot)$ for the distribution over Ω at time step t with initial state $x \in \Omega$. The *total variation distance* of this distribution from stationarity at time t with initial state x is

$$\Delta_x(t) = \frac{1}{2} \sum_{y \in \Omega} \left| P^t(x, y) - \pi(y) \right|,$$

and the *mixing time* of $\mathcal{M}$ is

$$\tau(\epsilon) = \max_{x \in \Omega} \tau_x(\epsilon), \text{ where } \tau_x(\epsilon) = \min\{t : \Delta_x(t') \leq \epsilon \text{ for all } t' \geq t\} \text{ for } \epsilon > 0\,.$$

The chain $\mathcal{M}$ is said to be *rapidly mixing* if its mixing time can be upper bounded by a polynomial in $\ln(|\Omega|/\epsilon)$.

It is well-known that the matrix P only has real eigenvalues $1 = \lambda_0 > \lambda_1 \geq \cdots \geq \lambda_{|\Omega|-1} > -1$. We may replace P by $(P+I)/2$ to make the chain *lazy*, and hence guarantee that all its eigenvalues are non-negative. In that case, by $\text{Gap}(P) = 1 - \lambda_1$ we denote the spectral gap of P. In this work all Markov chains involved are lazy. It is well known that one can use the spectral gap to give an upper bound on the mixing time of Markov chain. That is, it holds that

$$\tau_x(\epsilon) \leq \frac{1}{2(1 - \lambda_1(P))} \Big(\log \pi(x)^{-1} + 2 \log \Big(\frac{1}{2\epsilon} \Big) \Big),$$

as it follows directly from Proposition 1 in [55]. Furthermore, if one has two Markov chains $\mathcal{M} = (\Omega, P)$ and $\mathcal{M}' = (\Omega, P')$ both with stationary distribution π and there are constants c_1, c_2 such that $c_1 P(x, y) \leq P'(x, y) \leq c_2 P(x, y)$ for all $x, y \in \Omega$ with $x \neq y$. Then (see, e.g., [46]) it follows that

$$c_1 \,\text{Gap}(P) \leq \text{Gap}(P') \leq c_2 \,\text{Gap}(P). \tag{2.3}$$

The *state space graph* of the chain $\mathcal{M}$ is the directed graph $\mathbb{G} = \mathbb{G}(\mathcal{M})$ with node set Ω that contains the edges $(x, y) \in \Omega \times \Omega$ for which $P(x, y) > 0$ and $x \neq y$. that contains an edge $(x, y) \in \Omega \times \Omega$ if and only if $P(x, y) > 0$ and $x \neq y$ (denoted by $x \sim y$). Let $\mathcal{P} = \bigcup_{x \neq y} \mathcal{P}_{xy}$, where $\mathcal{P}_{xy}$ is the set of simple paths between x and y in the state space graph $\mathbb{G}$. A *flow* f in Ω is a function $\mathcal{P} \to [0, \infty)$ satisfying $\sum_{p \in \mathcal{P}_{xy}} f(p) = \pi(x)\pi(y)$ for all $x, y \in \Omega$, $x \neq y$. The flow f can be extended to the oriented edges $e = (z, z')$ of $\mathbb{G}$ by setting $f(e) = \sum_{p \in \mathcal{P} : e \in p} f(p)$, so that $f(e)$ is the total flow routed through $e \in E(\mathbb{G})$. Let $\text{length}(f) = \max_{p \in \mathcal{P} : f(p) > 0} |p|$ be the length of a longest flow-carrying path, and let $\text{load}(e) = f(e)/Q(e)$ be the *load* of the edge e, where $Q(e) = \pi(x)P(x, y)$ for $e = (x, y)$. If $\text{load}(f) = \max_{e \in E(\mathbb{G})} \text{load}(e)$ is the maximum load of the flow, it holds that $\text{Gap}(P)^{-1} \leq \text{load}(f)\,\text{length}(f)$ (see, e.g, [55]).

We will sometimes also work (implicitly) with the so-called *modified log-Sobolev constant* $\rho = \rho(P)$. This constant can also be used to upper bound the mixing time of a Markov chain. In particular, it holds that

$$\tau_x(\epsilon) \leq \frac{1}{\rho(P)} \left(\log \log \pi(x)^{-1} + \log \left(\frac{1}{2\epsilon^2} \right) \right),$$

see, e.g., [8]. Details on $\rho(P)$ are given in Appendix A.

2.2.1. Markov chain decomposition. We describe a Markov chain decomposition of Martin and Randall [46] that follows the decomposition framework of Madras and Randall [45]. Let $\mathcal{M} = (\Omega, P)$ be a Markov chain and $\bigcup_{i=1}^{q} \Omega_i$ be a partition of Ω for some $q \in \mathbb{N}$. We define the restriction Markov chains $\mathcal{M}_i = (\Omega_i, P_{\Omega_i})$ as follows. For $x \in \Omega_i$ we let $P_{\Omega_i}(x,y) = P(x,y)$ if $x, y \in \Omega_i$ with $x \neq y$, and $P_{\Omega_i}(x,x) = 1 - \sum_{y \in \Omega_i, y \neq x} P_{\Omega_i}(x,y)$. Furthermore, let $\partial_i(\Omega_j) = \{y \in \Omega_j : \exists x \in \Omega_i \text{ with } P(x,y) > 0\}$ be the set of elements in Ω_j that can be reached with positive probability in one transition of the chain $\mathcal{M}$ from some element in Ω_i.

Let $\mathcal{M}_{\text{MH}} = ([q], P_{\text{MH}})$ be (the Metropolis-Hastings variant of) the projection Markov chain on $[q] = \{1, \ldots, q\}$. That is, $P_{\text{MH}}(i,j) > 0$ if and only if $\partial_i(\Omega_j) \neq \emptyset$ and, in that case, for $i \neq j$,

$$P_{\text{MH}}(i,j) = \frac{1}{2\Delta} \min\left\{1, \frac{\pi(\Omega_j)}{\pi(\Omega_i)}\right\}, \tag{2.4}$$

where Δ is the maximum out-degree in the state space graph of $\mathcal{M}_{\text{MH}}$, while

$$P_{\text{MH}}(i,i) = 1 - \sum_{j \in [q] \setminus \{i\}} P_{\text{MH}}(i,j).$$

Note that $\mathcal{M}_{\text{MH}}$ has stationary distribution $\pi_{\text{MH}}(i) = \pi(\Omega_i)$ for $i \in \{1, \ldots, q\}$ and a holding probability of at least 1/2. We will use the following decomposition theorem from [46].

THEOREM 2.4 ([46], Corollary 3.3). *Suppose there exist $\beta > 0$ and $\gamma > 0$ such that $P(x,y) \geq \beta$ for all x, y that are adjacent in $\mathbb{G}(\mathcal{M})$, and $\pi(\partial_i(\Omega_j)) \geq \gamma\pi(\Omega_j)$ for all i, j that are adjacent in $\mathbb{G}(\mathcal{M}_{\text{MH}})$. Then* $\text{Gap}(P) \geq \beta\gamma \cdot \text{Gap}(P_{\text{MH}}) \cdot \min_{i=1,\ldots,q} \text{Gap}(P_{\Omega_i})$.

2.2.2. Load-exchange Markov chain. In this work, we will need a weighted version of the *base-exchange Markov chain* studied by Anari et al. [2]. Let π be a strongly log-concave probability distribution with $\pi(\boldsymbol{\alpha}) \propto w(\boldsymbol{\alpha})$ whose support forms an M-convex set C. We define the *(unit) load-exchange Markov chain* on $C \subseteq Z^n_{\geq 0}$:

> Assuming $\boldsymbol{\alpha} \in C$ is the current state of the *(unit) load-exchange Markov chain*:
>
> - Select an element $i \in [n]$ uniformly at random.
> - Pick an $\boldsymbol{\alpha}' \in C$ with $\boldsymbol{\alpha}' \geq \boldsymbol{\alpha} - \boldsymbol{e_i}$ with probability $\propto w(\boldsymbol{\alpha}')$ among all such $\boldsymbol{\alpha}'$.

Similarly to the base-exchange Markov chain [2], the above procedure defines an ergodic, time-reversible Markov chain with stationary distribution π over C given by $\pi(\boldsymbol{\alpha}) \propto w(\boldsymbol{\alpha})$. Using the notion of *polarization* for SLC polynomials [11], in combination with a simple Markov chain comparison argument (as in Appendix A.1), Corollary 2.5 can be shown. The proof (which is implicitly given in [43]), roughly speaking, uses a reduction to the case of matroids, after which a result of Cryan et al. [18] gives the desired result.

COROLLARY 2.5. *Let $\boldsymbol{\kappa} = (n, \ldots, n)$ and suppose that the d-homogeneous polynomial $g_{\boldsymbol{\kappa}}(x) = \sum_{\boldsymbol{\alpha} \in K} w(\boldsymbol{\alpha}) x^{\boldsymbol{\alpha}} \in \mathbb{R}[x_1, \ldots, x_n]$ is SLC. Then the transition matrix P of the load-exchange Markov chain on $\text{supp}(g_{\boldsymbol{\kappa}})$ satisfies $\rho(P) \geq 1/(n^2 d)$, where $\rho(P)$ is the modified log-Sobolev constant of P.*

2.2.3. Degree intervals and the switch-hinge flip Markov chain. A sequence of non-negative integers $\boldsymbol{d} = (d_1, \ldots, d_n)$ is called a *graphical degree sequence* if there exists a simple, undirected, labeled graph $G = (V, E)$ on nodes $V = [n]$, where node i

has degree d_i, for $i \in V$. Such a graph is called a *(graphical) realization of* $\boldsymbol{d}$. By $\mathcal{G}(\boldsymbol{d})$ we denote the set of all graphical realizations of $\boldsymbol{d}$, while by $\boldsymbol{d}(G)$ we denote the degree sequence of a graph G. For given vectors $\boldsymbol{\ell} = (\ell_1, \dots, \ell_n)$ and $\boldsymbol{u} = (u_1, \dots, u_n)$ of non-negative integers, we define $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u}) = \bigcup_{\boldsymbol{\ell} \leq \boldsymbol{d} \leq \boldsymbol{u}} \mathcal{G}(\boldsymbol{d})$ as the set of all graphical realizations G satisfying $\boldsymbol{\ell} \leq \boldsymbol{d}(G) \leq \boldsymbol{u}$, meaning $\ell_i \leq d_i(G) \leq u_i$ for all $i \in V$. For $m \in \mathbb{N}$, we define $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ as the set of all graphical realizations $G \in \mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ with precisely m edges, i.e., with $\sum_i d_i(G) = 2m$. Finally, we define the set of all degree sequences satisfying the degree interval constraints, and whose total sum of the degrees equals $2m$, as $\mathcal{D}_m = \{\boldsymbol{d} : \boldsymbol{\ell} \leq \boldsymbol{d} \leq \boldsymbol{u} \text{ and } \sum_i d_i = 2m\}$.

A *fully polynomial almost uniform sampler (FPAUS)* for sampling graphs with given degree intervals $[\boldsymbol{\ell}, \boldsymbol{u}]$ is an algorithm that, for any $\epsilon > 0$, outputs a graph $G \in \mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ according to a distribution $\tilde{\pi}$ such that $d_{\text{TV}}(\pi, \tilde{\pi}) \leq \epsilon$, where π is the uniform distribution over $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$, and runs in time polynomial in n and $\log(1/\epsilon)$. A *fully polynomial randomized approximation scheme (FPRAS)* for the problem is an algorithm that, for every ϵ, $\delta > 0$, outputs $|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$ up to a multiplicative factor $(1 + \epsilon)$ with probability at least $1 - \delta$, in time polynomial in n, $1/\epsilon$ and $\log(1/\delta)$. Analogous definitions hold for the set $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ for a given m.

First we define the *switch-hinge flip Markov chain* to uniformly sample elements from $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ based on two of the local operations of Figure 1.

Assuming $G \in \mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ is the current state of the *switch-hinge flip Markov chain*:

- With probability $2/3$, do nothing.
- With probability $1/6$, try to perform a *switch operation*: Choose an ordered tuple of distinct nodes (v, w, x, y) uniformly at random. If $\{w, v\}, \{x, y\} \in E(G)$, and $\{y, v\}, \{x, w\} \notin E(G)$, then delete $\{w, v\}, \{x, y\}$ from $E(G)$, and add $\{y, v\}, \{x, w\}$ to $E(G)$.
- With probability $1/6$, try to perform a *hinge flip operation*: Choose an ordered tuple of distinct nodes (v, w, x) uniformly at random. If $\{w, v\} \in E(G)$ and $\{w, x\} \notin E(G)$, then delete $\{w, v\}$ from and add $\{w, x\}$ to $E(G)$ if the resulting graph is in $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$.

Similarly, we define the *degree interval Markov chain* of Theorem 1.3, that can also perform addition/deletion operations.

Assuming $G \in \mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ is the current state of the *degree interval Markov chain*:

- With probability $1/2$, do nothing.
- Otherwise:
  - With probability $1/6$, try to perform a *switch operation*.
  - With probability $1/6$, try to perform a *hinge flip operation*.
  - With probability $1/6$, try to perform an *addition/deletion operation*: select an ordered tuple of distinct nodes (v, w) uniformly at random. If $\{v, w\} \in E(G)$, then delete it from $E(G)$ if the resulting graph is in $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$. Similarly, if $\{v, w\} \notin E(G)$, then add it to $E(G)$ if the resulting graph is in $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$.

Due to the symmetry of the transition probabilities, it is not hard to see that both chains are time-reversible with respect to the uniform distribution. Also because of the holding probability of at least $1/2$, the chains are aperiodic. Finally, by a simple counting argument, there exists polynomials $t(n), t'(n)$ such that $P_{\mathcal{G}(\boldsymbol{\ell},\boldsymbol{u})}(G,H) \geq 1/t(n)$ for all $G, H \in \mathcal{G}(\boldsymbol{\ell},\boldsymbol{u})$ with $P_{\mathcal{G}(\boldsymbol{\ell},\boldsymbol{u})}(G,H) > 0$ and $P_{\mathcal{G}_m(\boldsymbol{\ell},\boldsymbol{u})}(G,H) \geq 1/t'(n)$ for all $G, H \in \mathcal{G}_m(\boldsymbol{\ell},\boldsymbol{u})$ with $P_{\mathcal{G}_m(\boldsymbol{\ell},\boldsymbol{u})}(G,H) > 0$ respectively. The irreducibility of the chain (i.e., the fact that its state space is strongly connected) for the intervals of interest will follow implicitly from our analysis, in particular Lemmata 5.1 and 5.3.

2.3. Near-regular degree sequences. Let $r \geq 1$ be a given integer. A degree sequence $\boldsymbol{d}$ is said to be r-regular if $d_i = r$ for $i \in [n]$. For a fixed $0 \leq \alpha < 1$ we say that a degree sequence $\boldsymbol{d}$ is *(α, r)-near-regular* if $\max_i |d_i - r| \leq r^{\alpha}$. When we do not refer to a specific (α, r) pair, we just write about *near-regular* degree sequences. Note that r above can be a function of the length of a degree sequence. It will be convenient to refer to the class $\mathcal{F}_{(\alpha,r)}[n]$ of (α, r)-near-regular degree sequences of length at least n.

We state some properties of near-regular degree sequences that we will use later. The most important result is Theorem 2.6 below. We use a slightly different formulation than that of [44].[3] For any degree sequence $\boldsymbol{d} = (d_1, \ldots, d_n)$, define

$$\xi = \sum_i d_i/n, \quad \mu = \xi/(n-1), \quad \text{and} \quad \chi = \sum_i (d_i - \xi)^2/(n-1)^2\,.$$

Roughly speaking, the theorem states that if the distance between the degree sequence $\boldsymbol{d}$ and the ξ-regular sequence of the same size is not too large, then the expression in (2.5) is a good approximation for $|\mathcal{G}(\boldsymbol{d})|$. The absolute constant α in Theorem 1.1 is mostly restricted by the ϵ in Theorem 2.6.

THEOREM 2.6 (Liebenau and Wormald [44]). *There exists an absolute constant $\epsilon > 0$ such that for every sequence of degree sequences $\left(\boldsymbol{d}^{(n)}\right)_{n\in\mathbb{N}}$ with ξn even, $\max_{i\in[n]} |d_i^{(n)} - \xi| = o\left(n^{\epsilon}\min\{\xi, n-\xi-1\}^{1/2}\right)$, and $n^2\min\{\mu, 1-\mu\} \to \infty$, it holds that*

$$|\mathcal{G}(\boldsymbol{d})| \sim \bar{w}(\boldsymbol{d}) := \sqrt{2}\exp\left(\frac{1}{4} - \frac{\chi^2}{4\mu^2(1-\mu)^2}\right)\left(\mu^{\mu}(1-\mu)^{(1-\mu)}\right)^{\frac{n(n-1)}{2}}\prod_i \binom{n-1}{d_i^{(n)}}. \tag{2.5}$$

To be precise, there exists a non-negative function $\delta(n)$ with $\delta(n) \to 0$ as $n \to \infty$, so that the relative error in "$\sim$" is bounded above in absolute value by $\delta(n)$ for every such $\left(\boldsymbol{d}^{(n)}\right)_{n\in\mathbb{N}}$.

The growth condition $o\left(n^{\epsilon}\min\{\xi, n-\xi-1\}^{1/2}\right)$ in Theorem 2.6 gives rise to our restrictions on $[\ell_i, u_i]$ in Theorem 1.1. In particular, observe that under the condition that $0 \leq \alpha \leq 1/2$, it holds that $\min\{r, n-r-1\}^{\alpha} = o\left(n^{\epsilon}\min\{\xi, n-\xi-1\}^{1/2}\right)$ with ϵ and ξ as in the statement of Theorem 2.6.

Note that the existence of the asymptotic formula (2.5) suggests a straightforward approach for approximating the number of graphs in $\mathcal{G}(\boldsymbol{\ell},\boldsymbol{u})$: one could sum the formulae for all sequences $\mathbf{d}$ that satisfy $\boldsymbol{\ell} \leq \mathbf{d} \leq \mathbf{u}$ and that have an even sum. Nevertheless, this observation does not imply the existence of a FPRAS for the task of approximate counting. This is because we would need to sum superpolynomially many terms, even if we considered the (weighted) sum over sorted sequences, as long as $\max(u_i - \ell_i) = \omega(1)$.

In what follows, we also rely on the notion of strong stability introduced in [1] (and implicitly already used in [35]). A combinatorial definition of this notion is given below. It essentially states that any graph with a slightly perturbed degree sequence can easily be transformed into a graph with the desired degree sequence by flipping the edges on a short alternating path. An *alternating (u, v)-path in a graph G* is a (possibly non-simple)

[3] Our formulation is in line with the note after Conjecture 1.2 in [44].

edge-disjoint (u, v)-path (in the corresponding complete graph) alternating between edges and non-edges of G, starting with an edge adjacent to u, and ending with a non-edge adjacent to v; recall that a non-edge is an edge contained in the complement of $E(G)$. If $u = v$ we obtain an *alternating cycle*. To facilitate the definition of strong stability, let $\mathcal{G}'(\boldsymbol{d}) = \bigcup_{\boldsymbol{d}'} \mathcal{G}(\boldsymbol{d}')$ with $\boldsymbol{d}'$ ranging over all sequences $\boldsymbol{d}'$ satisfying $\sum_i d'_i = \sum_i d_i$ and $\sum_i |d'_i - d_i| = 2$, i.e., there exist κ, λ such that $d'_\kappa = d_\kappa + 1$, $d'_\lambda = d_\lambda - 1$, and $d'_i = d_i$ otherwise.

Definition 2.7 (Strong stability). *A class $\mathcal{D}$ of degree sequences is k-*strongly stable *if for all $\boldsymbol{d} \in \mathcal{D}$ and all $G \in \mathcal{G}'(\boldsymbol{d})$, there is an alternating (u, v)-path in G of length at most k, where u and v are the unique nodes with $\deg_G(u) = d_u + 1$ and $\deg_G(v) = d_v - 1$. We call $\mathcal{D}$* strongly stable *if there exists a constant $k \in \mathbb{N}$ for which $\mathcal{D}$ is k-strongly stable.*

Proposition 2.8. *Let $0 < \alpha \leq 1/2$ be a constant and assume that $2 \leq r(n) \leq (1-\sigma)n$ for some constant $0 < \sigma < 1$ and $n \in \mathbb{N}$. Then there exists some $n_1 \in \mathbb{N}$ so that the class $\mathcal{F}_{(\alpha,r)}[n_1]$ is 10-strongly stable.*

Proof. Let $n \geq n_1 = \lceil 10/\sigma^2 \rceil$. It is then a matter of simple calculations to verify that the condition $(d_{\max} - d_{\min} + 1)^2 \leq 4 d_{\min}(n - d_{\max} - 1)$ is satisfied for all $\boldsymbol{d} \in \mathcal{F}_{(\alpha,r)}[n_1]$, where $d_{\min}$ and $d_{\max}$ are the minimum and maximum value of $\boldsymbol{d}$, respectively. Sequences satisfying this condition are 10-strongly stable [1, 35]. □

The following two results hold for the class $\mathcal{F}_{(\alpha,r)}[n_1]$ of Proposition 2.8. Lemma 2.9 essentially states that if an edge is present in some graphical realization, then there exists a short alternating cycle to obtain a graphical realization with the same degree sequence not containing that edge. As a result, the subset of realizations in $\mathcal{G}(\boldsymbol{d})$ containing a given edge and the set of realizations not containing it are polynomially related in size.

Lemma 2.9. *Let $\boldsymbol{d} \in \mathcal{F}_{(\alpha,r)}[n_1]$. Suppose that $G \in \mathcal{G}(\boldsymbol{d})$ and let $\{u, v\} \in E(G)$ (resp. $\{u, v\} \notin E(G)$). Then there exists a graph $G' \in \mathcal{G}(\boldsymbol{d})$ with $\{u, v\} \notin E(G')$ (resp. $\{u, v\} \in E(G)$) and $E(G) \triangle E(G')$ is an alternating cycle of length at most 12. Similarly, suppose that $\{u, w\}, \{u, v\} \in E(G)$. Then there exists a graph $G' \in \mathcal{G}(\boldsymbol{d})$ with $\{u, w\} \in E(G')$ and $\{u, v\} \notin E(G')$, and $E(G) \triangle E(G')$ is an alternating cycle of length at most 12.*

Proof. Assume $n \geq n_1 = \lceil 10/\sigma^2 \rceil$ as in the proof of Proposition 2.8. Note that, in all cases below, the degree sequence $\boldsymbol{d}$ itself plays the role of being a perturbed degree sequence in the argument. By inspecting the proof of Proposition 2.8, this is allowed since n here is sufficiently large.

For the first case of the first part (i.e., when $\{u, v\} \in E(G)$), let y be such that $\{y, u\} \notin E(G)$. Such a non-edge is guaranteed to exist, as $n \geq n_1 > 2/\sigma$ and the maximum degree of any node will then be bounded away from $n - 2$. Also note that y has degree at least 2. By Proposition 2.8, we know that there exists some alternating (y, v)-path of length at most 10. Combining this path with the non-edge $\{y, u\}$ and the edge $\{u, v\}$, results in an alternating cycle of length at most 12. Hence, if we flip all the edges on this alternating cycle, we obtain a $G' \in \mathcal{G}(\boldsymbol{d})$ with the desired property.

For the second case of the first part (i.e., when $\{u, v\} \notin E(G)$), we pick a y such that $\{y, u\} \in E(G)$. By Proposition 2.8, we consider some alternating (v, y)-path (of length at most 10). Combining this path with the edge $\{y, u\}$ and the non-edge $\{u, v\}$, we again obtain an alternating cycle of length at most 12. By flipping this cycle, we get a $G' \in \mathcal{G}(\boldsymbol{d})$ with the desired property.

For the second part of the lemma, we make a similar, albeit more complicated, argument. We distinguish two cases and then consider subcases depending on the relative position of the edge $\{u, w\}$ with respect to some alternating path.

Case 1: *there exists* y *such that* $\{y,u\},\{y,v\} \notin E(G)$. Consider such a node y. By Proposition 2.8, there exists an alternating (y,v)-path of length at most 10. A key observation here is that this alternating (y,v)-path might contain the edge $\{u,w\}$. Of course, if this is not true, we proceed like in the first case of the first part above. So assume that the alternating (y,v)-path does contain $\{u,w\}$. If $\{u,w\}$ goes from w to u as we traverse the path from y to v, then by taking the remaining (u,v)-subpath of the alternating path together with the edge $\{u,v\}$ we get an alternating cycle (of length at most 8), that contains the edge $\{u,v\}$ but not the edge $\{u,w\}$. If $\{u,w\}$ goes from u to w as we traverse the path from y to v, then by taking the (y,u)-subpath preceding $\{u,w\}$ on the alternating path together with the edge $\{u,v\}$ and the non-edge $\{y,v\}$ we again get an alternating cycle (of length at most 10), that contains $\{u,v\}$ but not $\{u,w\}$.

In any case, by flipping the edges on the corresponding alternating cycle, we obtain a $G' \in \mathcal{G}(\boldsymbol{d})$ with the desired property.

Case 2: *for every* y *such that* $\{y,u\} \notin E(G)$ *we have* $\{y,v\} \in E(G)$. Consider a node x such that $\{x,v\} \notin E(G)$. Given the assumption of the current case, it must be that $\{x,u\} \in E(G)$. By Proposition 2.8, there exists an alternating (x,u)-path of length at most 10. If the edge $\{u,w\}$ is not contained in this alternating path, then the whole path together with the non-edge $\{x,v\}$ and the edge $\{u,v\}$ is an alternating cycle (of length at most 12) that contains $\{u,v\}$ but not $\{u,w\}$. Now, assume that the alternating (x,u)-path contains $\{u,w\}$. If $\{u,w\}$ goes from u to w as we traverse the path from x to u, then by taking the (x,u)-subpath preceding $\{u,w\}$ on the alternating path together with the edge $\{u,v\}$ and the non-edge $\{x,v\}$ we get an alternating cycle (of length at most 8), that contains $\{u,v\}$ but not $\{u,w\}$. Finally, if $\{u,w\}$ goes from w to u as we traverse the path from x to u, then by taking the remaining (u,u)-subpath of the alternating path together with the edges $\{u,v\},\{x,u\}$ and the non-edge $\{x,v\}$, we again get an alternating cycle (of length at most 10), that contains the edge $\{u,v\}$ but not the edge $\{u,w\}$.

In all subcases, by flipping the edges on the corresponding alternating cycle, we obtain a $G' \in \mathcal{G}(\boldsymbol{d})$ with the desired property. ∎

Furthermore, the *switch Markov chain* is rapidly mixing for the class $\mathcal{F}_{(\alpha,r)}[n_1]$. This follows directly from [1] where it is shown that the switch Markov chain is rapidly mixing for all strongly stable classes of degree sequences. In particular, we will use the following result.

COROLLARY 2.10 (Follows from [1]). *Let* $q(n) \geq 2$ *be a given polynomial and consider the* lazy switch Markov chain $\mathcal{M} = (\mathcal{G}(\boldsymbol{d}), P_{\mathcal{G}(\boldsymbol{d})})$ *for some* $\boldsymbol{d} \in \mathcal{F}_{(\alpha,r)}[n_1]$ *that proceeds as follows: For a given* $G \in \mathcal{G}(\boldsymbol{d})$

- *with probability* $1 - 1/q(n)$ *do nothing, and*
- *with probability* $1/q(n)$*, try to perform a switch operation.*

Then there exists a polynomial $p(n)$*, such that for any* $\boldsymbol{d} \in \mathcal{F}_{(\alpha,r)}[n_1]$ *we have* $\text{Gap}(P_{\mathcal{G}(\boldsymbol{d})}) \geq 1/p(n)$.[4]

3. Proof approach overview.

In this section we give a high-level overview of the proofs of Theorems 1.2 and 1.3. The idea is to decompose the degree interval Markov chain *twice*, using the addition/deletion and switch graph operations in Figure 1. Note that the second decomposition step suffices for proving Theorem 1.2, but both of them are needed in order to prove Theorem 1.3.

We first decompose $\mathcal{G}(\boldsymbol{\ell},\boldsymbol{u})$ based on the addition/deletion operation. Every part

[4] Note that $P_{\mathcal{G}(\boldsymbol{d})}$ depends on $r(n)$.

of the decomposition corresponds to a set $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ containing all graphs respecting the degree intervals $[\boldsymbol{\ell}, \boldsymbol{u}]$ and having exactly m edges, for some m. That is, there is a one-to-one correspondence between the possible values of m, and the parts of the decomposition. The Markov chain decomposition result of Theorem 2.4 tells us that if the switch hinge-flip chain is rapidly mixing for every m, and if it relatively "easy" to move between the different parts $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ by means of additions/deletions, then the degree interval chain is rapidly mixing on $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$. In the second step we carry out a similar decomposition, but now based on the hinge flip operation. That is, for a given m we decompose $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ in the sets $\mathcal{G}(\boldsymbol{d})$ for all sequences $\boldsymbol{d}$ which satisfy the interval constraints, and whose degrees sum up to $2m$. If the switch chain is rapidly mixing on every $\mathcal{G}(\boldsymbol{d})$, and we can move "easily" between the sets $\mathcal{G}(\boldsymbol{d})$ using hinge flip operations, then the switch hinge-flip Markov chain on $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ is also rapidly mixing. We continue with a formalization of these statements.

Let $\mathcal{T} = \{m_1, \dots, m_2\}$, where m_1 and m_2 are the minimum and maximum number of edges, respectively, that any $G \in \mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ could have; e.g., $m_1 = \lceil \frac{1}{2} \sum_i \ell_i \rceil$ and $m_2 = \lfloor \frac{1}{2} \sum_i u_i \rfloor$. It is not hard to see that these two edge-counts are indeed achievable for the intervals we consider in Theorem 1.1; this follows for example from the fact that the asymptotic formula in Theorem 2.6 is nonzero in those cases.

First we partition $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ into disjoint sets $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ for $m \in \mathcal{T}$. Recall that $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u}) = \left\{G \in \mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u}) : \sum_i d_i(G) = 2m\right\}$. The restriction Markov chains $\mathcal{M}_{\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})}$ are essentially given by restricting the original chain to only perform switch and hinge flip operations that respect the degree intervals on graphs with precisely m edges. Applying Theorem 2.4—with β and γ to be determined later—we get

$$\text{Gap}(P) \geq \beta\gamma \cdot \text{Gap}(P_{\mathcal{T}}) \cdot \min_{m \in \mathcal{T}} \text{Gap}(P_{\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})}), \tag{3.1}$$

where $P_{\mathcal{T}}$ is the transition matrix of the Metropolis-Hastings projection chain on $\mathcal{T}$, and P the transition matrix of the degree interval Markov chain. The goal will be to show that β and γ, as well as all the spectral gaps, can be lower bounded by an inverse polynomial function of the form $1/p(n)$ for some polynomial $p(n)$. This means that $\text{Gap}(P)$ is lower bounded by an inverse polynomial as well, which is equivalent to showing that the degree interval Markov chain is rapidly mixing (see Section 2.2).

Next we partition each $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ further into sets $\mathcal{G}(\boldsymbol{d})$ for sequences $\boldsymbol{d}$ in $\mathcal{D}_m(\boldsymbol{\ell}, \boldsymbol{u}) = \left\{\boldsymbol{d} : \sum_i d_i = 2m \text{ and } \boldsymbol{\ell} \leq \boldsymbol{d} \leq \boldsymbol{u}\right\}$. For simplicity, we drop the arguments and write $\mathcal{D}_m$ instead of $\mathcal{D}_m(\boldsymbol{\ell}, \boldsymbol{u})$. For this part of the decomposition we get a Metropolis-Hastings projection chain on the set $\mathcal{D}_m$. The restriction chains on $\mathcal{M}_{\mathcal{G}(\boldsymbol{d})}$ are the chains in which we essentially only apply switch operations on all graphs with degree sequence $\boldsymbol{d}$. This is precisely the switch Markov chain with some polynomially bounded holding probability (as defined in Corollary 2.10). Using one more time Theorem 2.4 for each m—again, with β_m and γ_m to be determined later—we have

$$\text{Gap}(P_{\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})}) \geq \beta_m\gamma_m \cdot \text{Gap}(P_{\mathcal{D}_m}) \cdot \min_{\boldsymbol{d} \in \mathcal{D}_m} \text{Gap}(P_{\mathcal{G}(\boldsymbol{d})}), \tag{3.2}$$

where $P_{\mathcal{D}_m}$ is the transition matrix of the Metropolis-Hastings chain on $\mathcal{D}_m$. This time, in order to show that the switch-hinge flip Markov chain is rapidly mixing, we need to bound γ_m, β_m, and all the spectral gaps by an inverse polynomial function.

Combining (3.1) and (3.2) we now get

$$\text{Gap}(P) \geq \beta\gamma \cdot \text{Gap}(P_{\mathcal{T}}) \cdot \min_{m \in \mathcal{T}} \left\{\beta_m\gamma_m \cdot \text{Gap}(P_{\mathcal{D}_m}) \cdot \min_{\boldsymbol{d} \in \mathcal{D}_m} \text{Gap}(P_{\mathcal{G}(\boldsymbol{d})})\right\}, \tag{3.3}$$

and in order to show that the degree interval Markov chain is rapidly mixing, we need to show that β, γ, all β_m and γ_m, and all spectral gaps can be lower bounded by $1/q(n)$ for some polynomial $q(n)$.

While this is what we are going to do for Theorem 1.3, recall that for Theorem 1.2 (directly) and Theorem 1.1 (through Theorem 1.2 and the reductions in Appendix B), we only show that the switch-hinge flip Markov chain is rapidly mixing. In that case, it suffices to show that β_m, γ_m, and the spectral gaps involved in (3.2) are polynomially bounded for any given m (and the polynomial bound is independent of m), i.e., we only need to globally consider the second decomposition step. A polynomial lower bound on β and each one of the β_m follows by the very definition of the degree interval Markov chain (see also the discussion after its definition). In order to bound γ and γ_m, for all m, we use Lemma 2.9. Roughly speaking, we need to show that we can move rather easily between realizations of two degree sequences $\boldsymbol{d}$ and $\boldsymbol{d}'$, with $\sum_i |d_i - d_i'| = 2$. The high-level idea for γ_m is to show that it is either directly possible to perform a hinge flip in order to transition from a graph G with degree sequence $\boldsymbol{d}$ to some G' with degree sequence $\boldsymbol{d}'$, or that G is not too far away from some other graph H with the same degree sequence $\boldsymbol{d}$ from which it is possible to directly move to some G' with degree sequence $\boldsymbol{d}'$ via a hinge flip. We take an analogous approach for bounding γ but in terms of addition/deletion operations rather than hinge flips.

The gaps of the chains $\mathcal{M}_{\mathcal{G}(\boldsymbol{d})}$ are globally bounded because of known rapid mixing results for the switch Markov chain [1] (Corollary 2.10). Therefore, in order to show Theorem 1.2, it remains to bound $\text{Gap}(P_{\mathcal{D}_m})$, which we do in Section 5.1; an outline is given in Section 3.1 below. For Theorem 1.3, we additionally need to bound $\text{Gap}(P_{\mathcal{T}})$, which we do in Section 5.2; a brief outline is given in Section 3.2.

3.1. Proving Theorem 1.2. The main technical challenge of Theorem 1.2 lies in proving that the resulting Metropolis-Hastings projection chain on $\mathcal{D}_m$ is rapidly mixing, i.e., that $\text{Gap}(P_{\mathcal{D}_m})$ can be polynomially bounded. We sometimes refer to this chain as the *hinge flip projection chain*. Note that for $\boldsymbol{d}, \boldsymbol{d}' \in \mathcal{D}_m$, with $||\boldsymbol{d} - \boldsymbol{d}'||_1 = 2$, it follows from (2.4) that

$$P_{\text{MH}}(\boldsymbol{d}, \boldsymbol{d}') \geq \frac{1}{2n^2} \min\left\{1, \frac{|\mathcal{G}(\boldsymbol{d}')|}{|\mathcal{G}(\boldsymbol{d})|}\right\},$$

by taking the obvious upper bound $\Delta \leq n^2$ in (2.4). So, intuitively, whether or not the hinge flip projection chain is rapidly mixing depends on the quantities $|\mathcal{G}(\boldsymbol{d})|$ for $\boldsymbol{d} \in \mathcal{D}_m$. To this end, we first argue, using a comparison argument, that if suffices to show that the load-exchange Markov chain on $\mathcal{D}_m$, i.e., the Markov chain that allows us to move between degree sequences by adjusting the degree of two nodes by 1 (while keeping the degree sums fixed), is rapidly mixing for the weights $w(\boldsymbol{d}) = |\mathcal{G}(\boldsymbol{d})|$. (It is not hard to see that $\mathcal{D}_m$ is in fact an M-convex set, as it can be seen as the collection of bases of a discrete polymatroid [33]. The so-called basis-exchange property for discrete polymatroids corresponds with (2.1) for indicator functions.) A Markov chain comparison argument, very informally speaking, proceeds by showing that if one Markov chain is rapidly mixing, and a second chain is very close to it (in terms of similar stationary distribution and transition probabilities), then the second chain is also rapidly mixing. In our setting, the comparison is based on the fact that both chains have the same stationary distribution π with $\pi(\boldsymbol{d}) \propto w(\boldsymbol{d})$, and the fact that their transition probabilities are polynomially related for the degree sequences that we are interested in (using Corollary A.2).

In order to show that the load-exchange Markov chain on $\mathcal{D}_m$ is rapidly mixing, we would like to use Corollary 2.5, which states that the load-exchange Markov chain is rapidly mixing if a polynomial identified with its stationary distribution satisfies the property of strong log-concavity (SLC). To be precise, we could apply Corollary 2.5 if, for given $\boldsymbol{\ell}, \boldsymbol{u}$ and m, the polynomial

$$h(x) = \sum_{\boldsymbol{d} \in \mathcal{D}_m} w(\boldsymbol{d}) \cdot x^{\boldsymbol{d}} = \sum_{\boldsymbol{d} \in \mathcal{D}_m} |\mathcal{G}(\boldsymbol{d})| \cdot x^{\boldsymbol{d}}$$

was SLC. This seems hard to prove (and might not be true in general). However, it turns out that when replacing the weights $w(\boldsymbol{d})$ by simplified versions, say $\bar{z}(\boldsymbol{d})$, of the approximations $\bar{w}(\boldsymbol{d})$ from the asymptotic formula (2.5) of Liebenau and Wormald [44], the resulting polynomial

$$\bar{f}(x) = \sum_{\boldsymbol{d} \in \mathcal{D}_m} \bar{z}(\boldsymbol{d}) \cdot x^{\boldsymbol{d}} \tag{3.4}$$

is in fact SLC, when considering the degree interval instances of Theorem 1.1.[5] We show this fact in Theorem 4.2 in Section 4 by observing that the polynomial in (3.4) is of the form (2.2) in Proposition 2.3, which is a general sufficient condition for a polynomial to be SLC [11].

The above implies that if we run the load-exchange Markov chain with the approximations $\bar{z}(\boldsymbol{d})$, it is in fact rapidly mixing with stationary distribution $\bar{\pi}$ given by $\bar{\pi}(\boldsymbol{d}) \propto \bar{z}(\boldsymbol{d})$. Now, the $\bar{z}(\boldsymbol{d})$ have the property that for some n_0 sufficiently large, it holds that for all $n \geq n_0$ and $\boldsymbol{d} \in \mathcal{D}_m$, where $\boldsymbol{d}$ is of length n,

$$\frac{1}{2}|\mathcal{G}(\boldsymbol{d})| \leq \bar{z}(\boldsymbol{d}) \leq 2e^{(19/\sigma)^2}|\mathcal{G}(\boldsymbol{d})|\,.$$

This also implies that

$$\frac{1}{2}e^{-(19/\sigma)^2}\pi(\boldsymbol{d}) \leq \bar{\pi}(\boldsymbol{d}) \leq 2e^{(19/\sigma)^2}\pi(\boldsymbol{d})\,.$$

One can then again use a Markov chain comparison argument to argue that the load-exchange Markov chain based on the original weights $w(\boldsymbol{d})$ is also rapidly mixing (by applying Corollary A.2). This in turn implies that the hinge flip projection chain is also rapidly mixing, which is what we wanted to show.

General framework. The approach described above for showing rapid mixing of the switch-hinge flip Markov chain might be applicable to other classes of degree interval instances. Informally speaking, the essential things that are needed are the following two things:

1. The degree sequences satisfying the interval constraints are strongly stable (see Def. 2.7).
2. The weights $|\mathcal{G}(\boldsymbol{d})|$ "approximately" give rise to an SLC polynomial.

The requirement of strong stability in the first point is needed for various reasons. First of all, it is a sufficient condition for the switch Markov chain (i.e., the restrictions chains in our decomposition) to be rapidly mixing [1]. Secondly, we rely on it when bounding the parameter γ in the Martin-Randall decomposition theorem (Theorem 2.4). Thirdly, strong stability is sufficient to argue that the transition probabilities of the load-exchange Markov chain, and the Metropolis-Hastings projection chain, are polynomially

[5] We remark at this point, that, although this is true for the regime considered in Theorem 1.1, this does not seem to be true for the general range in Theorem 2.6.

related (so that we can use a Markov chain comparison argument to compare their mixing times).

For the second point, even if the weights $|\mathcal{G}(\boldsymbol{d})|$ do not give rise to an SLC polynomial, one may still make things work. It suffices to find values $z(\boldsymbol{d})$ and polynomials q_1 and q_2 such that

$$\frac{1}{q_1(n)}|\mathcal{G}(\boldsymbol{d})| \leq z(\boldsymbol{d}) \leq q_2(n)|\mathcal{G}(\boldsymbol{d})|\,,$$

and for which

$$\bar{f}(x) = \sum_{\boldsymbol{d}\in\mathcal{D}_m} z(\boldsymbol{d}) \cdot x^{\boldsymbol{d}}$$

is SLC.

3.2. Proving Theorem 1.3. In order to prove Theorem 1.3, we additionally need to show that the projection chain on $\mathcal{T}$ is also rapidly mixing, i.e., that $\mathrm{Gap}(\mathcal{T})$ is polynomially bounded. In other words, we consider the Metropolis-Hastings projection Markov chain with state space $\{a,\ldots,b\}$, where $a = \lceil \frac{1}{2}\sum_i \ell_i \rceil$ and $b = \lfloor \frac{1}{2}\sum_i u_i \rfloor$, and $\pi(m) \propto |\mathcal{G}_m(\boldsymbol{\ell},\boldsymbol{u})|$ for $m \in \{a,\ldots,b\}$. This chain will sometimes be referred to as the *addition/deletion projection chain.* A sufficient condition for this Markov chain to be rapidly mixing is that the sequence $(w_m)_{m=a,\ldots,b}$ given by $w_m = |\mathcal{G}_m(\boldsymbol{\ell},\boldsymbol{u})|$ is log-concave, meaning that for every m $w_{m+1}w_{m-1} \leq w_m^2$. We show that this sequence is log-concave when the intervals have size at most one (corresponding to the statement of Theorem 1.3) by using a variation on an argument of Jerrum and Sinclair [36].

Remark 3.1. One might wonder if the theory of strongly log-concave polynomials can also be used to prove rapid mixing for degree intervals beyond size one. For example, one might consider a polynomial of the form

$$g(x_1,\ldots,x_n,y) = \sum_{m=m_1}^{m_2}\sum_{\boldsymbol{d}\in\mathcal{D}_m}\binom{n-1}{2m_2-2m}\bar{z}(\boldsymbol{d})\cdot y^{2m_2-2m}x^{\boldsymbol{d}}\,,$$

where m_1 and m_2 are the minimum and maximum number of edges that any graph in $\mathcal{D}(\ell,u)$ can have, respectively. This is then a $2m_2$-homogeneous polynomial. The problem that now occurs though, is that the domain of this polynomial, indexed by the tuples $(d_1,\ldots,d_n,2m_2-2m)$, can be shown not to be an M-convex set and, thus, g cannot be SLC.

4. SLC property in a restricted range of the Liebenau-Wormald result. Throughout this section, we consider m and n as fixed. Recall that for a given degree sequence $\boldsymbol{d} = (d_1,\ldots,d_n)$ we defined $\xi = \xi(n,m) = \sum_i d_i/n = 2m/n$, $\mu = \mu(n,m) = \xi/(n-1) = 2m/(n(n-1))$ and $\chi(\boldsymbol{d}) = \sum_i (d_i-\xi)^2/(n-1)^2$. Furthermore, in (2.5) we defined

$$\bar{w}(\boldsymbol{d}) = \sqrt{2}\exp\left(\frac{1}{4}-s(\boldsymbol{d})^2\right)\left(\mu^{\mu}(1-\mu)^{(1-\mu)}\right)^{n(n-1)/2}\prod_i\binom{n-1}{d_i}\,, \tag{4.1}$$

which is approximately the number of graphs with degree sequence $\boldsymbol{d}$ in case it is near-regular. Here we have

$$s(\boldsymbol{d}) = \frac{\chi(\boldsymbol{d})}{2\mu(1-\mu)}\,.$$

Ideally, we would like show that the weights $\bar{w}(\boldsymbol{d})$ give rise to an SLC polynomial,

$$\bar{h}(x) = \sum_{\boldsymbol{d}\in\mathcal{D}_m}\bar{w}(\boldsymbol{d})\cdot x^{\boldsymbol{d}}\,.$$

By inspecting the weights and the conditions of Proposition 2.3, it is not hard to see that it would be sufficient to argue that the function $1/4 - s(\boldsymbol{d})^2$ is M-convex. (The conditions of Proposition 2.3 are satisfied, note that always $0 \leq d_i \leq n-1$ and that for a degree sequence $\boldsymbol{d} \in \mathcal{D}_m$, it holds that $\sum_i d_i = 2m$, meaning the polynomial is $2m$-homogeneous.) Unfortunately, it turns out that this is not the case. Instead, we simply show that in the regime of Theorem 1.1, it holds that $s(\boldsymbol{d}) = O(1)$, and so, essentially, we can ignore the contribution $\exp(-s(\boldsymbol{d})^2)$ in (4.1) at the expense of a slightly worse bound on the mixing time. The resulting approximation formula is easily seen to be SLC, which intuitively follows from a discrete form of log-concavity of the binomial coefficients; see Theorem 4.2.

LEMMA 4.1. *Under the conditions on $[\boldsymbol{\ell}, \boldsymbol{u}]$ as in Theorem 1.1, with $0 < \sigma < 1$ and $2 \leq r = r(n) \leq (1-\sigma)n$, if n is large enough, then for any $\boldsymbol{\ell} \leq \boldsymbol{d} \leq \boldsymbol{u}$ it holds that*

$$0 \leq s(\boldsymbol{d}) \leq \frac{18n}{\sigma r^{1-2\alpha}(n-1)} \leq \frac{19}{\sigma}\,. \tag{4.2}$$

Proof. By the definition of $\chi(\boldsymbol{d})$, $s(\boldsymbol{d}) \geq 0$ always holds. To see the upper bound, let $n_2 = \max\left\{\lceil 10/\sigma^2 \rceil, \lceil 18/\sigma \rceil\right\}$ and note that the quantity $s(\boldsymbol{d})$ can be rewritten as

$$s(\boldsymbol{d}) = \frac{\chi(\boldsymbol{d})}{2\mu(1-\mu)} = \frac{n^2(n-1)^2}{(n-1)^2} \frac{\sum_i (d_i - \xi)^2}{2 \cdot 2m(n(n-1) - 2m)}\,, \tag{4.3}$$

where $2m = \sum_i d_i$. Note that $\sum_i (d_i - \xi)^2 \leq n(2\min\{r, n-r-1\}^\alpha)^2 = 4n\min\{r, n-r-1\}^{2\alpha}$. Moreover, we can bound m using the simple facts that $r - \min\{r, n-r-1\}^\alpha \geq r/4$ and $n^\alpha \leq \sigma n/2$ for $n \geq n_2$. The latter implies that

$$r + \min\{r, n-r-1\}^\alpha \leq (1-\sigma)n + (1-\sigma)n^\alpha \leq (1-\sigma)n + \sigma n/2 = (1-\sigma/2)n\,.$$

So, we have

$$n\frac{r}{4} \leq 2m \leq n\left(1 - \frac{\sigma}{2}\right)n\,,$$

and therefore,

$$2m(n(n-1) - 2m) \geq n\frac{r}{4}\left(1 - \left(1 - \frac{\sigma}{2}\right)\frac{n}{n-1}\right)n(n-1) \geq \frac{r\sigma n^2(n-1)}{9}\,, \tag{4.4}$$

where the last inequality holds because $\frac{n\sigma/2-1}{n-1} \geq \frac{4\sigma}{9}$ for $n \geq n_2$. By combining (4.3) and (4.4), we then get

$$s(\boldsymbol{d}) \leq \frac{n^2 \cdot 4n\min\{r, n-r-1\}^{2\alpha} \cdot 9}{2r\sigma n^2(n-1)} \leq \frac{18n}{\sigma r^{1-2\alpha}(n-1)}\,,$$

which completes the second inequality. The final inequality holds because $r \geq 2$ and $n/(n-1) \leq \frac{19}{18}$ for $n \geq n_2$. ∎

We next summarize the main result of this section, and give the remaining small technical steps of its proof. In a nutshell, it states that a simplified version of the Liebenau-Wormald formula which is within a constant factor from the original in (2.5) is approximately SLC in the regime of Theorem 1.1. Recall that $w(\boldsymbol{d}) = |\mathcal{G}(\boldsymbol{d})|$.

THEOREM 4.2. *For given $n, m \in \mathbb{N}$, $\boldsymbol{\ell}, u \in \mathbb{N}^n$ with $\boldsymbol{\ell} \leq \boldsymbol{u}$, and degree sequence $\boldsymbol{d}$ with $\sum_i d_i = 2m$ and $\boldsymbol{\ell} \leq \boldsymbol{d} \leq \boldsymbol{u}$, let*

$$\bar{z}(\boldsymbol{d}) = \sqrt{2}e^{\frac{1}{4}}\left(\mu^\mu(1-\mu)^{(1-\mu)}\right)^{n(n-1)/2} \prod_i \binom{n-1}{d_i}\,. \tag{4.5}$$

The resulting $2m$-homogeneous polynomial

$$\bar{f}(x) = \sum_{\boldsymbol{d}\in\mathcal{D}_m} \bar{z}(\boldsymbol{d}) \cdot x^{\boldsymbol{d}}$$

is SLC.

Furthermore, there exists an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and $m \geq n$, if the degree interval $[\boldsymbol{\ell}, \boldsymbol{u}]$ satisfies the conditions of Theorem 1.1, then

$$\frac{1}{2} w(\boldsymbol{d}) \leq \bar{w}(\boldsymbol{d}) \leq \bar{z}(\boldsymbol{d}) \leq e^{(19/\sigma)^2} \bar{w}(\boldsymbol{d}) \leq 2e^{(19/\sigma)^2} w(\boldsymbol{d})\,, \tag{4.6}$$

for every $\boldsymbol{\ell} \leq \boldsymbol{d} \leq \boldsymbol{u}$.

Proof. We first note that the factors $\sqrt{2}$ and $\left(\mu^\mu (1-\mu)^{(1-\mu)}\right)^{n(n-1)/2}$ can all be seen as non-negative scalars as n and m are given. This means, by Proposition 2.2, that it suffices to show that the polynomial with coefficients

$$e^{\frac{1}{4}} \prod_i \binom{n-1}{d_i}$$

is SLC.

Comparing this to the second polynomial in Proposition 2.3, it follows that we can simply choose ν to be the constant function $\nu(\boldsymbol{d}) = -\frac{1}{4}$ on its effective domain $\mathcal{D}_m$. (As mentioned earlier, intuitively it is SLC because the binomial coefficients satisfy a discrete form of log-concavity.) The inequalities

$$\bar{w}(\boldsymbol{d}) \leq \bar{z}(\boldsymbol{d}) \leq e^{(19/\sigma)^2} \bar{w}(\boldsymbol{d})$$

in (4.6) follow directly from Theorem 2.6 and Lemma 4.1. The outer two inequalities hold because for n sufficiently large $\bar{w}(d)$ approximates $\bar{w}(\boldsymbol{d})$ up to a multiplicative factor that converges to 1 (and, thus, will be at most 2 for all n beyond some $n_0 \in \mathbb{N}$). ∎

5. Decomposition of the degree interval Markov chain. In this section we give the missing details regarding the decomposition steps as outlined in Section 3.

5.1. Bounding β_m, γ_m and $\mathrm{Gap}(P_{\mathcal{D}_m})$ of inequality (3.2). Throughout this section we assume that some $m \in \{m_1, \ldots, m_2\}$ is fixed. Moreover, recall that we consider degree intervals of the form $[d_i, d_i + 1]$, or $[d_i, d_i]$, for $i \in [n]$. It is not hard to see that $\beta_m \geq (6n^4)^{-1}$. This rough polynomial bound follows directly from the transition probabilities of the degree interval Markov chain (see Section 2.2.3).

We first lower bound the γ_m in Lemma 5.1 below. By the definition of the hinge flip operation we have that for any $\boldsymbol{d}, \boldsymbol{d}' \in \mathcal{D}_m$, there is a strictly positive transition probability between $\boldsymbol{d}$ and $\boldsymbol{d}'$ if and only if $\sum_i |d_i - d_i'| = 2$.

The proof of Lemma 5.1 follows from Lemma 2.9, where it is shown that for a graph with a given degree sequence, we can always find a graph with a slightly perturbed degree sequence that is close to the former in terms of symmetric difference (when the original sequences satisfies strong stability).

LEMMA 5.1. *There exists a polynomial $q_1(n)$ such that, for any feasible m and for all $\boldsymbol{d}, \boldsymbol{d}' \in \mathcal{D}_m$ with $\sum_i |d_i - d_i'| = 2$, we have $\pi_{\mathcal{D}_m}\left(\partial_{\boldsymbol{d}}\left(\mathcal{G}\left(\boldsymbol{d}'\right)\right)\right) \geq \frac{1}{q_1(n)} \pi_{\mathcal{D}_m}\left(\mathcal{G}\left(\boldsymbol{d}'\right)\right).$*

Proof. Again assume that $n \geq n_1 = \lceil 10/\sigma^2 \rceil$. Let a and b be the unique nodes such that $d_a' = d_a + 1$ and $d_b' = d_b - 1$; note that the uniqueness of a, b follows from the condition $\sum_i |d_i - d_i'| = 2$. Let

$$\mathcal{H} = \{G \in \mathcal{G}(\boldsymbol{d}) : \exists c \in [n] \text{ such that } \{b, c\} \in E(G), \{a, c\} \notin E(G)\}\,,$$

and note that it has the property

$$|\partial_{\boldsymbol{d}}(\mathcal{G}(\boldsymbol{d}'))| \geq \frac{1}{n}|\mathcal{H}|\,. \tag{5.1}$$

To see this, note that for a given $G \in \mathcal{H}$, we can perform the hinge flip that removes the edge $\{b,c\}$ and adds the edge $\{a,c\}$ to obtain an element in $\mathcal{G}(\boldsymbol{d}')$. Moreover, there can be at most n graphs $G \in \mathcal{H}$ that map onto a given $G' \in \partial_{\boldsymbol{d}}(\mathcal{G}(\boldsymbol{d}'))$, as there are at most n choices for c.

Moreover, using the second part of Lemma 2.9, we show that

$$|\mathcal{H}| \geq \frac{1}{n^{12}}|\mathcal{G}(\boldsymbol{d})|\,. \tag{5.2}$$

To see this, note that for any $G \in \mathcal{G}(\boldsymbol{d})$, we have $d_b = d'_b + 1 \geq 0$ which implies that b has at least one neighbor c in G. Now, if $\{a,c\} \notin E(G)$ we obtain an element in $\mathcal{H}$; otherwise, by Lemma 2.9, we can find a graph G' close to G (in the sense that $|E(G) \triangle E(G')| \leq 12$) for which $\{a,c\} \notin E(G)$ while still $\{b,c\} \in E(G)$. As there are at most n^{12} graphs $G \in \mathcal{G}(\boldsymbol{d})$ that map to the same $G' \in \mathcal{H}$, the inequality (5.2) follows. Moreover, we also have $n^{10}|\mathcal{G}(\boldsymbol{d})| \geq |\mathcal{G}(\boldsymbol{d}')|$ which follows directly from Definition 2.7 and Proposition 2.8.

Combining the last observation with (5.1) and (5.2) then yields

$$|\partial_{\boldsymbol{d}}(\mathcal{G}(\boldsymbol{d}'))| \geq \frac{1}{q_1(n)}|\mathcal{G}(\boldsymbol{d}')|\,,$$

for $q_1(n) = n^{23}$. Dividing both sides by $\sum_{\boldsymbol{d}\in\mathcal{D}_m}|\mathcal{G}(\boldsymbol{d})|$, then gives the desired result. ◻

It remains to bound $\text{Gap}(P_{\mathcal{D}_m})$. As explained in Section 3.1, the first step is to carry out a comparison argument with the load-exchange Markov chain with weights $w(\boldsymbol{d}) = |\mathcal{G}(\boldsymbol{d})|$ (so that it will be sufficient to study the mixing time of the latter). Remember that both the hinge flip projection chain, as well as the load-exchange chain have stationary distribution $\pi(\boldsymbol{d}) \propto w(\boldsymbol{d})$.

In what follows we write $\mathcal{M}_{\mathcal{D}_m} = (\mathcal{D}_m, P)$ for the (Metropolis-Hastings) hinge flip projection chain, and $\mathcal{M}'_{\mathcal{D}_m} = (\mathcal{D}_m, P')$ for the load-exchange chain on $\mathcal{D}_m$.

Lemma 5.2. *There exists a polynomial $p(n)$ such that*

$$p(n)\,\text{Gap}(P_{\mathcal{D}_m}) \geq \text{Gap}(P'_{\mathcal{D}_m})\,.$$

for any $m \in \{m_1, \dots, m_2\}$.

Proof. It suffices to show that there exists polynomials p_1 and p_2 such that, whenever $\boldsymbol{d}, \boldsymbol{f} \in \mathcal{D}_m$ satisfy $||\boldsymbol{d} - \boldsymbol{f}||_1 = 2$, then

$$\frac{1}{p_1(n)} \leq P(\boldsymbol{d},\boldsymbol{f}), P'(\boldsymbol{d},\boldsymbol{f}) \leq \frac{1}{p_2(n)}\,. \tag{5.3}$$

This then implies that the transition probabilities $P(\boldsymbol{d},\boldsymbol{f})$ and $P'(\boldsymbol{d},\boldsymbol{f})$ are themselves polynomially related, i.e.,

$$\frac{p_2(n)}{p_1(n)} \leq \frac{P(\boldsymbol{d},\boldsymbol{f})}{P'(\boldsymbol{d},\boldsymbol{f})} \leq \frac{p_1(n)}{p_2(n)}\,.$$

In turn, this implies the existence of the desired polynomial $p(n)$ as both chains have the same stationary distribution and therefore their spectral gaps are polynomially related; see (2.3).

The existence of the polynomials in (5.3) follows from the fact that all near-regular degree sequences are strongly stable, which we will illustrate next: First of all, because of strong stability, we can always find polynomials $q_1(n)$ and $q_2(n)$ such that

$$\frac{1}{q_1(n)} \leq \frac{|\mathcal{G}(\boldsymbol{d}')|}{|\mathcal{G}(\boldsymbol{d})|} \leq \frac{1}{q_2(n)} \tag{5.4}$$

meaning that we find the desired bounds for $P(\boldsymbol{d}, \boldsymbol{f})$, i.e, for the transition probabilities of the Metropolis-Hastings hinge flip projection chain. Furthermore, in the load-exchange Markov chain we pick (in the second step) a new degree sequence $\boldsymbol{d}'$ proportional to $w(\boldsymbol{d}')$ over all possible choices of $\boldsymbol{d}'$ with $||\boldsymbol{d} - \boldsymbol{d}'||_1 = 2$ that respect the degree interval bounds. For a given $\boldsymbol{d}$, let $N(\boldsymbol{d})$ be the set of all such sequences $\boldsymbol{d}'$. Then the probability of transitioning to $\boldsymbol{d}'$ is (up to an additional polynomial factor because of the first step of the load-exchange Markov chain) equal to

$$\frac{|\mathcal{G}(\boldsymbol{d}')|}{\sum_{\boldsymbol{f} \in N(\boldsymbol{d})} |\mathcal{G}(\boldsymbol{f})|},$$

which can again be upper and lower bounded by a polynomial because of strong stability, as in (5.4), in combination with the fact that $|N(\boldsymbol{d})| \leq n^2$. ∎

Lemma 5.2 implies that we may focus on bounding $\text{Gap}(P'_{\mathcal{D}_m})$. Now, by the arguments given in Section 3.1 in combination with another simple comparison argument using (A.2) and Theorem 4.2, it suffices to bound $\text{Gap}(P''_{\mathcal{D}_m})$ where P'' is the transition matrix of the hinge flip Markov chain in which we replace the weights $w(\boldsymbol{d})$ by the approximations $\bar{z}(\boldsymbol{d})$ as in (4.5).

In Section 4, we showed that the polynomial in (4.6) is in fact SLC, so then Corollary 2.5 implies that the modified log-Sobolev constant of this chain can be lower bounded by a polynomial, which implies the same for the spectral gap by (A.1). This completes this section, and shows in particular that the switch-hinge flip Markov chain is rapidly mixing, which in turn completes the proof of Theorem 1.2.

5.2. Bounding β, γ and $\text{Gap}(P_\mathcal{T})$. Recall that $\mathcal{M}_\mathcal{T}$ is the Metropolis-Hastings chain on the index set $\mathcal{T} = \{m_1, \ldots, m_2\}$. For simplicity, we use $w_m = |\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$ to denote the number of feasible graphical realizations with m edges. Note that for any $m \in \mathcal{T}$ we have $\pi_\mathcal{T}(m) = w_m / \sum_{i \in \mathcal{T}} w_i$, and that $P_\mathcal{T}(m, m') > 0$ if and only if $|m - m'| \leq 1$. From the definition of the degree interval Markov chain, it immediately follows that $\beta \geq 1/q(n)$ for some polynomial $q(n)$. We lower bound γ in the following lemma following the same approach as for Lemma 5.1.

LEMMA 5.3. *There exists a polynomial $q_2(n)$ such that, for all $m, m' \in \mathcal{T}$ with $|m - m'| = 1$, we have $\pi_\mathcal{T}(\partial_m(\mathcal{G}_{m'})) \geq \frac{1}{q_2(n)} \pi_\mathcal{T}(\mathcal{G}_{m'})$.*

Proof. Assume that $m' = m + 1$ (the case $m' = m - 1$ is analogous). Let $G \in \mathcal{G}_{\boldsymbol{d}}$ for some $\boldsymbol{d} \in \mathcal{D}_m$. Note that $m' \geq m_1 + 1 > m_1$, which implies that there are nodes i and j whose degrees in G are not equal to the upper bound of their degree interval. Note that the set

$$\mathcal{H} = \{G \in \mathcal{G}_m : \exists i, j \in [n] \text{ with } d_i(G) < u_i, d_j(G) < u_j \text{ and } \{i, j\} \notin E(G)\}$$

has the property that

$$|\partial_m(\mathcal{G}_{m'})| \geq \frac{1}{m+1} |\mathcal{H}|. \tag{5.5}$$

In order to see this, note that for any $G \in \mathcal{H}$ we can add the edge $\{i, j\}$ (recall that these nodes depend on the choice of G) to obtain an element in $\mathcal{G}_{m'}$. On the other hand,there can be at most $m+1$ graphs G that map onto a given graph $H \in \mathcal{G}_{m'}$ using this procedure. This gives the inequality (5.5).

Moreover, using the first part of Lemma 2.9 and following the same argument as in the proof of Lemma 5.1 it can be shown that

$$|\mathcal{H}| \geq \frac{1}{n^{12}} |\mathcal{G}_m| \,. \tag{5.6}$$

To see this, note that for any graph $G \in \mathcal{G}_m$, nodes i and j with $d_i(G) < u_i$ and $d_j(G) < u_j$ always exist, as $m < m_2$. Moreover, if $\{i, j\} \in E(G)$ we know from Lemma 2.9 that there is a graph G' with the same degree sequence not containing edge $\{i, j\}$ close to G.

We next show that $|\mathcal{G}_m| \geq |\mathcal{G}_{m'}|/p(n)$ for some polynomial $p(n)$. To see this, note that for any $G' \in |\mathcal{G}_{m'}|$ there exist nodes x and y such that $d_x(G') > \ell_x$ and $d_y(G') > \ell_y$ as $m' = m + 1 > m_1$. If $\{x, y\} \in E(G')$ we can remove it to obtain an element of $|\mathcal{G}_m|$. Otherwise, again using Lemma 2.9 we can first find an element $G'' \in \mathcal{G}_{m'}$ close to G' that contains $\{x, y\}$ and the remove it. Combining this with (5.6) yields the existence of a polynomial $q_2(n)$ such that

$$|\mathcal{H}| \geq \frac{1}{q_2(n)} |\mathcal{G}_{m'}| \,.$$

Finally, combining the latter inequality with (5.5) and dividing both sides by $\sum_{m \in \mathcal{T}} w_m$, then gives the desired result. ∎

In order to show that $\mathcal{M}_{\mathcal{T}}$ is rapidly mixing or, in particular, that the gap $\text{Gap}(\mathcal{T})$ can be polynomially bounded, it is sufficient to show that the sequence $(w_m)_{m \in \mathcal{T}}$ is *log-concave*. Log-concavity means that for any $m \in \mathcal{T} \setminus \{m_1, m_2\}$, $w_{m-1}w_{m+1} \leq w_m^2$.

THEOREM 5.4. *The sequence $(w_m)_{m \in \mathcal{T}}$ is log-concave for all interval sequences $[\boldsymbol{\ell}, \boldsymbol{u}]$ for which $u_i \in \{\ell_i, \ell_i + 1\}$ for all $i \in [n]$.*

Proof. We follow the notation, terminology and general outline of the proof of Theorem 5.1 in [36]. Define $A = \mathcal{G}_{m+1} \times \mathcal{G}_{m-1}$ and $B = \mathcal{G}_m \times \mathcal{G}_m$. We will show that $|A| \leq |B|$, from which the claim follows.

Note that the symmetric difference of any two subgraphs of K_n can be decomposed into a collection of alternating cycles and simple paths. We will do this in a canonical way.[6] Fix some total order $\preceq_e$ on the edges of K_n. For two subgraphs G and G' we will call edges in $E(G) \backslash E(G')$ blue, and edges in $E(G') \backslash E(G)$ red. Around every node, we will pair up blue edges with red edges as much as possible. We do this by repeatedly selecting a node and pairing up the lowest ordered red and blue edge that have not yet been paired up. This yields a decomposition of the symmetric difference into i) alternating red-blue cycles, ii) alternating simple red-blue paths of even length (with same number of red and blue edges), iii) simple paths ending and starting with a red edge, iv) simple paths ending and starting with a blue edge. We call this the *canonical symmetric difference decomposition of $E(G) \triangle E(G')$ with respect to $\preceq_e$*, or simply the canonical decomposition of $E(G) \triangle E(G')$. We call a simple path a G-path if it contains one more edge of G than of G' (i.e., red edges are one more than blue edges), and a G'-path if it contains one more edge of G'. We emphasize that any path of odd length in the symmetric difference is of one of these two types.

[6] This decomposition is the main extra step we need compared to the proof of Theorem 5.1 in [36]. The symmetric difference of two matchings is by construction already a disjoint union of cycles and paths. This is also where the analysis breaks down in case the degree intervals have length at least two.

Now, for every pair $(G, G') \in A$ it holds that the number of G-paths exceeds the number of G'-paths by precisely two (as G has two edges more than G'). For this reason, we partition A into disjoint classes $\{A_r : r = 1, \ldots, m\}$ where

$$A_r = \{(G, G') \in A : \quad \begin{array}{l} \text{the canonical decomposition of } E(G) \triangle E(G') \\ \text{contains } r+1 \text{ } G\text{-paths and } r-1 \text{ } G'\text{-paths}\} \,. \end{array}$$

In order to prove $|A| \leq |B|$ it suffices to show $|A_r| \leq |B_r|$ for all r. We call a pair $(L, L') \in B$ *reachable* from $(G, G') \in A$ if and only if $E(G) \triangle E(G) = E(L) \triangle E(L')$ and L is obtained from G by taking some G-path in the canonical decomposition and flipping the parity of the edges with respect to G and G'. It is important to see that the canonical symmetric difference decomposition of the pairs (G, G') and (L, L') is the same because all degree intervals have length one. Note that the number of pairs in B_r reachable from a given $(G, G') \in A_r$ is precisely the number of G-paths in the canonical decomposition of G and G', which is $r + 1$. Conversely, any given $(L, L') \in B_r$ is reachable from precisely r pairs in A_r. Therefore, if $|A_r| > 0$, we have

$$\frac{|B_r|}{|A_r|} = \frac{r+1}{r} > 1 \,.$$

This proves the claim. ∎

We are ready to bound the spectral gap of $P_{\mathcal{T}}$. Note that Ω in the statement of Theorem 5.5 is actually $\mathcal{T}$. Recall that $|\mathcal{T}| = m_2 - m_1 + 1 \leq n/2 + 1 \leq n$. Moreover, the ratios w_i/w_j are also polynomially bounded for any $i, j \in \mathcal{T}$ with $|i - j| = 1$. This can be shown exactly as in the proofs of the Lemmata 5.1 and 5.3; see also Appendix B. As a result, it is sufficient to prove the statement in Theorem 5.5 below in order to bound the gap of $P_{\mathcal{T}}$.

THEOREM 5.5. *Let $(w_m)_{m \in \Omega}$ be a log-concave sequence of non-negative numbers and let $\mathcal{M} = (\Omega, P)$ be a Markov chain with transition probabilities*

$$P(i, j) = \begin{cases} \frac{1}{4} \min\{1, w_j/w_i\} & \text{if } |i - j| = 1 \,, \\ 0 & \text{if } |i - j| > 1 \,, \\ 1 - P(i, i-1) - P(i, i+1) & \text{if } i = j \,. \end{cases}$$

Then $1/\operatorname{Gap}(P) \leq 4\,|\Omega|^3 \max_{i,j:|i-j|=1} w_i/w_j$.[7]

Proof. First note that the stationary distribution π of $\mathcal{M}$ is proportional to the weights $(w_i)_{i \in \Omega}$, i.e., $\pi(i) = w_i / \sum_{p \in \Omega} w_p$, as desired. Consider the straightforward multi-commodity flow f in which we route $\pi(i)\,\pi(j)$ units of flow over the path $i \to (i+1) \to \cdots \to j$ if $i < j$, or $i \to (i-1) \to \cdots \to j$ if $i > j$. Recall from Section 2.2 that $\operatorname{Gap}(P)^{-1} \leq \operatorname{load}(f) \operatorname{length}(f)$, where $\operatorname{length}(f)$ is the length of a longest flow-carrying path and $\operatorname{load}(f)$ is the maximum load on eny edge of the state space graph of the chain. By the definition of the flow f, we have $\operatorname{length}(f) \leq |\Omega|$. Next we bound $\operatorname{load}(f)$.

We consider a fixed transition $e = (z, z+1)$. Note that the proof for transitions of the form $(z, z-1)$ is symmetric, since a sequence $(w_i)_{i \in \Omega}$ is log-concave if and only if the

[7] We suspect a similar result is true without the dependence on the w_i but this is not needed for our purpose.

sequence $(w_{|\Omega|-i+1})_{i\in\Omega}$ is log-concave. We have

$$\text{load}(e) = \sum_{1\leq i\leq z}\sum_{z<j\leq|\Omega|}\frac{\pi(i)\,\pi(j)}{\pi(z)P(z,z+1)} \leq 4\max_{i,j:|i-j|=1}\frac{w_i}{w_j}\sum_{1\leq i\leq z}\sum_{z<j\leq|\Omega|}\frac{\pi(i)\,\pi(j)}{\pi(z)}$$

$$= 4\max_{i,j:|i-j|=1}\frac{w_i}{w_j}\left(\sum_{p\in\Omega}w_p\right)^{-1}\sum_{1\leq i\leq z}\sum_{z<j\leq|\Omega|}\frac{w_iw_j}{w_z}\,. \tag{5.7}$$

Log-concavity of the sequence $(w_q)_{q\in\Omega}$ implies that for any fixed $i<j$, and any $a\in\mathbb{N}$ such that $i+a\leq j-a$, we have

$$w_iw_j \leq w_{i+a}w_{j-a}\,. \tag{5.8}$$

This follows from repeatedly applying the log-concavity condition. Indeed, log-concavity gives us $\frac{w_i}{w_{i+1}} \leq \frac{w_{i+1}}{w_{i+2}} \leq \ldots \leq \frac{w_{j-2}}{w_{j-1}} \leq \frac{w_{j-1}}{w_j}$ and thus $w_iw_j \leq w_{i+1}w_{j-1}$. By repeating this with $i+1$ and $j-1$ (i.e., by removing the outer terms) we get $\frac{w_{i+1}}{w_{i+2}} \leq \ldots \leq \frac{w_{j-2}}{w_{j-1}}$ and thus $w_{i+1}w_{j-1} \leq w_{i+2}w_{j-2}$. After a steps we get (5.8).

Now, for a fixed i and j in the double summation in (5.7), let a_{ij} be such that $w_{i+a_{ij}}$ or $w_{j-a_{ij}}$ (or both) equals w_z. Then (5.8) gives us that $w_iw_j \leq w_z\,w_p$ for some $p\in\Omega$. Note that for any choice of z, the double summation in (5.7) has at most $|\Omega|^2$ terms (as there are at most $|\Omega|$ choices for i and j). This implies that

$$\sum_{1\leq i\leq z}\sum_{z<j\leq|\Omega|}w_iw_j/w_z \leq |\Omega|^2\sum_{p\in\Omega}w_zw_p/w_z = |\Omega|^2\sum_{p\in\Omega}w_p\,.$$

Combining this inequality with (5.7), we obtain

$$\text{load}(e) \leq 4|\Omega|^2\max_{i,j:|i-j|=1}w_i/w_j\,,$$

and, thus, $1/\operatorname{Gap}(P) \leq \text{load}(f)\,\text{length}(f) \leq 4\,|\Omega|^3\max_{i,j:|i-j|=1}w_i/w_j$, as required. ∎

This then completes the proof of Theorem 1.3.

6. Discussion and future directions. We did not attempt to optimize the upper bounds on the mixing times of the Markov chains involved. Already for the *switch Markov chain* no low-degree polynomial upper bounds are known on its mixing time. For instance, the best known upper bound for r-regular graphs is $r^{23}n^8(rn\log(rn)+\log(1/\epsilon))$ [15,16]. This is a central issue for many MCMC approaches for sampling graphs with given degrees (or degree intervals in our case). Various non-MCMC approaches to the problem, see, e.g., [4,27,28,42,48,56], often have better running times, but only work for smaller classes of degree sequences or have weaker guarantees on the uniformity of the output than we require in our setting.

An interesting first direction for future work is determining whether the degree interval chain is rapidly mixing for more general instances. The most intriguing question from our point of view, however, is whether there is a black-box reduction implying that if the switch Markov chain is rapidly mixing for all degree sequences $\boldsymbol{d}$ satisfying $\boldsymbol{\ell}\leq\boldsymbol{d}\leq\boldsymbol{u}$, then the degree interval Markov chain is also rapidly mixing. Even more generally, can the problem of sampling graphs with given degree intervals always be reduced to the problem of sampling graphs with given degrees?

Further, one could explore other, non-MCMC, approaches for approximate sampling, especially when the degree ranges are relatively large. Can one come up with an algorithm in which resampling certain "bad events" (e.g., resampling edges adjacent to a node not satisfying its degree interval constraints) yields an exactly uniform sample, following the "partial rejection sampling" framework of Guo, Jerrum and Liu [31]? While this seems unlikely when sampling graphs with given degrees, we suspect it is possible for the problem of sampling graphs with (sufficiently large) given degree intervals.

REFERENCES

[1] G. Amanatidis and P. Kleer, *Rapid mixing of the switch Markov chain for strongly stable degree sequences and 2-class joint degree matrices*, in Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2019, pp. 966–985.

[2] N. Anari, K. Liu, S. Oveis Gharan, and C. Vinzant, *Log-concave polynomials II: high-dimensional walks and an FPRAS for counting bases of a matroid*, in Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, 2019, pp. 1–12.

[3] N. Anari, S. Oveis Gharan, and C. Vinzant, *Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids*, in Proceedings of the 59th IEEE Annual Symposium on Foundations of Computer Science, IEEE Computer Society, 2018, pp. 35–46.

[4] M. Bayati, J. H. Kim, and A. Saberi, *A sequential algorithm for generating random graphs*, Algorithmica, 58 (2010), pp. 860–910.

[5] E. A. Bender and E. R. Canfield, *The asymptotic number of labeled graphs with given degree sequences*, Journal of Combinatorial Theory, Series A, 24 (1978), pp. 296 – 307.

[6] I. Bezáková, N. Bhatnagar, and E. Vigoda, *Sampling binary contingency tables with a greedy start*, Random Structures & Algorithms, 30 (2007), pp. 168–205.

[7] J. Blitzstein and P. Diaconis, *A sequential importance sampling algorithm for generating random graphs with prescribed degrees*, Internet mathematics, 6 (2011), pp. 489–522.

[8] S. G. Bobkov and P. Tetali, *Modified logarithmic Sobolev inequalities in discrete settings*, Journal of Theoretical Probability, 19 (2006), pp. 289–336.

[9] B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European Journal of Combinatorics, 1 (1980), pp. 311–316.

[10] P. Brändén, *Personal communication*, 2020.

[11] P. Brändén and J. Huh, *Lorentzian polynomials*, Annals of Mathematics, 192 (2020), pp. 821–891.

[12] S. Caracciolo, A. Pelissetto, and A. D. Sokal, *Two remarks on simulated tempering*, Unpublished manuscript, (1992).

[13] C. J. Carstens and P. Kleer, *Speeding up switch Markov chains for sampling bipartite graphs with given degree sequence*, in Proceedings of the 22nd International Conference on Randomization and Computation, vol. 116 of LIPIcs, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018, pp. 36:1–36:18.

[14] A. C. Coolen, A. Annibale, and E. Roberts, *Generating random networks and graphs*, Oxford university press, 2017.

[15] C. Cooper, M. E. Dyer, and C. S. Greenhill, *Sampling regular graphs and a peer-to-peer network*, Combinatorics, Probability & Computing, 16 (2007), pp. 557–593.

[16] C. Cooper, M. E. Dyer, and C. S. Greenhill, *Corrigendum: Sampling regular graphs and a peer-to-peer network*, CoRR, abs/1203.6111 (2012), https://arxiv.org/abs/1203.6111.

[17] C. Cooper, M. E. Dyer, C. S. Greenhill, and A. J. Handley, *The flip Markov chain for connected regular graphs*, Discrete Applied Mathematics, (2018).

[18] M. Cryan, H. Guo, and G. Mousa, *Modified log-sobolev inequalities for strongly log-concave distributions*, in Proceedings of the 60th Annual Symposium on Foundations of Computer Science, IEEE Computer Society, 2019, pp. 1358–1370.

[19] P. Diaconis and L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, The Annals of Applied Probability, 6 (1996), pp. 695–750.

[20] M. E. Dyer, C. S. Greenhill, P. Kleer, J. Ross, and L. Stougie, *Sampling hypergraphs with given degrees*, Discrete Mathematics, 344 (2021), p. 112566.

[21] P. L. Erdős, E. Győri, T. R. Mezei, I. Miklós, and D. Soltész, *Half-graphs, other non-stable degree sequences, and the switch Markov chain*, Electronic Journal of Combinatorics, 28 (2021).

[22] P. L. Erdős, C. S. Greenhill, T. R. Mezei, I. Miklós, D. Soltész, and L. Soukup, *The mixing time of switch Markov chains: A unified approach*, European Journal of Combinatorics, 99 (2022), p. 103421.

[23] P. L. Erdős, T. R. Mezei, and I. Miklós, *Approximate sampling of graphs with near-p-stable degree intervals*, Annals of Combinatorics, 28 (2024), pp. 223–256.

[24] P. L. Erdős, T. R. Mezei, I. Miklós, and D. Soltész, *Efficiently sampling the realizations of bounded, irregular degree sequences of bipartite and directed graphs*, PLOS ONE, 13 (2018), pp. 1–20.
[25] P. L. Erdős, I. Miklós, and Z. Toroczkai, *A decomposition based proof for fast mixing of a Markov chain over balanced realizations of a joint degree matrix*, SIAM Journal on Discrete Mathematics, 29 (2015), pp. 481–499.
[26] T. Feder, A. Guetz, M. Mihail, and A. Saberi, *A local switch Markov chain on given degree graphs with application in connectivity of peer-to-peer networks*, in Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, 2006, pp. 69–76.
[27] P. Gao and N. C. Wormald, *Uniform generation of random regular graphs*, SIAM Journal of Computing, 46 (2017), pp. 1395–1427.
[28] P. Gao and N. C. Wormald, *Uniform generation of random graphs with power-law degree sequences*, in Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2018, pp. 1741–1758.
[29] S. Goel, *Modified logarithmic Sobolev inequalities for some models of random walk*, Stochastic processes and their applications, 114 (2004), pp. 51–79.
[30] C. S. Greenhill and M. Sfragara, *The switch Markov chain for sampling irregular graphs and digraphs*, Theoretical Computer Science, 719 (2018), pp. 1–20.
[31] H. Guo, M. Jerrum, and J. Liu, *Uniform sampling through the Lovász local lemma*, Journal of the ACM, 66 (2019), p. 18.
[32] L. Gurvits, *On multivariate Newton-like inequalities*, in Advances in combinatorial mathematics, Springer, 2009, pp. 61–78.
[33] J. Herzog and T. Hibi, *Discrete polymatroids*, Journal of Algebraic Combinatorics, 16 (2002), pp. 239–268.
[34] M. Jerrum, *Counting, sampling and integrating: algorithms and complexity*, Springer Science & Business Media, 2003.
[35] M. Jerrum, B. D. McKay, and A. Sinclair, *When is a graphical sequence stable?*, Random Graphs, Vol. 2, (1992), pp. 101–116.
[36] M. Jerrum and A. Sinclair, *Approximating the permanent*, SIAM Journal on Computing, 18 (1989), pp. 1149–1178.
[37] M. Jerrum and A. Sinclair, *Fast uniform generation of regular graphs*, Theoretical Computer Science, 73 (1990), pp. 91 – 100.
[38] M. Jerrum and A. Sinclair, *The Markov chain Monte Carlo method: An approach to approximate counting and integration*, in Approximation Algorithms for NP-hard Problems, PWS Publishing, 1997, pp. 482–520.
[39] M. Jerrum, A. Sinclair, and E. Vigoda, *A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries*, Journal of the ACM, 51 (2004), pp. 671–697.
[40] M. Jerrum, L. G. Valiant, and V. V. Vazirani, *Random generation of combinatorial structures from a uniform distribution*, Theoretical Computer Science, 43 (1986), pp. 169–188.
[41] R. Kannan, P. Tetali, and S. Vempala, *Simple Markov-chain algorithms for generating bipartite graphs and tournaments*, Random Structures and Algorithms, 14 (1999), pp. 293–308.
[42] J. H. Kim and V. H. Vu, *Generating random regular graphs*, Combinatorica, 26 (2006), pp. 683–708.
[43] P. Kleer, *Sampling from the gibbs distribution in congestion games*, Mathematics of Operations Research, 48 (2023), pp. 1846–1870.
[44] A. Liebenau and N. Wormald, *Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph*, Journal of the European Mathematical Society, 26 (2023), pp. 1–40.
[45] N. Madras and D. Randall, *Markov chain decomposition for convergence rate analysis*, Ann. Appl. Probab., 12 (2002), pp. 581–606.
[46] R. A. Martin and D. Randall, *Disjoint decomposition of Markov chains and sampling circuits in Cayley graphs*, Combinatorics, Probability & Computing, 15 (2006), pp. 411–448.
[47] B. D. McKay and N. C. Wormald, *Asymptotic enumeration by degree sequence of graphs of high degree*, European Journal of Combinatorics, 11 (1990), pp. 565 – 580.
[48] B. D. McKay and N. C. Wormald, *Uniform generation of random regular graphs of moderate degree*, Journal of Algorithms, 11 (1990), pp. 52 – 67.
[49] I. Miklós, P. L. Erdős, and L. Soukup, *Towards random uniform sampling of bipartite graphs with given degree sequence*, Electronic Journal of Combinatorics, 20 (2013).
[50] K. Murota, *Discrete convex analysis*, Mathematical Programming, 83 (1998), pp. 313–371.
[51] K. Murota, *Recent developments in discrete convex analysis*, in Research Trends in Combinatorial Optimization, 2009, pp. 219–260.
[52] B. P. Olding and P. J. Wolfe, *Inference for graphs and networks: Adapting classical tools to*

modern data, in Data Analysis for Network Cyber-Security, World Scientific, 2014, pp. 1–31.

[53] S. Oveis Gharan, *Lecture notes (lecture 1)*, CSE 599: Counting and Sampling, (2017), https://homes.cs.washington.edu/~shayan/courses/sampling/counting-1.pdf.

[54] S. Rechner, L. Strowick, and M. Müller-Hannemann, *Uniform sampling of bipartite graphs with degrees in prescribed intervals*, Journal of Complex Networks, 6 (2018), pp. 833–858.

[55] A. Sinclair, *Improved bounds for mixing rates of Markov chains and multicommodity flow*, Combinatorics, Probability and Computing, 1 (1992), pp. 351–370.

[56] A. Steger and N. C. Wormald, *Generating random regular graphs quickly*, Combinatorics, Probability & Computing, 8 (1999), pp. 377–396.

Appendix A. Modified log-Sobolev constant.

Let $\mathcal{M} = (\Omega, P)$ be a time-reversible Markov chain with stationary distribution π, and $f, g : \Omega \to \mathbb{R}_{\geq 0}$. Let $\mathbb{E}_\pi(f) = \sum_{x \in \Omega} \pi(x) f(x)$. Furthermore, define the entropy-like quantity

$$\mathrm{Ent}_\pi(f) = \mathbb{E}_\pi \left[f \log(f) - f \log(\mathbb{E}_\pi(f)) \right] ,$$

and the *Dirichlet form*

$$\mathcal{E}_P(f, g) = \frac{1}{2} \sum_{x \in \Omega} \sum_{y \in \Omega} \pi(x) P(x, y) [f(x) - f(y)][g(x) - g(y)] .$$

The *modified log-Sobolev constant* of the Markov chain $\mathcal{M}$ is defined by

$$\rho(P) = \inf \left\{ \frac{\mathcal{E}_P(f, \log(f))}{\mathrm{Ent}_\pi(f)} \;\middle|\; f : \Omega \to \mathbb{R}_{\geq 0},\ \mathrm{Ent}_\pi(f) \neq 0 \right\} .$$

As stated in Section 2, it holds that (see, e.g., [8])

$$\tau_x(\epsilon) \leq \frac{1}{\rho(P)} \left(\log \log \pi(x)^{-1} + \log \left(\frac{1}{2\epsilon^2} \right) \right) .$$

Furthermore, for any Markov chain it holds that

$$2(1 - \lambda_1(P)) \geq \rho(P) , \tag{A.1}$$

where $\lambda_1(P)$ is the second-largest eigenvalue of P (assuming the Markov chain is lazy).

A.1. Markov chain comparison.

A useful property of proving mixing time bounds through the modified log-Sobolev constant, is that it is easy to see that small perturbations in the transition probabilities and the stationary distribution only result in mild variations in the modified log-Sovolev constant of the resulting Markov chain (by means of a Markov chain comparison argument). Goel [29] states the following for the modified log-Sobolev constant, based on similar results for the other constants by Diaconis and Saloff-Coste [19]. The notation $W(\Omega, \pi)$ is used to denote the set of all (test) functions $f : \Omega \to \mathbb{R}_{\geq 0}$.

THEOREM A.1 (Lemma 4.1 [29]). *Let $\mathcal{M} = (\Omega, P)$ and $\mathcal{M}' = (\Omega', P')$ be two finite, reversible Markov chains with stationary distributions π and π', respectively, and modified log-Sobolev constant ρ and ρ', respectively. Assume there is a mapping $\phi : W(\Omega, \pi) \to W'(\Omega', \pi')$ mapping $f \to f'$ for $f : \Omega \to \mathbb{R}_{\geq 0}$, and constants $C, c > 0$ and $B \geq 0$ such that for all $f \in W(\Omega, \pi)$, we have*

$$\mathcal{E}_{P'}(f', \log f') \leq C \cdot \mathcal{E}_P(f, \log f) \quad \text{and} \quad c \cdot \mathit{Ent}_\pi(f) \leq \mathit{Ent}_{\pi'}(f') + B \cdot \mathcal{E}_P(f, \log f) .$$

Then

$$\frac{c\rho'}{C + B\rho'} \leq \rho .$$

COROLLARY A.2. *With $\mathcal{M}$ and $\mathcal{M}'$ as in Theorem A.1, if $\Omega = \Omega'$ and there exists a $0 < \delta < 1$ such that $(1 - \delta) P(x, y) \leq P'(x, y) \leq (1 + \delta) P(x, y)$ for all $x, y \in \Omega$, and $(1 - \delta)\pi(x) \leq \pi'(x) \leq (1 + \delta)\pi(x)$ for $x \in \Omega$, it directly follows that*

$$\frac{1}{\rho} \leq \frac{1 + \delta}{1 - \delta} \cdot \frac{1}{\rho'} . \tag{A.2}$$

Appendix B. Reductions for approximate sampling and counting. We first explain how Theorem 1.1 follows from Theorem 1.2. The induced approximate sampler from Theorem 1.2 can be turned into an approximate counter for $|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$ by standard techniques (Section B.1). Furthermore, this approximate counter can then be turned into an approximate counter for $|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$ by means of a simple reduction (Section B.2). In turn, the approximate counter for $|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$ can be transformed into an approximate sampler from $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$, again by a standard technique (Section B.3).

A subtle point is that it is not known whether the problem of sampling and counting from $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$, or $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$, is *self-reducible* [40]. This follows roughly from the fact that it is not known whether the problem of sampling/counting from $\mathcal{G}(\boldsymbol{d})$ is self-reducible in general. However, if one restricts to degree intervals $[\boldsymbol{\ell}, \boldsymbol{u}]$ for which both an FPRAS and FPAUS for $\mathcal{G}(\boldsymbol{d})$ is known for every $\boldsymbol{\ell} \leq \boldsymbol{d} \leq \boldsymbol{u}$, like the set of P-stable degree sequences [37], then standard reduction techniques for self-reducible problems can still be applied.

To give a more concrete intuition, in the reduction of Section B.1 the problem is that one needs to be able to compute the final factor in the telescoping product (B.1) efficiently, which we do not know how to do for an arbitrary degree sequence u (although we do know it for P-stable degree sequences by the results of Jerrum and Sinclair [37]).

B.1. From approximate sampling from $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ to approximating $|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$. Using a standard reduction technique, see, e.g., Chapter 12 in [38], we can turn our FPAUS into an FPRAS for counting the number of graphs with given degree intervals.

We first show how to express $|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$ as a *telescoping product*. We write

$$|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})| = \frac{|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|}{|\mathcal{G}_m(\boldsymbol{a}^1, \boldsymbol{u})|} \frac{|\mathcal{G}_m(\boldsymbol{a}^1, \boldsymbol{u})|}{|\mathcal{G}_m(\boldsymbol{a}^2, \boldsymbol{u})|} \cdots \frac{|\mathcal{G}_m(\boldsymbol{a}^{p-1}, \boldsymbol{u})|}{|\mathcal{G}_m(\boldsymbol{u})|} |\mathcal{G}_m(\boldsymbol{u})| \tag{B.1}$$

for a sequence of vectors $\boldsymbol{\ell} = \boldsymbol{a}^0, \boldsymbol{a}^1, \boldsymbol{a}^2, \ldots, \boldsymbol{a}^p = \boldsymbol{u}$, that are recursively defined as follows. We define $\boldsymbol{a}^{i+1}$ by choosing the lowest indexed nodes[8] v and w for which $a_v^i < u_v^i$ and $a_w^i < u_w^i$, and then setting

$$a_j^{i+1} = \begin{cases} a_j^i + 1 & \text{if } j \in \{v, w\}\,, \\ a_j^i & \text{otherwise}\,. \end{cases}$$

It is clear that there is some $p \leq 2\sum_i u_i$ such that this procedure gives $\boldsymbol{a}^p = \boldsymbol{u}$. Also, if $c = \max_i(u_i - \ell_i)$, all intermediate degree intervals $[\boldsymbol{a}^i, \boldsymbol{u}]$ also have length at most c component-wise, as $u_j - a_j^i \leq u_j - \ell_j \leq c$. Finally, note that we have

$$\boldsymbol{a}^0 < \boldsymbol{a}^1 < \cdots < \boldsymbol{a}^{p-1} < \boldsymbol{a}^p, \tag{B.2}$$

where for two sequences $\boldsymbol{a}$ and $\boldsymbol{b}$, we write $\boldsymbol{a} < \boldsymbol{b}$ if $\boldsymbol{a} \leq \boldsymbol{b}$ and $a_i < b_i$ for at least one $i \in \{1, \ldots, n\}$.

In order to approximate the size of $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$, it will be sufficient to approximate the ratios in the telescoping product, as well as the last factor $|\mathcal{G}_m(\boldsymbol{u})|$. The latter can be approximated by employing, e.g., the approximate counting scheme of Jerrum and Sinclair [37] for *P-stable* degree sequences. For approximating the ratios, we need the following two (sufficient) components: (1) the existence of a FPAUS and (2) the fact that the ratios can be polynomially bounded. This implies that polynomially many samples are enough in order to estimate the ratio up to the desired accuracy.

[8] A number $i \in [n]$ is lower indexed than $j \in [n]$ if $i < j$.

We sketch how to formalize this argument. Using strong stability, i.e., Proposition 2.8, and very similar reasoning as in the proofs of Lemmas 5.3 and 5.1, it is easy to show that all ratios in (B.1) are upper bounded by some polynomial $p_2(n)$ (independent of $\boldsymbol{\ell}$ and $\boldsymbol{u}$). By setting

$$\nu_i = \frac{|\mathcal{G}_m(\boldsymbol{a}^i, \boldsymbol{u})|}{|\mathcal{G}_m(\boldsymbol{a}^{i+1}, \boldsymbol{u})|},$$

this just means $1 \leq \nu_i \leq p_2(n)$. Moreover, using (B.2), it follows that, for $i = 0, \ldots, p-1$, we have $\mathcal{G}_m(\boldsymbol{a}^{i+1}, \boldsymbol{u}) \subseteq \mathcal{G}_m(\boldsymbol{a}^i, \boldsymbol{u})$. If we define X_i to be the indicator variable of the event that a random sample from $\mathcal{G}_m(\boldsymbol{a}^i, \boldsymbol{u})$ is indeed contained in $\mathcal{G}_m(\boldsymbol{a}^{i+1}, \boldsymbol{u})$, then $\nu_i = 1/\mathbb{E}(X_i)$. The high-level idea is now to show that polynomially many samples from the sampler (the switch hinge-flip Markov chain) not only suffice for an accurate estimate for ν_i but, crucially, they suffice for an accurate estimate of the product

$$\prod_{i=0}^{p-1} \nu_i = \frac{|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|}{|\mathcal{G}_m(\boldsymbol{u})|}$$

up to a factor $(1 \pm \epsilon/3)$. This can be done by standard arguments, e.g., see Chapter 12 of [38] or Chapter 3.2 of [34]. Finally, as mentioned above, we may use the approximate counter from [37] for approximating $|\mathcal{G}_m(\boldsymbol{u})|$ up to a factor $(1 \pm \epsilon/3)$. This then implies that we can also approximate $|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$ up to a factor $(1 \pm \epsilon)$.

B.2. From approximating $|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$ to approximating $|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$. In order to provide an FPRAS for approximating $|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$, it suffices to give an FPRAS for approximating $|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$ for every $\frac{1}{2}\sum_i \ell_i \leq m \leq \frac{1}{2}\sum_i u_i$. Recall that $\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})$ is the set of graphs with degree intervals $[\boldsymbol{\ell}, \boldsymbol{u}]$ and for which the total number of edges is equal to m.

LEMMA B.1. *Suppose there is an FPRAS for approximating $|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$ for every $m \in \mathbb{N}$ such that $\frac{1}{2}\sum_i \ell_i \leq m \leq \frac{1}{2}\sum_i u_i$. Then there is an FPRAS for approximating $|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$.*

Proof. We write $a = \frac{1}{2}\sum_i \ell_i$ and $b = \frac{1}{2}\sum_i u_i$. Note that there are at most $b - a \leq n^2$ possible choices for m, and that

$$|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})| = \sum_{m=a}^{b} |\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|.$$

Now, for every m use the given FPRAS for approximating $|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$ with $\delta' = \delta/n^2$. It outputs a number c_m satisfying $(1-\epsilon)|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})| \leq c_m \leq (1+\epsilon)|\mathcal{G}_m(\boldsymbol{\ell}, \boldsymbol{u})|$ with probability at least $1 - \delta/n^2$. Then $c = \sum_m c_m$ satisfies $(1-\epsilon)|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})| \leq c \leq (1+\epsilon)|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$ with probability at least

$$(1 - \delta/n^2)^{b-a} \geq (1 - \delta/n^2)^{n^2} \geq 1 - \delta.$$

This completes the proof. ∎

B.3. From approximating $|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$ to approximate sampling from $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$. Again, using a reduction inspired by a similar one for self-reducible problems, see, e.g., [40], we can turn our FPRAS for computing $|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|$ into an FPAUS for sampling from $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$. Note that if $\boldsymbol{\ell} = \boldsymbol{u}$, we can simply use the approximate sampler from Jerrum and Sinclair [37] when $\boldsymbol{\ell}$ is near-regular.

As long as $\boldsymbol{\ell} \neq \boldsymbol{u}$ there is some i such that $\ell_i < u_i$. We can partition the set $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ based on whether or not the degree of a graphical realization G with $\boldsymbol{\ell} \leq \boldsymbol{d}(G) \leq \boldsymbol{u}$ satisfies $d_i = \ell_i$ or $d_i \geq \ell_i + 1$. For a given vector $\boldsymbol{z} = (z_1, \ldots, z_n) \in \mathbb{R}^n$ and $z_i' \in \mathbb{R}$ we write $(z_i', \boldsymbol{z}_{-i}) = (z_1, \ldots, z_{i-1}, z_i', z_{i+1}, \ldots, z_n)$. We then have $\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})$ as the disjoint union

$$\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u}) = \mathcal{G}(\boldsymbol{\ell}, (\ell_i, \boldsymbol{u}_{-i})) \cup \mathcal{G}((\ell_i + 1, \boldsymbol{\ell}_{-i}), \boldsymbol{u}).$$

Roughly speaking the idea is to use the approximation scheme, and approximate the marginal probabilities

$$\frac{|\mathcal{G}(\boldsymbol{\ell}, (\ell_i, \boldsymbol{u}_{-i}))|}{|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|} \quad \text{and} \quad \frac{|\mathcal{G}((\ell_i + 1, \boldsymbol{\ell}_{-i}), \boldsymbol{u})|}{|\mathcal{G}(\boldsymbol{\ell}, \boldsymbol{u})|}$$

up to a sufficient accuracy. Then we sample one of the sets $\mathcal{G}(\boldsymbol{\ell}, (\ell_i, \boldsymbol{u}_{-i}))$ or $\mathcal{G}((\ell_i + 1, \boldsymbol{\ell}_{-i}), \boldsymbol{u})$ according to these—sufficiently accurate—marginals, and keep applying this procedure recursively. However, this procedure only gives a poly$(1/\epsilon)$ dependence and not the desired $\log(1/\epsilon)$ dependence. This can be achieved by using a slightly different version of the above in combination with rejection sampling. See, e.g., [53] for this idea in the context of (approximately) sampling and counting matchings from a given graph.

We then repeat this step until the lower and upper bound defining the intervals are equal. Note that this step is only carried out a polynomial number of times. After this we have, roughly speaking, sampled a degree sequence $\boldsymbol{d}$ with $\boldsymbol{\ell} \leq \boldsymbol{d} \leq \boldsymbol{u}$ according to the (approximately) correct marginal probability. After this we can use the approximate sampler from [37] to sample from $\mathcal{G}(\boldsymbol{d})$ (or, e.g., simply the switch Markov chain).