

# RANDOM GRAPHS WITH GIVEN VERTEX DEGREES AND SWITCHINGS

SVANTE JANSON

ABSTRACT. Random graphs with a given degree sequence are often constructed using the configuration model, which yields a random multigraph. We may adjust this multigraph by a sequence of switchings, eventually yielding a simple graph. We show that, assuming essentially a bounded second moment of the degree distribution, this construction with the simplest types of switchings yields a simple random graph with an almost uniform distribution, in the sense that the total variation distance is  $o(1)$ . This construction can be used to transfer results on distributional convergence from the configuration model multigraph to the uniform random simple graph with the given vertex degrees. As examples, we give a few applications to asymptotic normality. We show also a weaker result yielding contiguity when the maximum degree is too large for the main theorem to hold.

## 1. INTRODUCTION

We consider random graphs with vertex set  $[n] := \{1, \dots, n\}$  and a given degree sequence  $\mathbf{d} = (d_1, \dots, d_n)$ . In particular, we define  $G(n, \mathbf{d})$  to be the random (simple) graph with degree sequence  $\mathbf{d}$  chosen uniformly at random among all such graphs. We will consider asymptotic results as  $n \rightarrow \infty$ , where the degree sequence  $\mathbf{d} = \mathbf{d}_n = (d_i^{(n)})_1^n$  depends on  $n$ , but usually we omit  $n$  from the notation.

The standard methods to construct a random graph with a given degree sequence begin with the *configuration model*, which was introduced by Bollobás [7]. (See [6; 31] for related models and arguments.) As is well-known, this method yields a random multigraph, which we denote by  $G^*(n, \mathbf{d})$ , with the given degree sequence  $\mathbf{d}$ ; see Section 3.1. This random multigraph may contain loops and multiple edges; however, in the present paper (as in many others), we will consider asymptotic results as  $n \rightarrow \infty$ , where  $\mathbf{d} = \mathbf{d}_n$  satisfies (at least)

$$\sum_{i=1}^n d_i = \Theta(n), \quad \sum_{i=1}^n d_i^2 = O(n), \quad (1.1)$$

and then (see e.g. the proof of Lemma 4.1) the expected number of loops and multiple edges is  $O(1)$ , which might seem insignificant when  $n$  is large. (Recall that  $\Theta(n)$  means a number in the interval  $[cn, Cn]$  for some constants  $c, C > 0$ .)

---

*Date:* 28 January, 2019; revised 30 January, 2019.

*2010 Mathematics Subject Classification.* 05C80, 60C05.

Partly supported by the Knut and Alice Wallenberg Foundation.

In fact, we are mainly interested in the more regular case where, as  $n \rightarrow \infty$ ,

$$\frac{1}{n} \sum_{i=1}^n d_i \rightarrow \mu, \quad \frac{1}{n} \sum_{i=1}^n d_i^2 \rightarrow \mu_2 \quad (1.2)$$

for some  $\mu, \mu_2 \in (0, \infty)$ . Obviously, (1.2) implies (1.1). Conversely, if (1.1) holds, then there is always a subsequence satisfying (1.2). It follows, see Section 4.4 for details, that for our purposes (1.1) and (1.2) are essentially equivalent. We will thus use the more general (1.1) in the theorems.

In some applications, the random multigraph  $G^*(n, \mathbf{d})$  may be at least as good as the simple graph  $G(n, \mathbf{d})$ . For example, this may be the case in an application where the random graph is intended to be an approximation of an unknown graph “in real life”; then the multigraph model may be just as good as an approximation. On the other hand, if we, as is often the case, really want a random simple graph (i.e., no loops or multiple edges), then there are several ways to proceed.

The standard method, at least in the sparse case studied in the present paper, is to condition  $G^*(n, \mathbf{d})$  on the event that it is a simple graph; it is a fundamental fact of the configuration model construction (implicit in [7]) that this yields a random simple graph  $G(n, \mathbf{d})$  with the uniform distribution over all graphs with the given degree sequence. This method has been very successful in many cases. In particular, under the condition (1.1) on  $\mathbf{d}$ ,

$$\liminf_{n \rightarrow \infty} \mathbb{P}(G^*(n, \mathbf{d}) \text{ is simple}) > 0, \quad (1.3)$$

see e.g. [16; 18], and then any result on convergence in probability for  $G^*(n, \mathbf{d})$  immediately transfers to  $G(n, \mathbf{d})$ . (See also Bollobás and Rioridan [8], where this method is used, with more complicated arguments, also in cases with  $\mathbb{P}(G^*(n, \mathbf{d}) \text{ is simple}) \rightarrow 0$ .) However, as is also well-known, results on convergence in distribution do not transfer so easily, and further arguments are needed. (See [21], [29], [19] for examples where this has succeeded, with more or less complicated extra arguments.)

Another method to create a simple graph from  $G^*(n, \mathbf{d})$  is to erase all loops and merge any set of parallel edges into a single edge. This creates a simple random graph, but typically its degree sequence is not exactly the given sequence  $\mathbf{d}$ . Nevertheless, this *erased configuration model* may be as useful as  $G^*(n, \mathbf{d})$  in some applications. This construction is studied in Britton, Deijfen and Martin-Löf [9] and van der Hofstad [15, Section 3], but will not be considered further in the present paper where we insist on the degree sequence being exactly  $\mathbf{d}$ .

In the present paper, we consider a different method, where we also adjust  $G^*(n, \mathbf{d})$  to make it simple, but this time we keep the degree sequence  $\mathbf{d}$  exact by using switchings instead of erasing. More precisely, we process the loops and multiple edges in  $G^*(n, \mathbf{d})$  one by one. For each such bad edge, we chose another edge, uniformly at random, and switch the endpoints of these two edges, thus replacing them by another pair of edges. See Section 3.2 for details. Assuming (1.1), this typically gives a simple graph after a single pass through the bad edges (Theorem 3.2); if not, we repeat until a simple graph is obtained. We denote the resulting graph by  $\widehat{G}(n, \mathbf{d})$  and call it the

*switched configuration model.* The idea to use switchings in this context goes back to McKay [25] for the closely related problem of counting simple graphs with a given degree sequence (assuming  $\max d_i = O(n^{1/4})$ , but not (1.1)), and was made explicit for generating  $G(n, \mathbf{d})$  by McKay and Wormald [26] (using somewhat different switchings). See the survey by Wormald [33] for further uses of switchings. Recent refinements of the method, extending it to larger classes of degree sequences by employing more types of switchings, are given in Gao and Wormald [10, 11, 12]. We will not use these recent refinements (that have been developed to handle also rather dense graphs); instead we focus on the simple case when (1.1) holds and only a few switchings are needed; we also use only the oldest and simplest types of switchings, used already by McKay [25] (called *simple switchings* in [33]). Although the switching method for this case has been known for a long time, it seems to have been somewhat neglected. Our purpose is to show that the switching method is powerful also in this case, and that it complements the conditioning method discussed above for the purpose of proving asymptotic results for  $G(n, \mathbf{d})$ .

**Remark 1.1.** From the point of view of constructing a random simple graph with given degree sequence by simulation, the standard approach using conditioning means that we sample the multigraph  $G^*(n, \mathbf{d})$ ; if it happens to be simple, we accept it, and otherwise we discard it completely and start again, repeating until a simple graph is found. (See e.g. [32].) The approach in the present paper is instead to keep most of the multigraph even when it is not simple, and resample only a few edges. The disadvantage is that the result  $\widehat{G}(n, \mathbf{d})$  is not perfectly uniformly random, but Theorem 2.1 below shows that is a good approximation, and asymptotically correct. The advantage is that  $\widehat{G}(n, \mathbf{d})$  typically does not differ much from  $G^*(n, \mathbf{d})$ , and thus we often can show estimates of the type (2.10) in Corollary 2.3 below.

**Remark 1.2.** In e.g. [26; 11; 12], an exactly uniformly distributed simple graph (i.e.,  $G(n, \mathbf{d})$ ) is constructed by combining switchings with rejection sampling, meaning that we may, with some carefully calculated probabilities, abort the construction and restart. (Cf. the conditioning method where, as discussed in Remark 1.1, we restart as soon as anything is wrong, instead of trying to fix it by switchings.) Our focus is not on actual concrete construction of instances of  $G(n, \mathbf{d})$  by simulation, but rather to have a method of construction that can be used theoretically to study properties of  $G(n, \mathbf{d})$ , and for our purposes the approximate uniformity given by Theorem 2.1 is good enough. (And better, since the method is simpler.)

**Remark 1.3.** Switchings have also recently been used (in a different way) by Athreya and Yogeshwaran [1] to prove asymptotic normality for statistics of  $G^*(n, \mathbf{d})$  (in a subcritical case) using martingale methods.

**Remark 1.4.** For convenience, we state the results for a sequence  $\mathbf{d}_n$  of degree sequences where  $\mathbf{d}_n$  has length (number of vertices)  $n$ . More generally, one might consider a subsequence, or other sequences of degree sequences  $\mathbf{d}_j$  with lengths  $n_j \rightarrow \infty$ . This will be used in the proofs, see Section 4.4.

The main results are stated in Section 2, and proved in Sections 4–7. A few applications are given in Section 8.

## 2. NOTATION AND MAIN RESULTS

**2.1. Some notation.** Unspecified limits are as  $n \rightarrow \infty$ ; w.h.p. (with high probability) means with probability tending to 1 as  $n \rightarrow \infty$ .  $\xrightarrow{d}$  and  $\xrightarrow{p}$  denote convergence in distribution and probability, respectively.

If  $X_n$  are random variables and  $a_n$  are positive numbers, then  $X_n = O_p(a_n)$  means  $\lim_{K \rightarrow \infty} \sup_n \mathbb{P}(|X_n| > Ka_n) = 0$ , and  $X_n = o_p(a_n)$  means  $\sup_n \mathbb{P}(|X_n| > \varepsilon a_n) = 0$  for every  $\varepsilon > 0$ ; thus  $X_n = o_p(a_n) \iff X_n/a_n \xrightarrow{p} 0$ .

Given a degree sequence  $\mathbf{d} = (d_i)_1^n$ , we let

$$d_{\max} := \max_{1 \leq i \leq n} d_i, \quad (2.1)$$

$$N := \sum_{i=1}^n d_i. \quad (2.2)$$

Thus a graph with degree sequence  $\mathbf{d}$  has  $n$  vertices and  $N/2$  edges. Note that (1.1) implies  $N = \Theta(n)$  and  $d_{\max} = O(n^{1/2})$ .

If  $\mathcal{S}$  is a measurable space, then  $\mathcal{M}(\mathcal{S})$  is the Banach space of finite signed measures on  $\mathcal{S}$ , and  $\mathcal{P}(\mathcal{S})$  is the subset of probability measures. If  $\lambda, \nu \in \mathcal{P}(\mathcal{S})$ , then their *total variation distance* is defined by

$$d_{\text{TV}}(\lambda, \nu) := \sup_{A \subseteq \mathcal{S}} |\lambda(A) - \nu(A)| = \frac{1}{2} \|\lambda - \nu\|_{\mathcal{M}(\mathcal{S})} \quad (2.3)$$

(where we tacitly only consider measurable  $A$ ). If  $X$  and  $Y$  are random elements of  $\mathcal{S}$  with distributions  $\lambda$  and  $\nu$ , we also write

$$d_{\text{TV}}(X, Y) := d_{\text{TV}}(\lambda, \nu) = \sup_{A \subseteq \mathcal{S}} |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|. \quad (2.4)$$

If  $\mathcal{S}$  is e.g. a separable metric space (for example, as in our applications, a discrete finite or countable set), then

$$d_{\text{TV}}(X, Y) = \min \mathbb{P}(X' \neq Y'), \quad (2.5)$$

taking the minimum over all *couplings*  $(X', Y')$  of  $X$  and  $Y$ , i.e., pairs of random variables  $X', Y'$  (defined on the same probability space) such that  $X' \stackrel{d}{=} X$  and  $Y' \stackrel{d}{=} Y$ . (See e.g. [4, Appendix A.1] or [17, Section 4].)

If  $\mathcal{S}_n$ ,  $n \geq 1$ , is a sequence of measurable spaces, and  $X_n$  and  $Y_n$  are random variables with values in  $\mathcal{S}_n$ , then  $X_n$  and  $Y_n$  are *contiguous* if for any sequence of measurable sets (events)  $\mathcal{E}_n \subseteq \mathcal{S}_n$ ,

$$\mathbb{P}(X_n \in \mathcal{E}_n) \rightarrow 0 \iff \mathbb{P}(Y_n \in \mathcal{E}_n) \rightarrow 0. \quad (2.6)$$

See e.g. [23, Section 9.6] and [17].

If  $G$  is a (multi)graph, we let  $E(G)$  denote its edge set and  $e(G) := |E(G)|$  its number of edges (counted with multiplicity).

$P_k$  denotes a path with  $k$  edges and  $k + 1$  vertices, and  $C_k$  a cycle with  $k$  vertices,  $k \geq 1$ . In particular,  $C_1$  is a loop, and  $C_2$  is a pair of parallel edges. We denote the disjoint union of (unlabelled) graphs by  $+$ , and write e.g.  $2P_2$  for  $P_2 + P_2$ .

$C$  and  $c$  denote positive constants that may be different at each occurrence. (They typically depend on the sequence of degree sequences, but they do not depend on  $n$ .)

**2.2. Main results.**  $\widehat{G}(n, \mathbf{d})$  is, by construction, a random simple graph with the given degree sequence  $\mathbf{d}$ . However, it does not have a uniform distribution over all such graphs, i.e., it will not be equal to the desired random graph  $G(n, \mathbf{d})$ ; see Example 3.5. Nevertheless, our main result is the following theorem, which says in a strong form that  $\widehat{G}(n, \mathbf{d})$  has asymptotically the same distribution as  $G(n, \mathbf{d})$ ; in the notation of [17],  $\widehat{G}(n, \mathbf{d})$  and  $G(n, \mathbf{d})$  are *asymptotically equivalent*. Hence,  $\widehat{G}(n, \mathbf{d})$  is a useful approximation of  $G(n, \mathbf{d})$ , and as stated formally in Corollary 2.2 below, results on both convergence in probability and convergence in distribution that can be proved for  $\widehat{G}(n, \mathbf{d})$  transfer to  $G(n, \mathbf{d})$ . Proofs are given in Section 6.

**Theorem 2.1.** *Assume that  $\mathbf{d} = (d_i^{(n)})_1^n$  depends on  $n$  and satisfies the conditions (1.1) and*

$$d_{\max} = o(n^{1/2}). \quad (2.7)$$

*Then, as  $n \rightarrow \infty$ ,*

$$d_{\text{TV}}(\widehat{G}(n, \mathbf{d}), G(n, \mathbf{d})) \rightarrow 0. \quad (2.8)$$

*In other words, there exists a coupling of  $\widehat{G}(n, \mathbf{d})$  and  $G(n, \mathbf{d})$  such that*

$$\mathbb{P}(\widehat{G}(n, \mathbf{d}) \neq G(n, \mathbf{d})) \rightarrow 0. \quad (2.9)$$

**Corollary 2.2.** *Assume that  $\mathbf{d}$  satisfies (1.1) and (2.7). Suppose that  $X_n = f_n(G(n, \mathbf{d}))$  for some function  $f_n$  of labelled simple graphs, and let  $\widehat{X}_n = f_n(\widehat{G}(n, \mathbf{d}))$ . If  $\tau$  is a constant such that  $\widehat{X}_n \xrightarrow{\text{P}} \tau$  as  $n \rightarrow \infty$ , then also  $X_n \xrightarrow{\text{P}} \tau$ . More generally, if  $Y$  is a random variable such that  $\widehat{X}_n \xrightarrow{\text{d}} Y$  as  $n \rightarrow \infty$ , then also  $X_n \xrightarrow{\text{d}} Y$ .*

Moreover,  $\widehat{G}(n, \mathbf{d})$  is obtained from  $G^*(n, \mathbf{d})$  using only a few switchings. Hence it is often easy to prove the estimate (2.10) below, and then the next corollary shows that results on convergence in distribution for  $G^*(n, \mathbf{d})$  transfer to  $\widehat{G}(n, \mathbf{d})$ , using  $\widehat{G}(n, \mathbf{d})$  as an intermediary in the proof.

**Corollary 2.3.** *Assume that  $\mathbf{d}$  satisfies (1.1) and (2.7). Suppose that  $X_n = f_n(G(n, \mathbf{d}))$  for some function  $f_n$ , which is defined more generally for labelled multigraphs, and let  $X_n^* = f_n(G^*(n, \mathbf{d}))$ . Suppose also that*

$$f_n(\widehat{G}(n, \mathbf{d})) - f_n(G^*(n, \mathbf{d})) \xrightarrow{\text{P}} 0. \quad (2.10)$$

*If  $Y$  is a random variable such that  $X_n^* \xrightarrow{\text{d}} Y$  as  $n \rightarrow \infty$ , then also  $X_n \xrightarrow{\text{d}} Y$ .*

We show in Example 3.7 that the condition  $\max_i d_i = o(n^{1/2})$  is needed in Theorem 2.1 and its corollaries above. However, we will also show the following weaker statement without this assumption. The proof is given in Section 7.

**Theorem 2.4.** *Assume that  $\mathbf{d} = (d_i^{(n)})_1^n$  depends on  $n$  and satisfies (1.1). Then, as  $n \rightarrow \infty$ , the random graphs  $\widehat{G}(n, \mathbf{d})$  and  $G(n, \mathbf{d})$  are contiguous. In other words, any sequence of events  $\mathcal{E}_n$  that holds w.h.p. for  $\widehat{G}(n, \mathbf{d})$  holds also w.h.p. for  $G(n, \mathbf{d})$ , and conversely.*

### 3. THE CONSTRUCTION OF $\widehat{G}(n, \mathbf{d})$

**3.1. The configuration model.** The well-known *configuration model* was introduced by Bollobás [7] to generate a random multigraph with a given degree sequence  $\mathbf{d} = (d_i)_1^n$ . ( $N := \sum_i d_i$  is assumed to be even.) The construction works by assigning a set of  $d_i$  *half-edges* to each vertex  $i$ ; this gives a total of  $N$  half-edges. A perfect matching of the half-edges is called a *configuration*, and defines a multigraph in the obvious way: each pair of half-edges in the matching is regarded as an edge in the multigraph. We say that the configuration *projects* to a multigraph. We choose a configuration uniformly at random, and let  $G^*(n, \mathbf{d})$  be the corresponding multigraph.

We denote the half-edges at a vertex  $i$  by  $i_1, \dots, i_{d_i}$ .

Note that the mapping from configurations to multigraphs is not injective, since we may permute the half-edges at each vertex without changing the multigraph. Nevertheless, we often informally identify a configuration and the corresponding multigraph, and we use graph theory language for configurations too. In particular, a pair  $\{i_\alpha, j_\beta\}$  in a configuration  $\Gamma$  is called an edge in  $\Gamma$ , with endpoints  $i$  and  $j$ , and may be written  $i_\alpha j_\beta$ ; similarly, the particular case  $\{i_\alpha, i_\beta\}$  is called a loop, two pairs (edges)  $\{i_\alpha, j_\beta\}$  and  $\{i_\gamma, j_\delta\}$  are said to be parallel; a configuration is simple if it has no loops or parallel edges.

**3.2. The switched configuration model.** We construct the switched configuration model by first constructing a random configuration  $\Gamma_0$  and the corresponding multigraph  $G^*(n, \mathbf{d})$  as above. Formally, we will do the switchings in the configuration, where all edges are uniquely labelled; they induce corresponding switchings in the multigraph, and informally we may think of the multigraph only.

We say that an edge in a configuration or multigraph is *bad* if it is a loop or if it is parallel to another edge. If there is no bad edge in  $\Gamma_0$  (or equivalently, in  $G^*(n, \mathbf{d})$ ), then  $G^*(n, \mathbf{d})$  is simple, and we accept it as it is. Otherwise, we choose a bad edge in  $\Gamma_0$ , say  $i_\alpha j_\beta$  (where  $i = j$  in the case of a loop), and choose another edge  $k_\gamma \ell_\delta$  uniformly at random among all other edges in  $\Gamma_0$ ; we also order the two half-edges  $k_\gamma$  and  $\ell_\delta$  randomly. We then make a *switching*, and replace the two edges  $i_\alpha j_\beta$  and  $k_\gamma \ell_\delta$  by the new edges  $i_\alpha \ell_\delta$  and  $k_\gamma j_\beta$ . This gives a new configuration  $\Gamma_1$  on the same set of half-edges, and thus a new multigraph  $G_1$  that still has the same degree sequence  $\mathbf{d}$ . Moreover, we have removed one bad edge (in the case of parallel edges, also another bad edge may have become good); however, it is possible that we have created a new bad edge (or several). If the new configuration  $\Gamma_1$  has no bad edge, then the corresponding multigraph  $G_1$  is simple and we stop; otherwise we pick a bad edge in  $\Gamma_1$ , and repeat until we obtain a simple graph.  $\widehat{G}(n, \mathbf{d})$  is defined to be the simple random graph we have when we terminate.

**Remark 3.1.** The description above is somewhat incomplete, since we have not specified which bad edge we switch, if there is more than one. We assume that we have some fixed rule for this, e.g. the lexicographically first bad edge, or a random one; different rules may yield somewhat different

final distributions, and thus formally different random graphs  $\widehat{G}(n, \mathbf{d})$ , see Example 3.6, but our results hold for any such rule. See Lemma 4.2.

Of course, if the degree sequence  $\mathbf{d}$  is not graphic, i.e., no simple graph with this degree sequence exists, then the switching process will never terminate. We conjecture that if the sequence  $\mathbf{d}$  is graphic, then the switching process almost surely terminates, but we leave this as an open problem, see Remark 3.4. However, we show in Section 4 the following theorem, which is enough for our purposes; it shows that assuming (1.1), the process w.h.p. terminates very quickly.

**Theorem 3.2.** *Assume (1.1). Then, during the construction of  $\widehat{G}(n, \mathbf{d})$ , w.h.p. no new bad edges are created and the process terminates after  $O_p(1)$  switchings.*

**Remark 3.3.** It is easily seen that if we switch a bad edge  $i_\alpha j_\beta$  with an edge  $k_\gamma \ell_\delta$  that has a vertex in common with  $i_\alpha j_\beta$ , i.e.,  $\{i, j\} \cap \{k, \ell\} \neq \emptyset$ , then there will always be a new bad edge created. It is therefore reasonable to modify the construction by choosing the edge  $k_\gamma \ell_\delta$  uniformly at random among all edges vertex-disjoint from the bad edge  $i_\alpha j_\beta$ . (Provided this is possible, which it is e.g. if the maximum degree is  $< N/4$ , as is the case for all large  $n$  when (1.1) holds.)

Theorem 3.2 implies that assuming (1.1), w.h.p. we never switch two edges with a common vertex; hence the modified version will w.h.p. yield exactly the same result  $\widehat{G}(n, \mathbf{d})$ , and consequently Theorems 2.1 and 2.4 holds for the modified construction too.

Example 3.5 shows that the modified construction does not yield exactly the same distribution of  $\widehat{G}(n, \mathbf{d})$ , nor the uniform distribution.

**Remark 3.4.** Theorem 3.2 shows that, under our conditions, the switching process w.h.p. terminates with a simple graph after a finite number of switchings. A different question is whether it always terminates, for a given  $n$  and a graphic degree sequence  $\mathbf{d}$ . First, it is easy to see that the process might loop and never terminate, even using the modification in Remark 3.3, see Example 3.8; however, in that example at least, this has probability 0. Hence the right question is whether the process terminates a.s. (i.e., with probability 1).

It can be shown, see Sjöstrand [30], that there always exists a sequence of switchings leading to a simple graph. It follows that if we choose the bad edge to switch at random, then the process terminates a.s. with a simple graph. (Note that the switching process is a finite-state Markov process, where the simple graphs are absorbing states.) We conjecture that the same holds for any rule choosing the bad edge to switch, but this remains an open problem.

For completeness, if the switching process does not terminate, we define  $\widehat{G}(n, \mathbf{d})$  by restarting with a new random configuration. (This makes no difference for our results.)

### 3.3. Examples.

**Example 3.5.** We consider a small example, both to illustrate the construction and to show that it does not yield perfect uniformity.

Let  $n = 6$  and the degree sequence  $\mathbf{d} = (2, 2, 1, 1, 1, 1)$ . Thus  $N = 8$  and there are  $N/2 = 4$  edges. There are  $7!! = 105$  different configurations of 5 different isomorphism types, as shown in Figure 1. Of these configurations, 72 yield simple graphs: 48 of the type  $P_3 + P_1$  and 24 of the type  $2P_2$ . The random simple graph  $G(6, \mathbf{d})$  thus has the distribution

$$\mathbb{P}(G(6, \mathbf{d}) \cong P_3 + P_1) = \frac{2}{3}, \quad \mathbb{P}(G(6, \mathbf{d}) \cong 2P_2) = \frac{1}{3}. \quad (3.1)$$

The remaining 33 configurations yield non-simple graphs: 24  $C_1 + P_2 + P_1$ , 6  $C_2 + 2P_1$  and 3  $2C_1 + 2P_1$ . It is easily seen that modifying a graph  $C_1 + P_2 + P_1$  by switching the bad edge and another one, randomly chosen, gives  $P_3 + P_1$  or  $2P_2$  with the same probabilities  $\frac{2}{3}$  and  $\frac{1}{3}$  as in (3.1), while  $C_2 + 2P_1$  may give  $P_3 + P_1$  but never  $2P_2$  (it may also give  $2C_1 + 2P_1$  or  $C_2 + 2P_1$  again if we switch the two parallel edges with each other; then further switchings are needed); the final possibility  $2C_1 + 2P_1$  will give  $C_1 + P_2 + P_1$  or  $C_2 + 2P_1$  after the first switching. It follows that if we continue until we have a simple graph  $\widehat{G}(n, \mathbf{d})$ , then the probability of it being  $P_3 + P_1$  is strictly larger than  $\frac{2}{3}$ ; hence  $\widehat{G}(n, \mathbf{d})$  and  $G(n, \mathbf{d})$  do not have the same distribution. An elementary but uninteresting calculation shows that for this example,

$$\mathbb{P}(\widehat{G}(6, \mathbf{d}) \cong P_3 + P_1) = \frac{24}{35}, \quad \mathbb{P}(\widehat{G}(6, \mathbf{d}) \cong 2P_2) = \frac{11}{35}. \quad (3.2)$$

The same holds also if we modify the construction by always switching with an edge disjoint from the bad one (see Remark 3.3), although the exact probabilities will be different:  $\frac{31}{45}$  and  $\frac{14}{45}$ .

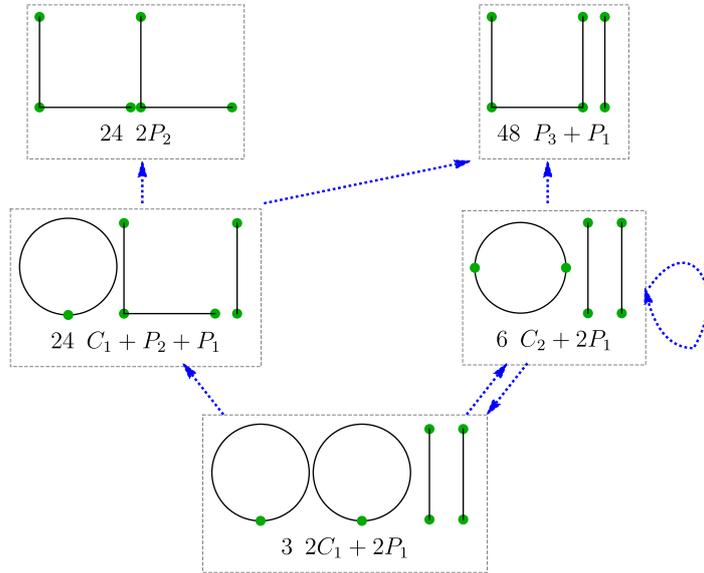


FIGURE 1. Multigraphs (unlabelled) given by the degree sequence  $(2, 2, 1, 1, 1, 1)$  in Example 3.5.

**Example 3.6.** As another example, let  $n = 7$  and  $\mathbf{d} = (2, 2, 2, 1, 1, 1, 1)$ ; thus there are  $N/2 = 5$  edges. Suppose that a realization of the configuration

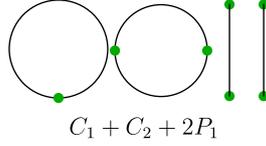


FIGURE 2. A multigraph given by the degree sequence  $(2, 2, 2, 1, 1, 1, 1)$  in Example 3.6.

model yields the multigraph  $C_1 + C_2 + 2P_1$  in Figure 2. There are three bad edges: one loop and two parallel edges.

If we choose to first switch the loop, then with probability  $\frac{1}{2}$  we switch it with one of the parallel edges, yielding the simple graph  $C_3 + 2P_1$ . If instead we switch the loop with one of the isolated edges, the result is  $C_2 + P_2 + P_1$ , which after a second switching yields either a simple graph  $P_4 + P_1$  or  $P_3 + P_2$ , or (if we switch the two parallel edges with each other)  $C_2 + P_2 + P_1$  again or  $2C_1 + P_2 + P_1$ ; in the latter cases, we obtain  $P_4 + P_1$  or  $P_3 + P_2$  after further switchings. Hence, the probability that the final graph contains a cycle  $C_3$  is  $\frac{1}{2}$ .

On the other hand, if we begin by switching one of the parallel edges, then there are three possibilities:

- (i) With probability  $\frac{1}{4}$ , we switch with the loop, and obtain again  $C_3 + 2P_1$ .
- (ii) With probability  $\frac{1}{2}$  we switch with an isolated edge and obtain  $C_1 + P_3 + P_1$  and after a second switching either  $P_4 + P_1$  or  $P_3 + P_2$ .
- (iii) With probability  $\frac{1}{4}$  we switch the two parallel edges with each other, yielding either (with probability  $\frac{1}{8}$  each) the same multigraph  $C_1 + C_2 + 2P_1$  and we restart, or  $3C_1 + 2P_1$ . In the latter case, we next switch a loop with either (probability  $\frac{1}{2}$  each) another loop, yielding  $C_1 + C_2 + 2P_1$  and we restart, or with an isolated edge, yielding  $2C_1 + P_2 + P_1$ , which eventually yields either  $P_4 + P_1$  or  $P_3 + P_2$ .

Summing up this case, we start again with a graph  $C_1 + C_2 + 2P_1$  with probability  $\frac{3}{16}$ , and thus the total probability that we end with case (i) is  $\frac{1}{4} + \frac{3}{16} = \frac{4}{13}$ .

Consequently, conditioned on this realisation of  $G^*(7, \mathbf{d})$ , the probability that the final graph  $\widehat{G}(7, \mathbf{d})$  has a cycle  $C_3$  is  $\frac{1}{2}$  or  $\frac{4}{13}$ , depending on our choice for the first switching. This shows that the order of the switchings matters. However, Theorem 2.1 is valid in any case, and thus such choices make no difference asymptotically.

With the modification in Remark 3.3, we never switch two parallel edges with each other so some possibilities disappear in this example; the final probabilities are  $\frac{1}{2}$  and  $\frac{1}{3}$ , but the conclusion remains the same.

**Example 3.7.** Fix  $a > 0$ , consider for simplicity only even  $n$ , and let  $\mathbf{d} := (m, m, 1, \dots, 1)$ , where  $m := \lfloor \sqrt{an} \rfloor$ . Thus all vertices except 1 and 2 have degree 1. Note that (1.1) holds, but not (2.7).

Let  $L_1$  and  $L_2$  be the numbers of loops at 1 and 2, and let  $M_{12}$  be the number of edges  $12$  in the multigraph  $G^*(n, \mathbf{d})$ . Note that besides these edges,  $G(n, \mathbf{d})$  contains  $m - 2L_j - M_{12}$  edges from  $j$  to a leaf ( $j = 1, 2$ ),

and a perfect matching of all remaining vertices. In particular, there are  $n/2 - O(n^{1/2})$  isolated edges.

The multigraph  $G^*(n, \mathbf{d})$  is simple if  $L_1 = L_2 = 0$  and  $M_{12} \leq 1$ . It is easy to see, e.g. by the method of moments, that asymptotically,  $L_1 \xrightarrow{d} \text{Po}(a/2)$ ,  $L_2 \xrightarrow{d} \text{Po}(a/2)$ , and  $M_{12} \xrightarrow{d} \text{Po}(a)$ , jointly with independent limits. When we construct  $\widehat{G}(n, \mathbf{d})$  by switchings, there is thus only  $O_p(1)$  bad edges; moreover, w.h.p. each switching will be with one of the  $n/2 - O(n^{1/2})$  isolated edges. In this case, no new bad edge is created by the switchings, and we reach a simple graph  $\widehat{G}(n, \mathbf{d})$  after  $L_1 + L_2 + (M_{12} - 1)_+$  switchings; furthermore, no edge 12 is created by the switchings. It follows that w.h.p.  $\widehat{G}(n, \mathbf{d})$  has an edge 12 if and only if  $M_{12} \geq 1$ . Consequently,

$$\mathbb{P}(12 \in E(\widehat{G}(n, \mathbf{d}))) = \mathbb{P}(M_{12} \geq 1) + o(1) \rightarrow \mathbb{P}(\text{Po}(a) \geq 1) = 1 - e^{-a}. \quad (3.3)$$

On the other hand, a simple graph with degree sequence  $\mathbf{d}$  has either

- (i) no edge 12 and  $m$  edges from each of 1 and 2 to leaves  $k \geq 3$ , together with a perfect matching of the remaining  $n - 2m - 2$  vertices.
- (ii) an edge 12 and  $m - 1$  edges from each of 1 and 2 to leaves  $k \geq 3$ , together with a perfect matching of the remaining  $n - 2m$  vertices.

Let the numbers of graphs of these two types be  $N_0$  and  $N_1$ . Then

$$N_0 = \binom{n-2}{m} \binom{n-2-m}{m} (n-2m-3)!! \quad (3.4)$$

$$N_1 = \binom{n-2}{m-1} \binom{n-1-m}{m-1} (n-2m-1)!! \quad (3.5)$$

and a simple calculation yields

$$\frac{N_1}{N_0} = \frac{m^2}{n-2m} \rightarrow a. \quad (3.6)$$

Hence,

$$\mathbb{P}(12 \in E(G(n, \mathbf{d}))) = \frac{N_1}{N_0 + N_1} \rightarrow \frac{a}{1+a}. \quad (3.7)$$

Comparing (3.3) and (3.7), we see that the limits differ, and thus Theorem 2.1 does not hold for this example. Similarly, Corollaries 2.2 and 2.3 fail, for example if  $f_n(G)$  is the indicator of the event that the multigraph  $G$  contains an edge where both endpoints have degrees  $\geq 2$ . This example shows that the condition (2.7) cannot be omitted from Theorem 2.1 and its corollaries.

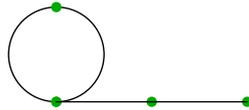


FIGURE 3. A multigraph given by the degree sequence  $(3, 2, 1, 1)$  in Example 3.8.

**Example 3.8.** Let  $\mathbf{d} = (3, 2, 2, 1)$ , and suppose that the initial multigraph  $G^*(4, \mathbf{d})$  has edges 12, 12, 13, 34, see Figure 3. If we switch one of the parallel

edges 12 with 34, then we may get a simple graph, but we may also create another edge 13 and get the edge set 12, 13, 13, 24. The latter multigraph is isomorphic to the original one, and we may continue and cycle between these two multigraphs for ever. Hence, there is no deterministic guarantee that the switching process always leads to a simple graph. Note also the modification in Remark 3.3 does not help; we still can make the same switchings.

However, note that in this example, switching the same edges but in different orientations yields a simple graph. Since we order the half-edges at random when switching, the infinite sequence of switchings above has probability 0; more generally, it is easily verified that for this example, a.s. the process terminates after a finite number of switchings.

#### 4. THE DISTRIBUTION OF $\widehat{G}(n, \mathbf{d})$

**4.1. More notation.** Let  $S$  be number of switchings used in the construction. Let  $\Gamma_k$  ( $0 \leq k \leq S$ ) be the configuration after  $k$  switchings, and let  $G_k$  be the corresponding multigraph. Thus  $G^*(n, \mathbf{d}) = G_0$  and  $\widehat{G}(n, \mathbf{d}) = G_S$ .

Let  $\mathcal{B}_k$  be the set of endpoints of bad edges in  $G_k$ , and let  $\mathcal{A}_k$  be the set of their neighbours in  $G_k$ .

Let  $b_k$  be the bad edge in  $\Gamma_{k-1}$  chosen for the  $k$ th switching, and let  $e_k$  be the (random) other edge used in that switching.

An  $m$ -edge (in a graph or configuration) is a set of  $m$  parallel edges that are not loops, and such that there are no further edges parallel to them. (I.e., the multiplicity of the edge equals  $m$ .) Let  $L$  be the number of loops in  $\Gamma_0$  (i.e., in  $G_0 = G^*(n, \mathbf{d})$ ), and let  $M_m$  ( $m \geq 2$ ) be the number of  $m$ -edges. Furthermore, let

$$M := \sum_{m \geq 2} \binom{m}{2} M_m, \quad (4.1)$$

the number of pairs of parallel edges in  $\Gamma_0$ .

Let  $\mathfrak{G}_n$  be the set of all simple graphs on  $[n]$  with degree sequence  $\mathbf{d}$ . Let  $\lambda \in \mathcal{P}(\mathfrak{G}_n)$  be the distribution of  $G(n, \mathbf{d})$ , i.e., the uniform distribution on  $\mathfrak{G}_n$ , and let  $\widehat{\lambda} \in \mathcal{P}(\mathfrak{G}_n)$  be the distribution of  $\widehat{G}(n, \mathbf{d})$ .

We sometimes tacitly assume that  $n$  (and thus  $N$ ) is large enough to avoid trivialities (such as division by 0).

**4.2. Silver and golden.** We say that the construction of  $\widehat{G}(n, \mathbf{d})$  is *silver* if

- (S1) No new bad edge is created during the construction.
- (S2) No additional edge  $e_k$  used for a switching has an endpoint in  $\mathcal{B}_0$ .

The construction is *golden* if it is silver and furthermore

- (G1)  $G^*(n, \mathbf{d})$  has no triple edges. I.e.,  $M_m = 0$  for  $m \geq 3$ .
- (G2) The loops and double edges in  $G^*(n, \mathbf{d})$  are vertex-disjoint.
- (G3) The additional edges  $e_k$  used for the switchings are vertex-disjoint with each other.

Let  $\mathcal{S}$  and  $\mathcal{G}$  be the events that the construction is silver or golden, respectively, and let  $\mathcal{S}^c$  and  $\mathcal{G}^c$  be their complements. Furthermore, let  $\mathcal{S}_s := \mathcal{S} \cap \{S = s\}$  and  $\mathcal{G}_s := \mathcal{G} \cap \{S = s\}$  ( $s \geq 0$ ).

In a silver construction, each  $m$ -edge is reduced to a single (good) edge by  $m - 1$  switchings, and thus

$$S = L + \sum_{m \geq 2} (m - 1)M_m. \quad (4.2)$$

In particular, in a golden construction, recalling (4.1),

$$S = L + M_2 = L + M. \quad (4.3)$$

In a silver construction, (4.2) and (4.1) yield the inequality

$$S \leq L + M. \quad (4.4)$$

**Lemma 4.1.** (i) *If (1.1) holds, then the construction of  $\widehat{G}(n, \mathbf{d})$  is w.h.p. silver.*

(ii) *If (1.1) and (2.7) hold, then the construction of  $\widehat{G}(n, \mathbf{d})$  is w.h.p. golden.*

*Proof.* (i): A well-known simple calculation shows that, assuming (1.1),

$$\mathbb{E} L = \sum_i \frac{d_i(d_i - 1)}{2(N - 1)} \leq \frac{1}{N} \sum_{i=1}^n d_i(d_i - 1) \leq \frac{1}{N} \sum_{i=1}^n d_i^2 = O(1), \quad (4.5)$$

$$\mathbb{E} M = \sum_{i < j} \frac{d_i(d_i - 1)d_j(d_j - 1)}{2(N - 1)(N - 3)} \leq C \left( \frac{1}{N} \sum_{i=1}^n d_i(d_i - 1) \right)^2 = O(1). \quad (4.6)$$

Hence,  $\mathbb{E}(L + M) = O(1)$ , i.e., there is a constant  $C$  such that

$$\mathbb{E}(L + M) \leq C. \quad (4.7)$$

Fix a large integer  $K$ , and assume that  $L + M \leq K$ . Then  $|\mathcal{B}_0| \leq L + 2M \leq 2K$ . Let  $k \geq 0$ , and suppose that the construction has been silver for the first  $k$  switchings (in the obvious sense). Thus no new bad edges have been created and hence  $\mathcal{B}_k \subseteq \mathcal{B}_0$ . By the Cauchy–Schwarz inequality and (1.1), the number of half-edges belonging to vertices in  $\mathcal{B}_0$  is

$$\sum_{i \in \mathcal{B}_0} d_i \leq \left( |\mathcal{B}_0| \sum_{i \in \mathcal{B}_0} d_i^2 \right)^{1/2} \leq (2K)^{1/2} \left( \sum_{i=1}^n d_i^2 \right)^{1/2} = O(n^{1/2}), \quad (4.8)$$

and thus

$$|\mathcal{A}_k| \leq \sum_{i \in \mathcal{B}_k} d_i \leq \sum_{i \in \mathcal{B}_0} d_i = O(n^{1/2}). \quad (4.9)$$

Furthermore, the number of half-edges belonging to vertices in  $\mathcal{A}_k$  is, by the argument in (4.8) together with (4.9),

$$\sum_{i \in \mathcal{A}_k} d_i \leq \left( |\mathcal{A}_k| \sum_{i=1}^n d_i^2 \right)^{1/2} = O(|\mathcal{A}_k|^{1/2} n^{1/2}) = O(n^{3/4}). \quad (4.10)$$

It follows from (4.8) and (4.10) that when we pick a random edge  $e_{k+1}$  for the next switching, the probability that it has an endpoint in  $\mathcal{B}_0$  or  $\mathcal{A}_k$  is  $O(n^{-1/4}) = o(1)$ . Hence, w.h.p. we switch with an edge  $e_{k+1}$  not having any endpoint in  $\mathcal{B}_k \cup \mathcal{A}_k$ , and it is easy to see that then no new bad edge is created. Furthermore, w.h.p.  $e_{k+1}$  has no endpoint in  $\mathcal{B}_0$ . Consequently, w.h.p. the construction remains silver for the  $(k + 1)$ th switching too. Since

only  $L + M \leq K$  switchings are needed, it follows by induction that w.h.p. the construction is silver until the end.

We have shown that for every fixed  $K$ ,  $\mathbb{P}(\mathcal{S}^c \cap \{L + M \leq K\}) \rightarrow 0$ . Hence, using also Markov's inequality and (4.7),

$$\begin{aligned} \mathbb{P}(\mathcal{S}^c) &\leq \mathbb{P}(\mathcal{S}^c \cap \{L + M \leq K\}) + \mathbb{P}(L + M > K) \leq o(1) + \frac{\mathbb{E}(L + M)}{K} \\ &\leq o(1) + \frac{C}{K}. \end{aligned} \quad (4.11)$$

Thus  $\limsup_{n \rightarrow \infty} \mathbb{P}(\mathcal{S}^c) \leq C/K$ . Since  $K$  is arbitrary,  $\mathbb{P}(\mathcal{S}^c) \rightarrow 0$ .

(ii): The expected number of triples of parallel edges in  $G^*(n, \mathbf{d})$  is at most, using (1.1) and (2.7),

$$\sum_{i < j} d_i^3 d_j^3 \frac{C}{N^3} \leq C \frac{d_{\max}^2}{N^3} \left( \sum_i d_i^2 \right)^2 \leq C \frac{d_{\max}^2}{n} = o(1). \quad (4.12)$$

Hence, (G1) holds w.h.p.

Similar calculations show that the expected number of pairs of 2 loops, a loop and a double edge, or 2 double edges, are  $o(1)$ . Hence, (G2) holds w.h.p.

Finally, fix  $k \geq 1$ . Given  $e_1, \dots, e_{k-1}$ , these have (at most)  $2(k-1)$  endpoints. There are at most  $2(k-1)d_{\max}$  edges with an endpoint in this set. Since  $e_k$  is drawn at random among the  $N/2 - 1$  edges distinct from  $b_k$ , the probability that  $e_k$  is not vertex-disjoint from  $e_1, \dots, e_{k-1}$  is at most  $2(k-1)d_{\max}/(N/2 - 1) = o(1)$ . Note also that if the construction is silver and  $L + M \leq K$ , then at most  $K$  switchings are done by (4.4). It follows that for any fixed  $K$ ,

$$\mathbb{P}(\mathcal{G}^c \cap \mathcal{S} \cap \{L + M \leq K\}) = o(1). \quad (4.13)$$

The argument in (and after) (4.11) shows that  $\mathbb{P}(\mathcal{G}^c \cap \mathcal{S}) \rightarrow 0$ . Hence, using also part (i),  $\mathbb{P}(\mathcal{G}^c) \leq \mathbb{P}(\mathcal{G}^c \cap \mathcal{S}) + \mathbb{P}(\mathcal{S}^c) \rightarrow 0$ .  $\square$

*Proof of Theorem 3.2.* Immediate from Lemma 4.1(i), since a silver construction creates no new bad edges by (S1) and uses  $S \leq L + M$  switchings by (4.4), so  $\mathbb{E}S \leq \mathbb{E}(L + M) = O(1)$  by (4.7) and thus  $S = O_p(1)$ .  $\square$

**4.3. The choice of a bad edge.** As said in Remark 3.1, the random graph  $\widehat{G}(n, \mathbf{d})$  may depend on the (unspecified) rule for choosing the bad edge for each switching. However, all rules yield asymptotically the same result, at least provided (1.1) holds.

**Lemma 4.2.** *Assume (1.1). Let  $\widehat{G}_1(n, \mathbf{d})$  and  $\widehat{G}_2(n, \mathbf{d})$  be created by using two different rules for choosing the bad edge for each switching. Then  $d_{\text{TV}}(\widehat{G}_1(n, \mathbf{d}), \widehat{G}_2(n, \mathbf{d})) \rightarrow 0$ .*

*Proof.* Suppose that we have a silver construction of  $\widehat{G}_1(n, \mathbf{d})$ ; then only edges that are bad already in  $\Gamma_0$  will be switched. It follows that we may couple the two constructions of  $\widehat{G}(n, \mathbf{d})$ , starting with the same  $\Gamma_0$ , such that if the construction of  $\widehat{G}_1(n, \mathbf{d})$  is silver, then, in the sequence of graphs  $G_0, \dots, G_S = \widehat{G}(n, \mathbf{d})$ , exactly the same switchings are made in both constructions, although perhaps in different order; consequently, the

two constructions yield the same  $\widehat{G}(n, \mathbf{d})$ . (On the level of configurations, the switchings may differ, because of the choice of one parallel edge out of several.) Consequently, for this coupling, by Lemma 4.1(i),

$$\mathbb{P}(\widehat{G}_1(n, \mathbf{d}) \neq \widehat{G}_2(n, \mathbf{d})) = o(1). \quad (4.14)$$

This proves the lemma by (2.5).  $\square$

Lemma 4.2 implies that if Theorem 2.1 or Theorem 2.4 holds for some rule, then it holds for any rule. We may thus for the proofs below assume that we each time choose the bad edge in  $\Gamma_k$  that is first according to the following order.

- (B1) First the loops, in lexicographic order.
- (B2) The  $m$ -edges in lexicographic order of their endpoints, and for each  $m$ -edge its  $m$  edges in  $\Gamma_k$  in lexicographic order.

The exact definition of the lexicographic order in these cases is left to the reader. In fact, any fixed order would do.

**4.4. The subsubsequence principle.** Next we note that it suffices to prove that Theorems 2.1 and 2.4 always hold for some subsequence. This is a general argument, which we repeat for convenience: Suppose that Theorem 2.1 fails; then there exists a sequence  $\mathbf{d}_n$  satisfying the assumptions and such that (2.8) fails; thus there exists  $\varepsilon > 0$  and a subsequence such that  $d_{\text{TV}}(\widehat{G}(n, \mathbf{d}), G(n, \mathbf{d})) \geq \varepsilon$  for every  $n$  in the subsequence. But by assumption we can find a subsubsequence such that Theorem 2.1 holds, a contradiction. The proof for Theorem 2.4 is essentially the same.

In particular, assuming (1.1), by selecting a suitable subsequence we may in the remainder of the proofs assume that (1.2) holds for some  $\mu > 0$  and  $\mu_2 < \infty$ . Let  $\nu := \mu_2 - \mu$ ; then (1.2) implies  $N/n \rightarrow \mu$  and

$$\frac{1}{n} \sum_{i=1}^n d_i(d_i - 1) \rightarrow \nu, \quad \frac{1}{N} \sum_{i=1}^n d_i(d_i - 1) \rightarrow \frac{\nu}{\mu}. \quad (4.15)$$

If  $\nu = 0$ , then (4.15) and (4.5)–(4.6) show that  $\mathbb{E}L \rightarrow 0$  and  $\mathbb{E}M \rightarrow 0$ . Consequently, w.h.p.  $L = M = 0$ , so  $G^*(n, \mathbf{d})$  is simple, in which case  $\widehat{G}(n, \mathbf{d}) = G^*(n, \mathbf{d})$ ; thus  $d_{\text{TV}}(\widehat{G}(n, \mathbf{d}), G^*(n, \mathbf{d})) \rightarrow 0$ . Furthermore, by the fact that  $G(n, \mathbf{d})$  has the same distribution as  $G^*(n, \mathbf{d})$  conditioned on being simple, we can couple  $G(n, \mathbf{d})$  and  $G^*(n, \mathbf{d})$  such that they are equal when  $G^*(n, \mathbf{d})$  is simple; thus by (2.5),

$$d_{\text{TV}}(G(n, \mathbf{d}), G^*(n, \mathbf{d})) \leq \mathbb{P}(G^*(n, \mathbf{d}) \text{ is not simple}) \rightarrow 0. \quad (4.16)$$

Hence Theorems 2.1 and 2.4 follow trivially when  $\nu = 0$ . Consequently, in the proofs we may assume  $\nu > 0$ .

**4.5. Silver constructions.** Consider a silver construction. Each switching of a loop creates a copy of  $\mathbf{P}_2$ , and each set of  $m - 1$  switchings of an  $m$ -edge creates  $m - 1$  copies of  $\mathbf{P}_3$  having a common middle edge, which is the one edge remaining of the original  $m$  parallel ones. Colour the created copies of  $\mathbf{P}_2$  and  $\mathbf{P}_3$  *red*; these are regarded as (not necessarily disjoint) subgraphs of the configuration. Note that the non-leaves in the red paths belong to  $\mathcal{B}_0$ , while a leaf is an endpoint of some  $e_k$  and thus, by (S2), lies outside  $\mathcal{B}_0$ . By

(S2), the edges in the red paths will not be used by later switchings, and thus the red paths remain as subgraphs of  $\Gamma_S$ , and thus of  $G_S = \widehat{G}(n, \mathbf{d})$ .

The red paths do not have to be vertex-disjoint. However, by the remarks above and (S1)–(S2), the set  $\mathcal{R}$  of red paths in  $\Gamma_S$ , or equivalently in  $G_S = \widehat{G}(n, \mathbf{d})$ , has the following properties.

- (P1) Each red path has length 2 or 3, i.e., is a copy of  $P_2$  or  $P_3$ .
- (P2) The red paths are edge-disjoint, except that several red paths  $P_3$  may share the same middle edge.
- (P3) An leaf of a red path is not a non-leaf of another red path.

Define the *gap* of a red path as its pair of endpoints. This is the pair of endpoints of the edge  $e_k$  used to create this red path. By (S1) and (S2),  $e_k$  was a good edge in  $\Gamma_0$ , i.e., there was no parallel edge in  $\Gamma_0$ . Thus the red paths have also the properties:

- (P4) The gaps of the red paths are distinct pairs of vertices.
- (P5) The gaps of the red paths are non-edges in  $G_S$ .

Furthermore, by (B2):

- (P6) If a red  $P_3$  is given by the edges  $i_\alpha j_\beta, j_\gamma k_\delta, k_\varepsilon \ell_\zeta$  in  $\Gamma_S$ , then necessarily the edge  $j_\gamma k_\delta$  comes after  $j_\beta k_\varepsilon$  in the lexicographic order.

Conversely, in a silver construction, the red paths in  $\Gamma_S$  determine precisely the switchings that have been made; hence they together with  $\Gamma_S$  determine the initial configuration  $\Gamma_0$  and also, by (B1)–(B2), the order of the switchings. Moreover, given any simple configuration  $\Gamma$  (on the given set of half-edges) with corresponding graph  $G$  and a set of red paths in  $\Gamma$  (or  $G$ ) satisfying (P1)–(P6) (with obvious notational changes here and below:  $\Gamma_S$  is replaced by  $\Gamma$  and  $G_S$  by  $G$ ), there exists a unique initial configuration  $\Gamma_0$  and a unique silver sequence of switchings, satisfying (B1)–(B2), that yields  $\Gamma_S = \Gamma$  with the given red paths. Each such history with a given number  $S$  of switchings has the same probability.

Consequently, dropping “red”:

**Claim 1.** *For a fixed  $s \geq 0$ , the conditional probability  $\mathbb{P}(\Gamma_S = \Gamma \mid \mathcal{S}_s)$  is proportional to the number of sets of  $s$  paths in  $\Gamma$  that satisfy (P1)–(P6).*

We project to graphs. For a simple graph  $G$  and a configuration  $\Gamma$  projecting to  $G$ , there is an obvious bijection between sets of paths in  $G$  and sets of paths in  $\Gamma$ . Note that the conditions (P1)–(P5) depend only on the paths in  $G$ , while (P6) depends also on the specific configuration  $\Gamma$ .

Given a simple graph  $G \in \mathfrak{G}_n$ , and a set  $\mathcal{R}$  of  $s$  paths in  $G$  that satisfy (P1)–(P5), let the *weight*  $w(\mathcal{R}; G)$  be the probability that the lifting of  $\mathcal{R}$  to paths in a configuration  $\Gamma$ , chosen uniformly at random among all configuration  $\Gamma$  that project to  $G$ , satisfies also (P6). Furthermore, for  $G \in \mathfrak{G}_n$  and  $s \geq 0$ , let  $\zeta_{S,s}(G)$  be the sum of the weights of all sets of  $s$  paths in  $G$  that satisfy (P1)–(P5).

Recall that each simple graph  $G \in \mathfrak{G}_n$  is the projection of the same number  $A := \prod_i d_i!$  of configurations. It follows that the number of pairs  $(\Gamma, \mathcal{R})$  where  $\Gamma$  is a configuration projecting to  $G$  and  $\mathcal{R}$  is a set of  $s$  paths in  $\Gamma$  satisfying (P1)–(P6) equals  $\sum_{\mathcal{R}} w(\mathcal{R}; G)A = A\zeta_{S,s}(G)$ . Consequently, Claim 1 implies:

**Claim 2.** For a fixed  $s \geq 0$ , and all  $G \in \mathfrak{G}_n$ , the conditional probability  $\mathbb{P}(\widehat{G}(n, \mathbf{d}) = G \mid \mathcal{S}_s)$  is proportional to  $\zeta_{\mathcal{S},s}(G)$ .

Let  $\widehat{\lambda}_{\mathcal{S},s}$  be the distribution of  $\widehat{G}(n, \mathbf{d})$  conditioned on  $\mathcal{S}_s$ , i.e.,

$$\widehat{\lambda}_{\mathcal{S},s}\{G\} = \mathbb{P}(\widehat{G}(n, \mathbf{d}) = G \mid \mathcal{S}_s). \quad (4.17)$$

Claim 2 says that the probability  $\widehat{\lambda}_{\mathcal{S},s}(G)$  is proportional to  $\zeta_{\mathcal{S},s}(G)$ , and thus to  $\zeta_{\mathcal{S},s}(G)\lambda\{G\}$ , recalling that  $\lambda$  is the uniform distribution. To find the normalizing constant, recall that  $\lambda$  is the distribution of  $G(n, \mathbf{d})$ , which implies, using the notation  $Z_{\mathcal{S},s} := \zeta_{\mathcal{S},s}(G(n, \mathbf{d}))$ ,

$$\sum_{G \in \mathfrak{G}_n} \zeta_{\mathcal{S},s}(G)\lambda\{G\} = \mathbb{E}[\zeta_{\mathcal{S},s}(G(n, \mathbf{d}))] = \mathbb{E} Z_{\mathcal{S},s}. \quad (4.18)$$

Consequently, since  $\widehat{\lambda}_{\mathcal{S},s}$  is a probability measure,

$$\widehat{\lambda}_{\mathcal{S},s}\{G\} = \frac{\zeta_{\mathcal{S},s}(G)}{\mathbb{E} Z_{\mathcal{S},s}} \lambda\{G\}. \quad (4.19)$$

**4.6. Golden constructions.** For the proof of Theorem 2.1, we simplify and consider only golden constructions. In a golden construction, it follows from (G1)–(G3) that the red paths are vertex-disjoint. Conversely, a silver construction yielding vertex-disjoint red paths is golden.

It follows that Claims 1 and 2 above hold also if we replace  $\mathcal{S}_s$  by  $\mathcal{G}_s$  and consider only sets  $\mathcal{R}$  of vertex-disjoint paths, so  $\zeta_{\mathcal{S},s}(G)$  is replaced by  $\zeta_{\mathcal{G},s}(G)$ , defined as the total weight of all sets of  $s$  vertex-disjoint paths in  $G$  that satisfy (P1)–(P5). Thus, in analogy to (4.17)–(4.19), letting  $\widehat{\lambda}_{\mathcal{G},s}$  be the distribution of  $\widehat{G}(n, \mathbf{d})$  conditioned on  $\mathcal{G}_s$ , and  $Z_{\mathcal{G},s} := \zeta_{\mathcal{G},s}(G(n, \mathbf{d}))$ ,

$$\widehat{\lambda}_{\mathcal{G},s}\{G\} := \mathbb{P}(\widehat{G}(n, \mathbf{d}) = G \mid \mathcal{G}_s) = \frac{\zeta_{\mathcal{G},s}(G)}{\mathbb{E} Z_{\mathcal{G},s}} \lambda\{G\}. \quad (4.20)$$

Hence,

$$\begin{aligned} \|\widehat{\lambda}_{\mathcal{G},s} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} &= \sum_{G \in \mathfrak{G}_n} |\widehat{\lambda}_{\mathcal{G},s}\{G\} - \lambda\{G\}| = \sum_{G \in \mathfrak{G}_n} \left| \frac{\zeta_{\mathcal{G},s}(G) - \mathbb{E} Z_{\mathcal{G},s}}{\mathbb{E} Z_{\mathcal{G},s}} \right| \lambda\{G\} \\ &= \frac{\mathbb{E} |Z_{\mathcal{G},s} - \mathbb{E} Z_{\mathcal{G},s}|}{\mathbb{E} Z_{\mathcal{G},s}}. \end{aligned} \quad (4.21)$$

This will be studied in the following sections. We first find the weights  $w(\mathcal{R}; G)$ .

**Lemma 4.3.** Suppose that  $G \in \mathfrak{G}_n$  and that  $\mathcal{R}$  consists of  $\ell \geq 0$  paths  $P_2$  and  $m \geq 0$  paths  $P_3$  in  $G$ , all vertex-disjoint. Then the weight  $w(\mathcal{R}; G) = 2^{-m}$ .

*Proof.* Given any configuration projecting to  $G$ , we obtain all other such configurations by permuting the half-edges at each vertex. Hence, if  $H \cong P_3$  is a path in  $G$ , then it follows by symmetry, permuting only the half-edges at the two central vertices of  $H$ , that the probability is  $\frac{1}{2}$  that (P6) holds for the lift of  $H$  to a random configuration  $\Gamma$  that projects to  $G$ . Furthermore, for disjoint  $H_1, \dots, H_m$ , the corresponding events are independent. The result follows.  $\square$

For  $\ell, m \geq 0$  and  $G \in \mathfrak{G}_n$ , let  $\zeta_{\ell,m}(G)$  be the number of sets  $\{F_1, \dots, F_\ell, H_1, \dots, H_m\}$  of vertex-disjoint paths in  $G$  such that each  $F_i \cong P_2$  and each  $H_j \cong P_3$ , and (P5) holds. (We ignore the order of  $F_1, \dots, F_\ell$  and  $H_1, \dots, H_m$ .) Note that (P2)–(P4) holds for any set of vertex-disjoint paths. Hence, Lemma 4.3 yields

$$\zeta_{\mathcal{G},s}(G) = \sum_{\ell+m=s} 2^{-m} \zeta_{\ell,m}(G). \quad (4.22)$$

**Remark 4.4.** The proof of Lemma 4.3 is easily extended to the more general sets of paths in Section 4.5. If the paths  $P_3$  in  $\mathcal{R}$  are grouped according to their middle edges, with  $m_k$  groups of  $k-1$  paths having the same middle edge (and thus coming from switchings of a  $k$ -edge in  $\Gamma_0$ ),  $k \geq 2$ , then  $w(\mathcal{R}; G) = \prod_k k^{-m_k}$ . We will not use this formula, and omit the details.

### 5. SOME SUBGRAPH COUNTS IN $G^*(n, \mathbf{d})$

The equations (4.21) and (4.22) show that it suffices to show good estimates for the special subgraph counts  $\zeta_{\ell,m}(G(n, \mathbf{d}))$ . In order to do so, we use the standard method to study the random multigraph  $G^*(n, \mathbf{d})$  instead (cf. Section 1). Let  $Z_{\ell,m}^* := \zeta_{\ell,m}(G^*(n, \mathbf{d}))$ .

For two multigraphs  $H$  and  $G$ , let  $\psi_H(G)$  be the number of subgraphs of  $G$  isomorphic to  $H$ . Define

$$X_H := \psi_H(G^*(n, \mathbf{d})). \quad (5.1)$$

Note that  $L = X_{C_1}$  and  $M = X_{C_2}$ . We are mainly interested in the case when  $H = P_2$  or  $P_3$ , and we write  $X_k := X_{P_k}$ .

We begin with an estimate that does not require (2.7).

**Lemma 5.1.** *Assume that  $\mathbf{d}$  satisfies (1.1). Then, for every fixed  $\ell, m \geq 0$ ,*

$$\mathbb{E}(X_2^\ell X_3^m) = O(n^{\ell+m}). \quad (5.2)$$

*Proof.* First, deterministically using (1.1), since there are at most  $\binom{d_i}{2}$  copies of  $P_2$  with middle vertex  $i$ ,

$$X_2 \leq \sum_{i=1}^n \binom{d_i}{2} = O(n). \quad (5.3)$$

We estimate  $X_3$  too from above by overcounting. For  $i, j \in [n]$  and  $\alpha \in [d_i]$ ,  $\beta \in [d_j]$ , let  $I_{i,\alpha,j,\beta}$  be the indicator of the event that the half-edges  $i_\alpha$  and  $j_\beta$  form an edge. Then

$$X_3 \leq \widehat{X}_3 := \sum_{i < j} \sum_{\alpha=1}^{d_i} \sum_{\beta=1}^{d_j} (d_i - 1)(d_j - 1) I_{i,\alpha,j,\beta}. \quad (5.4)$$

We show, by induction on  $m$ , that for every fixed  $m \geq 0$ ,

$$\mathbb{E} \widehat{X}_3^m = O(n^m). \quad (5.5)$$

This is trivial for  $m = 0$ , and for  $m = 1$  we have (cf. the similar (4.6))

$$\mathbb{E} \widehat{X}_3 = \sum_{i < j} \sum_{\alpha=1}^{d_i} \sum_{\beta=1}^{d_j} (d_i - 1)(d_j - 1) \mathbb{E} I_{i,\alpha,j,\beta} = \sum_{i < j} \frac{d_i(d_i - 1)d_j(d_j - 1)}{N - 1}$$

$$\leq \frac{C}{n} \left( \sum_{i=1}^n d_i(d_i - 1) \right)^2 = O(n). \quad (5.6)$$

For the induction step, note that (5.4) yields the expansion

$$\mathbb{E} \widehat{X}_3^m = \sum_{i_1 < j_1, \dots, i_m < j_m} \sum_{\alpha_1, \dots, \alpha_m} \sum_{\beta_1, \dots, \beta_m} \mathbb{E} \prod_{k=1}^m (d_{i_k} - 1)(d_{j_k} - 1) I_{i_k, \alpha_k, j_k, \beta_k}. \quad (5.7)$$

First, let  $1 \leq k < \ell \leq m$ , and consider all terms in (5.7) where  $(i_k, \alpha_k, j_k, \beta_k) = (i_\ell, \alpha_\ell, j_\ell, \beta_\ell)$ . In these terms,  $I_{i_\ell, \alpha_\ell, j_\ell, \beta_\ell}$  is redundant, and  $(d_{i_\ell} - 1)(d_{j_\ell} - 1) < d_{\max}^2 \leq Cn$ , and eliminating these factors yields  $\mathbb{E} \widehat{X}_3^{m-1}$ . Hence, the induction hypothesis shows that the contribution of these terms is at most  $Cn \mathbb{E} \widehat{X}_3^{m-1} = O(n^m)$ .

Summing over all pairs  $(k, \ell)$  still yields  $O(n^m)$ .

The remaining terms in (5.7) are those where the  $m$  quadruples  $(i_k, \alpha_k, j_k, \beta_k)$  are distinct. In this case, either

$$\mathbb{E} \prod_{k=1}^m I_{i_k, \alpha_k, j_k, \beta_k} = \frac{1}{(N-1)(N-3) \cdots (N-2m+1)} \quad (5.8)$$

$$\leq \frac{C}{(N-1)^m} = C \prod_{k=1}^m \mathbb{E} I_{i_k, \alpha_k, j_k, \beta_k}, \quad (5.9)$$

(where  $C = C_m$  depends on  $m$ ), or  $\prod_{k=1}^m I_{i_k, \alpha_k, j_k, \beta_k} = 0$  identically because of conflicts. Consequently, the sum of these terms in (5.7) is at most

$$\sum_{i_1 < j_1, \dots, i_m < j_m} \sum_{\alpha_1, \dots, \alpha_m} \sum_{\beta_1, \dots, \beta_m} C \prod_{k=1}^m (d_{i_k} - 1)(d_{j_k} - 1) \mathbb{E} I_{i_k, \alpha_k, j_k, \beta_k} = C (\mathbb{E} \widehat{X}_3)^m. \quad (5.10)$$

By (5.6), this too is  $O(n^m)$ , which completes the induction step and proves (5.5). The result follows by (5.3) and (5.5).  $\square$

**Lemma 5.2.** *Assume that  $\mathbf{d}$  satisfies (1.2) and (2.7). Then*

$$\frac{X_2}{n} \xrightarrow{\mathbb{P}} \frac{\nu}{2}, \quad (5.11)$$

$$\frac{X_3}{n} \xrightarrow{\mathbb{P}} \frac{\nu^2}{2\mu}. \quad (5.12)$$

*Proof.* First, the overcount in (5.3) comes from the loops and multiple edges, and we have the estimate

$$0 \leq \sum_{i=1}^n \binom{d_i}{2} - X_2 \leq 2d_{\max} X_{C_1} + 2X_{C_2} = 2d_{\max} L + 2M. \quad (5.13)$$

By (4.5)–(4.6),  $L, M = O_{\mathbb{P}}(1)$ , and thus (5.13) yields, using (2.7),

$$\frac{X_2}{n} = \frac{1}{n} \sum_{i=1}^n \binom{d_i}{2} + o_{\mathbb{P}}(n^{-1/2}) = \frac{1}{n} \sum_{i=1}^n \binom{d_i}{2} + o_{\mathbb{P}}(1). \quad (5.14)$$

Hence, (5.11) follows by (4.15).

For  $X_3$ , we consider again  $\widehat{X}_3$  defined in (5.4). This too is defined by overcounting, and it is easily seen that

$$0 \leq \widehat{X}_3 - X_3 \leq 2d_{\max}^2 X_{C_1} + 4d_{\max} X_{C_2} + 3X_{C_3}. \quad (5.15)$$

Again,  $X_{C_1} = L = O_p(1)$  and  $X_{C_2} = M = O_p(1)$  by (4.5)–(4.6), and a similar calculation shows  $\mathbb{E} X_{C_3} = O(1)$  and thus  $X_{C_3} = O_p(1)$ . Furthermore,  $d_{\max} = o(n^{1/2})$  by (2.7). Hence, (5.15) implies

$$\widehat{X}_3 - X_3 = o_p(n). \quad (5.16)$$

Consequently, it suffices to consider  $\widehat{X}_3$ . We have, by (5.6),

$$\mathbb{E} \widehat{X}_3 = \frac{1}{2(N-1)} \left( \left( \sum_{i=1}^n d_i(d_i-1) \right)^2 - \sum_{i=1}^n d_i^2(d_i-1)^2 \right). \quad (5.17)$$

Furthermore, by (2.7) and (1.1),

$$\sum_{i=1}^n d_i^2(d_i-1)^2 \leq d_{\max}^2 \sum_{i=1}^n d_i^2 = o(n^2). \quad (5.18)$$

Using (5.18) and (4.15) in (5.17) yields

$$\frac{\mathbb{E} \widehat{X}_3}{n} \rightarrow \frac{\nu^2}{2\mu}. \quad (5.19)$$

Finally, we estimate the variance of  $\widehat{X}_3$ . We use again the representation (5.4). We have the covariances

$$\text{Cov}(I_{i,\alpha,j,\beta}, I_{k,\gamma,\ell,\delta}) = \begin{cases} \frac{1}{N-1} - \left(\frac{1}{N-1}\right)^2, & \{i_\alpha, j_\beta\} = \{k_\gamma, \ell_\delta\}, \\ -\left(\frac{1}{N-1}\right)^2, & |\{i_\alpha, j_\beta\} \cap \{k_\gamma, \ell_\delta\}| = 1, \\ \frac{1}{(N-1)(N-3)} - \left(\frac{1}{N-1}\right)^2, & |\{i_\alpha, j_\beta\} \cap \{k_\gamma, \ell_\delta\}| = 0. \end{cases} \quad (5.20)$$

Hence, whenever  $\{i_\alpha, j_\beta\} \neq \{k_\gamma, \ell_\delta\}$ ,

$$\text{Cov}(I_{i,\alpha,j,\beta}, I_{k,\gamma,\ell,\delta}) \leq \frac{2}{(N-1)^2(N-3)} \leq \frac{C}{n} \mathbb{E} I_{i,\alpha,j,\beta} \mathbb{E} I_{k,\gamma,\ell,\delta}, \quad (5.21)$$

and it follows from (5.4) that, using (5.6) and (2.7),

$$\begin{aligned} \text{Var} \widehat{X}_3 &\leq \sum_{i < j} \sum_{\alpha=1}^{d_i} \sum_{\beta=1}^{d_j} (d_i-1)^2 (d_j-1)^2 \mathbb{E} I_{i,\alpha,j,\beta} + \frac{C}{n} (\mathbb{E} \widehat{X}_3)^2 \\ &\leq d_{\max}^2 \mathbb{E} \widehat{X}_3 + \frac{C}{n} (\mathbb{E} \widehat{X}_3)^2 = O(nd_{\max}^2) + O(n) = o(n^2). \end{aligned} \quad (5.22)$$

Consequently,  $(\widehat{X}_3 - \mathbb{E} \widehat{X}_3)/n \xrightarrow{p} 0$ , which together with (5.19) implies

$$\frac{\widehat{X}_3}{n} \xrightarrow{p} \frac{\nu^2}{2\mu}. \quad (5.23)$$

Finally, this and (5.16) imply (5.12).  $\square$

**Remark 5.3.** Lemma 5.2 may fail without the assumption (2.7). For an example, consider again Example 3.7. Then  $\widehat{X}_3 = (m-1)^2 M_{12}$  and by an estimate similar to (5.15),  $X_3 = (m-1)^2 M_{12} + o_p(n)$ ; hence  $X_3/n = aM_{12} + o_p(1) \xrightarrow{d} a \text{Po}(a)$ . Thus  $X_3/n$  does not converge in probability to a constant.

**Remark 5.4.** Under the stronger assumption  $\sum_i d_i^m = O(n)$  for every  $m < \infty$ , it is shown in [19, Theorem 3.10] that (for example)  $X_2$  and  $X_3$  are asymptotically normal, with variance of order  $n$ . We do not know whether that holds under the weaker assumptions in Lemma 5.2.

**Lemma 5.5.** *Assume that  $\mathbf{d}$  satisfies (1.2) and (2.7). Then, for any  $\ell, m \geq 0$ ,*

$$n^{-\ell-m} Z_{\ell,m}^* \xrightarrow{p} \alpha_{\ell,m} := \frac{1}{\ell! m!} \left(\frac{\nu}{2}\right)^\ell \left(\frac{\nu^2}{2\mu}\right)^m. \quad (5.24)$$

*Proof.*  $\ell! m! Z_{\ell,m}^*$  counts ordered sequences of subgraphs  $F_1, \dots, F_\ell, H_1, \dots, H_m$  of vertex-disjoint paths in  $G^*(n, \mathbf{d})$  such that each  $F_i \cong P_2$  and each  $H_j \cong P_3$ , and (P5) holds. We may overcount and estimate this by  $X_2^\ell X_3^m$ ; we can also estimate the error by

$$\begin{aligned} 0 &\leq X_2^\ell X_3^m - \ell! m! Z_{\ell,m}^* \\ &\leq \binom{\ell}{2} A_{22} X_2^{\ell-2} X_3^m + \ell m A_{23} X_2^{\ell-1} X_3^{m-1} + \binom{m}{2} A_{33} X_2^\ell X_3^{m-2} \\ &\quad + \ell B_2 X_2^{\ell-1} X_3^m + m B_3 X_2^\ell X_3^{m-1}, \end{aligned} \quad (5.25)$$

where  $A_{jk}$  is that number of pair of paths  $F \cong P_j$  and  $F' \cong P_k$  in  $G^*(n, \mathbf{d})$  such that  $F \cap F' \neq \emptyset$ , and  $B_j$  is the number of paths  $P_j$  such that (P5) does not hold, i.e. paths  $P_j$  that are part of a cycle  $C_{j+1}$ .

We estimate  $A_{jk}$  and  $B_j$ . First, we have  $B_2 \leq 3X_{C_3}$  and  $B_3 \leq 4X_{C_4}$ . Calculations similar to (4.6) show  $\mathbb{E}X_{C_j} = O(1)$  for any fixed  $j$  (as said in the proof of Lemma 5.2 for  $j = 3$ ), and thus

$$B_2 = O_p(1), \quad B_3 = O_p(1). \quad (5.26)$$

Fix  $j, k \in \{2, 3\}$ . We make a decomposition

$$A_{jk} = \sum_{H \in \mathcal{H}} A_{jk}(H), \quad (5.27)$$

where  $A_{jk}(H)$  is the number of pairs  $(F, F')$  of paths in  $G^*(n, \mathbf{d})$  such that  $F \cap F' \neq \emptyset$ ,  $F \cong P_j$ ,  $F' \cong P_k$  and  $F \cup F' \cong H$ , and  $\mathcal{H}$  is the (finite) set of unlabelled multigraphs that can be written as a union of two paths of lengths  $j$  and  $k$ . Given  $F \cup F'$ , there is  $O(1)$  choices of  $F$  and  $F'$ , and thus

$$A_{ij}(H) \leq CX_H. \quad (5.28)$$

For  $r \geq 0$ , let

$$\Delta_r := \sum_{i=1}^n d_i^r. \quad (5.29)$$

Let  $H \in \mathcal{H}$  have  $q$  vertices with degrees  $\delta_1, \dots, \delta_q$ . Then the number of possible copies of  $H$  in  $G^*(n, \mathbf{d})$ , counted in the corresponding configuration, is at most

$$\sum_{i_1, \dots, i_q} \prod_{k=1}^q d_{i_k}^{\delta_k} = \prod_{k=1}^q \Delta_{\delta_k}, \quad (5.30)$$

and each such copy occurs with probability  $O(N^{-e(H)}) = O(n^{-e(H)})$ . Since  $e(H) = \frac{1}{2} \sum_k \delta_k$ ,

$$\mathbb{E} X_H \leq C n^{-e(H)} \prod_{k=1}^q \Delta_{\delta_k} = C \prod_{k=1}^q n^{-\delta_k/2} \Delta_{\delta_k}. \quad (5.31)$$

By (1.1),  $\Delta_1 = O(n)$  and  $\Delta_2 = O(n)$ . Furthermore, for  $\delta > 2$ ,  $\Delta_\delta \leq d_{\max}^{\delta-2} \Delta_2 \leq C d_{\max}^{\delta-2} n$ . Hence, using (2.7),

$$n^{-\delta/2} \Delta_\delta \leq \begin{cases} C n^{1/2}, & \delta = 1, \\ C, & \delta = 2, \\ C d_{\max}^{\delta-2} n^{1-\delta/2} = o(1), & \delta > 2. \end{cases} \quad (5.32)$$

Let  $h_1$  be the number of vertices in  $H$  with degree 1. Since  $H$  is a connected union of two paths,  $H$  has no isolated vertices, and it follows from (5.31) and (5.32) that  $\mathbb{E} X_H = O(n^{h_1/2})$ . Furthermore,  $h_1 \leq 4$ , and if  $h_1 = 4$ , then there is some vertex with degree  $> 2$ . If  $h_1 \leq 3$ , then  $\mathbb{E} X_H = O(n^{3/2})$ , and if  $h_1 = 4$  and some vertex has degree  $\delta > 2$ , then (5.31) and (5.32) imply  $\mathbb{E} X_H \leq C n^{4/2} (d_{\max}/n^{1/2})^{\delta-2} = o(n^2)$ .

Hence,  $\mathbb{E} X_H = o(n^2)$  for every  $H \in \mathcal{H}$ , and thus  $\mathbb{E} A_{jk}(H) = o(n^2)$  by (5.28), and finally  $\mathbb{E} A_{jk} = o(n^2)$  by (5.27), which implies

$$A_{jk} = o_p(n^2). \quad (5.33)$$

Since Lemma 5.2 implies  $X_2, X_3 \leq Cn$  w.h.p., it follows from (5.26) that the last two terms in (5.25) are  $O_p(n^{\ell+m-1})$ , and from (5.33) that the remaining terms on the right-hand side are  $o_p(n^{\ell+m})$ . Consequently, (5.25) implies

$$X_2^\ell X_3^m - \ell! m! Z_{\ell,m}^* = o_p(n^{\ell+m}) \quad (5.34)$$

and thus

$$n^{-\ell-m} Z_{\ell,m}^* = \frac{1}{\ell! m!} \left( \frac{X_2}{n} \right)^\ell \left( \frac{X_3}{n} \right)^m + o_p(1) \quad (5.35)$$

and the result (5.24) follows from Lemma 5.2.  $\square$

**Lemma 5.6.** *Assume that  $\mathbf{d}$  satisfies (1.2) and (2.7). Let  $\ell, m \geq 0$ . Then,*

$$\mathbb{E} |Z_{\ell,m}^* - \alpha_{\ell,m} n^{\ell+m}| = o(n^{\ell+m}). \quad (5.36)$$

*Proof.* Let  $Y_n := n^{-\ell-m} Z_{\ell,m}^*$ . We have  $Y_n \leq n^{-\ell-m} X_2^\ell X_3^m$  by (5.25), and thus by Lemma 5.1,

$$\mathbb{E}[Y_n^2] \leq n^{-2\ell-2m} \mathbb{E}[X_2^{2\ell} X_3^{2m}] = O(1). \quad (5.37)$$

Hence, the sequence  $Y_n$  ( $n \geq 1$ ) is uniformly integrable (see e.g. [13, Theorem 5.4.2]), and thus (5.24) implies  $L^1$ -convergence [13, Theorem 5.5.4], i.e.,

$$\mathbb{E}|Y_n - \alpha_{\ell,m}| \rightarrow 0, \quad (5.38)$$

which is equivalent to (5.36).  $\square$

## 6. PROOF OF THEOREM 2.1

We now transfer the results in Section 5 to the simple random graph  $G(n, \mathbf{d})$ .

**Lemma 6.1.** *The results in Lemmas 5.1, 5.2, 5.5, 5.6 hold also conditioned on the event that  $G^*(n, \mathbf{d})$  is simple. In other words, the corresponding results for  $G(n, \mathbf{d})$  hold as well.*

*Proof.* An immediate consequence of (1.3).  $\square$

Let  $Z_{\ell,m} := \zeta_{\ell,m}(G(n, \mathbf{d}))$ .

**Lemma 6.2.** *Assume that  $\mathbf{d}$  satisfies (1.2) and (2.7). Let  $\ell, m \geq 0$ . Then,*

$$\mathbb{E} Z_{\ell,m} = (\alpha_{\ell,m} + o(1))n^{\ell+m}, \quad (6.1)$$

$$\mathbb{E}|Z_{\ell,m} - \mathbb{E} Z_{\ell,m}| = o(n^{\ell+m}). \quad (6.2)$$

*Proof.* By Lemmas 5.6 and 6.1,

$$\mathbb{E}|Z_{\ell,m} - \alpha_{\ell,m}n^{\ell+m}| = o(n^{\ell+m}). \quad (6.3)$$

This implies (6.1), and then (6.1) and (6.3) yield (6.2).  $\square$

*Proof of Theorem 2.1.* As said in Section 4.4, we may assume that (1.2) holds and  $\nu > 0$ . Fix  $s \geq 0$ . Recalling the notations  $Z_{\mathcal{G},s} := \zeta_{\mathcal{G},s}(G(n, \mathbf{d}))$  and  $Z_{\ell,m} := \zeta_{\ell,m}(G(n, \mathbf{d}))$ , we see that (4.22) and (6.2) imply that

$$\mathbb{E}|Z_{\mathcal{G},s} - \mathbb{E} Z_{\mathcal{G},s}| = o(n^s). \quad (6.4)$$

Furthermore, (4.22) and (6.1) imply

$$\mathbb{E} Z_{\mathcal{G},s} \geq \mathbb{E} Z_{s,0} = (\alpha_{s,0} + o(1))n^s. \quad (6.5)$$

Hence, using (4.21), noting that  $\alpha_{s,0} > 0$ ,

$$\|\widehat{\lambda}_{\mathcal{G},s} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} = \frac{\mathbb{E}|Z_{\mathcal{G},s} - \mathbb{E} Z_{\mathcal{G},s}|}{\mathbb{E} Z_{\mathcal{G},s}} \rightarrow 0. \quad (6.6)$$

Let  $p_s := \mathbb{P}(\mathcal{G}_s)$ , and  $p^c := \mathbb{P}(\mathcal{G}^c)$ . Then  $p^c + \sum_s p_s = 1$ , and, recalling (4.20) and letting  $\widehat{\lambda}_{\mathcal{G}^c}$  be the distribution of  $\widehat{G}(n, \mathbf{d})$  conditioned on  $\mathcal{G}^c$ ,

$$\widehat{\lambda} = p^c \widehat{\lambda}_{\mathcal{G}^c} + \sum_{s=0}^{\infty} p_s \widehat{\lambda}_{\mathcal{G},s}. \quad (6.7)$$

Consequently, for any  $K \geq 1$ , using (4.21),

$$\begin{aligned} \|\widehat{\lambda} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} &= \left\| p^c (\widehat{\lambda}_{\mathcal{G}^c} - \lambda) + \sum_{s \geq 0} p_s (\widehat{\lambda}_{\mathcal{G},s} - \lambda) \right\|_{\mathcal{M}(\mathfrak{G}_n)} \\ &\leq p^c \|\widehat{\lambda}_{\mathcal{G}^c} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} + \sum_{s \geq 0} p_s \|\widehat{\lambda}_{\mathcal{G},s} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} \end{aligned}$$

$$\leq 2p^c + \sum_{s=0}^K \|\widehat{\lambda}_{\mathcal{G},s} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} + \sum_{s>K} 2p_s. \quad (6.8)$$

By Lemma 4.1(ii),  $p^c = o(1)$ , and by (6.6),  $\|\widehat{\lambda}_{\mathcal{G},s} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} = o(1)$  for every fixed  $s$ . Furthermore, recall that  $\mathcal{G}_s$  implies  $L + M = S = s$ , see (4.3). Hence, for any fixed  $K$ , (6.8) implies, using (4.7),

$$\begin{aligned} \|\widehat{\lambda} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} &\leq o(1) + o(1) + \sum_{s>K} 2\mathbb{P}(L + M = s) = o(1) + 2\mathbb{P}(L + M > K) \\ &\leq o(1) + \frac{2\mathbb{E}(L + M)}{K} \leq o(1) + \frac{C}{K}. \end{aligned} \quad (6.9)$$

Thus  $\limsup_{n \rightarrow \infty} \|\widehat{\lambda} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} \leq C/K$ , and then letting  $K \rightarrow \infty$  yields  $\|\widehat{\lambda} - \lambda\|_{\mathcal{M}(\mathfrak{G}_n)} \rightarrow 0$ . This completes the proof of (2.8) by (2.3). The final sentence follows by (2.5).  $\square$

*Proof of Corollary 2.2.* Using a coupling such that (2.9) holds, we have  $\mathbb{P}(X_n \neq \widehat{X}_n) \rightarrow 0$ , and the conclusion follows.  $\square$

*Proof of Corollary 2.3.* Let  $\widehat{X}_n := f_n(\widehat{G}(n, \mathbf{d}))$ . Then the assumption (2.10) says  $\widehat{X}_n - X_n^* \xrightarrow{P} 0$ . We have also assumed  $X_n^* \xrightarrow{d} Y$ , and it follows that  $\widehat{X} \xrightarrow{d} Y$ . Hence the result follows by Corollary 2.2.  $\square$

## 7. PROOF OF THEOREM 2.4

Since we do not assume (2.7), the construction is not necessarily golden w.h.p. (There may be e.g. triple edges in  $G^*(n, \mathbf{d})$ .) Hence we use in this section silver constructions. Recall the notation  $Z_{\mathcal{S},s} := \zeta_{\mathcal{S},s}(G(n, \mathbf{d}))$ .

We define, cf. (5.1), for a multigraph  $H$ ,

$$Y_H := \psi_H(G(n, \mathbf{d})). \quad (7.1)$$

**Lemma 7.1.** *Assume that  $\mathbf{d}$  satisfies (1.2) and  $\nu > 0$ . Let  $s \geq 0$ . Then,*

$$\mathbb{E} Z_{\mathcal{S},s} \geq cn^s, \quad (7.2)$$

$$\mathbb{E} Z_{\mathcal{S},s}^2 \leq Cn^{2s}, \quad (7.3)$$

*Proof.* (i): By selecting subsequences, see Section 4.4, we may assume that the limit  $b := \lim_{n \rightarrow \infty} d_{\max}/n^{1/2}$  exists. We consider two cases.

(ia):  $b = 0$ . This means  $d_{\max} = o(n^{1/2})$ , i.e., (2.7) holds; thus Lemma 6.2 and its consequence (6.5) hold. By definition,  $\zeta_{\mathcal{S},s}(G) \geq \zeta_{\mathcal{G},s}(G)$  for all  $G \in \mathfrak{G}_n$ . Hence,  $Z_{\mathcal{S},s} \geq Z_{\mathcal{G},s}$ , and (7.2) follows by (6.5).

(ib):  $b > 0$ . We may assume that  $d_{\max} = d_1$ . Then, for large  $n$ ,  $d_1 = d_{\max} \geq \frac{b}{2}n$ . Say that a half-edge at vertex 1 is *green* if it is not part of a loop or multiple edge, and let  $W$  be the number of green edges. Then  $W \geq d_1 - 2L - 2M \geq \frac{b}{2}n^{1/2} - O_p(1)$ , and thus w.h.p.  $W \geq \frac{b}{3}n^{1/2}$ . A pair of green half-edges defines a path of length 2 in  $G(n, \mathbf{d})$ , with 1 as midpoint. Hence, a sequence of  $2s$  distinct green half-edges defines  $s$  paths of length 2. This set of paths satisfies (P1)–(P4) and (trivially) (P6). It fails to satisfy (P5) only if one of the paths is part of a  $\mathcal{C}_3$ , and the number of such

sequences is  $O(Y_{C_3} W^{2s-2})$ . We have, by the usual conditioning argument with (1.3), and a simple calculation (used also in the proof of Lemma 5.2),

$$\mathbb{E} Y_{C_3} = \mathbb{E}(X_{C_3} \mid G^*(n, \mathbf{d}) \text{ is simple}) \leq C \mathbb{E} X_{C_3} = O(1). \quad (7.4)$$

Hence, w.h.p.  $Y_{C_3} \leq n^{1/2}$ . Consequently, w.h.p., crudely,

$$Z_{\mathcal{S},s} \geq \binom{W}{2s} - CW^{2s-2}Y_{C_3} \geq cb^{2s}n^s - O(n^{s-1/2}) \geq cn^s, \quad (7.5)$$

which implies (7.2).

(ii): The weights  $w(\mathcal{R}; G) \in [0, 1]$ , and thus  $\zeta_{\mathcal{S},s}(G)$  is at most the number of sets of  $s$  paths of lengths 2 or 3 in  $G$ . Hence,

$$Z_{\mathcal{S},s} = \zeta_{\mathcal{S},s}(G(n, \mathbf{d})) \leq \sum_{\ell=0}^s Y_{P_2}^\ell Y_{P_3}^{s-\ell}. \quad (7.6)$$

Thus, for fixed  $s \geq 0$ , recalling  $X_k = X_{P_k}$  in Section 5 and Lemma 5.1, and (1.3),

$$\begin{aligned} \mathbb{E} Z_{\mathcal{S},s}^2 &\leq C \sum_{\ell=0}^{2s} \mathbb{E}[Y_{P_2}^\ell Y_{P_3}^{2s-\ell}] = C \sum_{\ell=0}^{2s} \mathbb{E}(X_2^\ell X_3^{2s-\ell} \mid G^*(n, \mathbf{d}) \text{ is simple}) \\ &\leq \frac{C}{\mathbb{P}(G^*(n, \mathbf{d}) \text{ is simple})} \sum_{\ell=0}^{2s} \mathbb{E}(X_2^\ell X_3^{2s-\ell}) = O(n^{2s}). \end{aligned} \quad (7.7)$$

□

*Proof of Theorem 2.4.* Again, as said in Section 4.4, we may assume that (1.2) holds and  $\nu > 0$ .

Recall the definition (2.6), and fix an arbitrary sequence of subsets  $\mathcal{E}_n \subseteq \mathfrak{G}_n$ . (An event for these random graphs may be identified with a subset of  $\mathfrak{G}_n$ .)

(i): Suppose that

$$\mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n) \rightarrow 0. \quad (7.8)$$

The event that  $G^*(n, \mathbf{d})$  is simple is the same as  $S = 0$  (i.e., no switchings are made), and in this case  $\widehat{G}(n, \mathbf{d}) = G^*(n, \mathbf{d})$ . Hence,

$$\begin{aligned} \mathbb{P}(G(n, \mathbf{d}) \in \mathcal{E}_n) &= \mathbb{P}(G^*(n, \mathbf{d}) \in \mathcal{E}_n \mid S = 0) = \frac{\mathbb{P}(G^*(n, \mathbf{d}) \in \mathcal{E}_n \text{ and } S = 0)}{\mathbb{P}(S = 0)} \\ &= \frac{\mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n \text{ and } S = 0)}{\mathbb{P}(S = 0)} \leq \frac{\mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n)}{\mathbb{P}(S = 0)} \rightarrow 0, \end{aligned} \quad (7.9)$$

by the assumption (7.8) and (1.3).

(ii): Suppose conversely that

$$\mathbb{P}(G(n, \mathbf{d}) \in \mathcal{E}_n) \rightarrow 0. \quad (7.10)$$

Fix  $s \geq 0$  and let now  $p_s := \mathbb{P}(\mathcal{S}_s)$ . Then (4.17) yields

$$\mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n \text{ and } \mathcal{S}_s) = p_s \mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n \mid \mathcal{S}_s) = p_s \widehat{\lambda}_{\mathcal{S},s}(\mathcal{E}_n). \quad (7.11)$$

Furthermore, (4.19) implies

$$\widehat{\lambda}_{\mathcal{S},s}(\mathcal{E}_n) = \sum_{G \in \mathcal{E}_n} \frac{\zeta_{\mathcal{S},s}(G)}{\mathbb{E} Z_{\mathcal{S},s}} \lambda\{G\} = \frac{\mathbb{E}[Z_{\mathcal{S},s} \mathbf{1}\{G(n, \mathbf{d}) \in \mathcal{E}_n\}]}{\mathbb{E} Z_{\mathcal{S},s}}. \quad (7.12)$$

The Cauchy–Schwarz inequality and (7.12) yield, using (7.2)–(7.3) and the assumption (7.10),

$$\widehat{\lambda}_{\mathcal{S},s}(\mathcal{E}_n) \leq \frac{(\mathbb{E}[Z_{\mathcal{S},s}^2] \mathbb{P}(G(n, \mathbf{d}) \in \mathcal{E}_n))^{1/2}}{\mathbb{E} Z_{\mathcal{S},s}} \leq C \mathbb{P}(G(n, \mathbf{d}) \in \mathcal{E}_n)^{1/2} \rightarrow 0. \quad (7.13)$$

Hence, for every fixed  $s \geq 0$ , (7.11) and (7.13) imply

$$\mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n \text{ and } \mathcal{S}_s) \rightarrow 0. \quad (7.14)$$

We now argue similarly to the final part of the proof of Theorem 2.1. For every fixed  $K \geq 1$ , using also Lemma 4.1(i), (4.4) and (4.7),

$$\begin{aligned} \mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n) &\leq \mathbb{P}(\mathcal{S}^c) + \sum_{s=0}^{\infty} \mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n \text{ and } \mathcal{S}_s) \\ &\leq \mathbb{P}(\mathcal{S}^c) + \sum_{s=0}^K \mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n \text{ and } \mathcal{S}_s) + \sum_{s>K} \mathbb{P}(\mathcal{S}_s) \\ &= o(1) + o(1) + \mathbb{P}(\mathcal{S} \text{ and } S > K). \\ &\leq o(1) + \mathbb{P}(L + M > K) \leq o(1) + \frac{\mathbb{E}(L + M)}{K} \\ &\leq o(1) + \frac{C}{K}. \end{aligned} \quad (7.15)$$

Consequently,  $\limsup_{n \rightarrow \infty} \mathbb{P}(\widehat{G}(n, \mathbf{d}) \in \mathcal{E}_n) \leq C/K$ , and then letting  $K \rightarrow \infty$  yields (7.8).  $\square$

## 8. APPLICATIONS

Let the random variable  $D_n$  be the degree of a uniformly random vertex, and note that (1.2) can be written  $\mathbb{E} D_n \rightarrow \mu$  and  $\mathbb{E} D_n^2 \rightarrow \mu_2$ . We will in the applications below use the standard assumption that there exists a random variable  $D$  such that

$$D_n \xrightarrow{d} D. \quad (8.1)$$

We will also sometimes assume that (1.2) is strengthened to

$$\mathbb{E} D_n^2 \rightarrow \mathbb{E} D^2 < \infty. \quad (8.2)$$

Equivalently, assuming (8.1), the sequence  $D_n^2$  is uniformly integrable, see [13, Theorem 5.5.9]. Note that this implies that (1.2) holds with  $\mu = \mathbb{E} D$ . It is also easy to see that (8.1)–(8.2) imply (2.7).

**Example 8.1.** Assume (8.1)–(8.2) and  $\mathbb{P}(D = 1) > 0$ . Assume also that  $\nu - \mu = \mathbb{E} D(D - 2) > 0$ ; this is the *supercritical* case where there is w.h.p. a giant component of order  $\Theta(n)$  in both  $G^*(n, \mathbf{d})$  and  $G(n, \mathbf{d})$ , see Molloy and Reed [27, 28] with refinements in, e.g., [22], [8], [24].

Let  $|\mathcal{C}_k| = |\mathcal{C}_k(G)|$  be the order of the  $k$ th largest component in a multigraph  $G$ . It was proved by Barbour and Röllin [5] (under somewhat stronger assumptions), with a different proof in [19] (under the conditions here), that the size  $|\mathcal{C}_1|$  of the giant component is asymptotically normal for  $G^*(n, \mathbf{d})$ :

$$\frac{|\mathcal{C}_1(G^*(n, \mathbf{d}))| - \mathbb{E}|\mathcal{C}_1(G^*(n, \mathbf{d}))|}{\sqrt{n}} \xrightarrow{d} N(0, \sigma^2), \quad (8.3)$$

where the asymptotic variance  $\sigma^2$  was calculated explicitly by Ball and Neal [3].

It is shown in [19], by a non-trivial extra argument, that (8.3) holds also for  $G(n, \mathbf{d})$ . We can now replace that argument, and give a simpler proof of asymptotic normality for  $G(n, \mathbf{d})$ .

Consider the components of a multigraph as sets of vertices (ignoring the edges). A switching will either leave all components unchanged, or it will merge two components. Hence, a sequence of  $S$  switchings will change the size  $|\mathcal{C}_1|$  of the largest component by at most  $S|\mathcal{C}_2|$ ; consequently,

$$|\mathcal{C}_1(\widehat{G}(n, \mathbf{d}))| - |\mathcal{C}_1(G^*(n, \mathbf{d}))| \leq S|\mathcal{C}_2(G^*(n, \mathbf{d}))|. \quad (8.4)$$

Furthermore, under our assumptions, [8, Theorem 2] and [19, Lemma 9.4] imply  $|\mathcal{C}_2(G^*(n, \mathbf{d}))| \leq C \log n$  w.h.p., while Theorem 3.2 yields  $S = O_p(1)$ . Hence,  $S|\mathcal{C}_2(G^*(n, \mathbf{d}))| = o_p(n^{1/2})$ , and (8.4) shows that (2.10) holds for  $f_n(G) := n^{-1/2}(|\mathcal{C}_1(G)| - \mathbb{E}|\mathcal{C}_1(G^*(n, \mathbf{d}))|)$ . Consequently, Corollary 2.3 applies and shows that (8.3) implies

$$\frac{|\mathcal{C}_1(G(n, \mathbf{d}))| - \mathbb{E}|\mathcal{C}_1(G^*(n, \mathbf{d}))|}{\sqrt{n}} \xrightarrow{d} N(0, \sigma^2), \quad (8.5)$$

Furthermore, if  $X_n$  denotes the left-hand side of (8.3), then also  $\mathbb{E} X_n^2 \rightarrow \sigma^2$  [5; 19], and thus  $X_n^2$  are uniformly integrable [13, Theorem 5.5.9]. Hence, using (1.3),  $X_n^2$  are uniformly integrable also conditioned on  $G^*(n, \mathbf{d})$  being simple, and thus the mean and variance converge in (8.5) too. In particular,  $\mathbb{E}|\mathcal{C}_1(G(n, \mathbf{d}))| - \mathbb{E}|\mathcal{C}_1(G^*(n, \mathbf{d}))| = o(\sqrt{n})$ , and thus  $\mathbb{E}|\mathcal{C}_1(G^*(n, \mathbf{d}))|$  can be replaced by  $\mathbb{E}|\mathcal{C}_1(G(n, \mathbf{d}))|$  in (8.5).

**Remark 8.2.** Ball [2] has proved related results on asymptotic normality for the size of SIR epidemics on  $G^*(n, \mathbf{d})$ . As a special case, he obtains asymptotic normality of the size of the giant component for (bond or site) percolation in  $G^*(n, \mathbf{d})$  (in the supercritical case).

It seems possible that at least some of these results too can be extended to  $G(n, \mathbf{d})$  by our Corollary 2.3, but it remains to verify (2.10) for them.

**Example 8.3.** Let  $T$  be a fixed tree, and let  $n_T(G)$  be the number of components isomorphic to  $T$  in a (multi)graph  $G$ . Assume (8.1)–(8.2). Then, by Barbour and Röllin [5] (under somewhat stronger assumptions), and [19] (with a different proof)

$$\frac{n_T(G^*(n, \mathbf{d})) - \mathbb{E}n_T(G^*(n, \mathbf{d}))}{\sqrt{n}} \xrightarrow{d} N(0, \sigma_T^2), \quad (8.6)$$

for some  $\sigma_T^2 \geq 0$  (with  $\sigma_T^2 > 0$  except in some rather trivial cases). It was shown in [19] that (8.6) holds also for  $G(n, \mathbf{d})$ , again with an extra argument; we can now replace that by a simpler proof.

A switching can change at most two components, and thus  $n_T$  is changed by at most 2; hence, using Theorem 3.2,

$$|n_T(\widehat{G}(n, \mathbf{d})) - n_T(G^*(n, \mathbf{d}))| \leq 2S = O_p(1). \quad (8.7)$$

Consequently, Corollary 2.3 applies as in Example 8.1, and shows that (8.6) holds also for  $G(n, \mathbf{d})$ .

**Example 8.4.** Assume (8.1),  $\mathbb{P}(D \notin \{0, 2\}) > 0$  and (1.1). Assume also

$$\varepsilon_n := \frac{\mathbb{E} D_n(D_n - 2)}{\mathbb{E} D_n} = O(n^{-1/3}(\mathbb{E} D_n^3)^{2/3}), \quad (8.8)$$

which means that we are in the *critical window*, and that

$$d_{\max} = o(n^{1/3}(\mathbb{E} D_n^3)^{1/3}), \quad (8.9)$$

which easily is seen to imply (2.7). Then, Hatami and Molloy [14, Theorem 1.1] (under somewhat stronger conditions) showed that  $|\mathcal{C}_1|$  is of the order  $\Upsilon_n := n^{2/3}(\mathbb{E} D_n^3)^{-1/3}$ . Moreover, see Janson, van der Hofstad and Luczak [20, Theorem 2.12],  $|\mathcal{C}_1|/\Upsilon_n$  is bounded in probability, but not bounded by a constant w.h.p.; in other words:

(i) For any  $\delta > 0$  there exists  $K = K(\delta)$  such that

$$\mathbb{P}(|\mathcal{C}_1| > K\Upsilon_n) \leq \delta. \quad (8.10)$$

(ii) For any  $K < \infty$ ,

$$\liminf_{n \rightarrow \infty} \mathbb{P}(|\mathcal{C}_1| > K\Upsilon_n) > 0. \quad (8.11)$$

Both parts hold for both  $G^*(n, \mathbf{d})$  and  $G(n, \mathbf{d})$ ; however, there is a technical difference in the proofs. Part (i) is proved, by both [14] and [20] (with different methods) first for  $G^*(n, \mathbf{d})$ , and the result for  $G(n, \mathbf{d})$  then follows immediately by the standard conditioning argument and (1.3).

Part (ii) is also proved (by [20]) first for  $G^*(n, \mathbf{d})$ , but here we cannot use conditioning directly, and a rather long extra argument is given in [20, Section 6.3]. We can now replace this extra argument by Theorem 2.4.

Note first that, as said in Example 8.1, switchings can only merge components, but never break them; hence, switchings can only increase  $|\mathcal{C}_1|$ , and thus if (8.11) holds for  $G^*(n, \mathbf{d})$ , then it holds for  $\widehat{G}(n, \mathbf{d})$  too. Suppose that (8.11) fails for  $G(n, \mathbf{d})$ . Then there exists a subsequence where the probability tends to 0, and the contiguity in Theorem 2.4 shows that the same holds for  $\widehat{G}(n, \mathbf{d})$ ; a contradiction.

#### ACKNOWLEDGEMENT

I thank Xing Shi Cai for making the pictures.

#### REFERENCES

- [1] Siva Athreya & D. Yogeshwaran: Central limit theorem for statistics of subcritical configuration models. Preprint, 2018. [arXiv:1808.06778](#)
- [2] Frank Ball: Central limit theorems for SIR epidemics and percolation on configuration model random graphs. Preprint, 2018. [arXiv:1812.03105](#)

- [3] Frank Ball & Peter Neal: The asymptotic variance of the giant component of configuration model random graphs. *Ann. Appl. Probab.* **27** (2017), no. 2, 1057–1092. MR 3655861
- [4] A. D. Barbour, Lars Holst & Svante Janson: *Poisson Approximation*. Oxford University Press, Oxford, 1992. MR 1163825
- [5] A. D. Barbour & Adrian Röllin: Central limit theorems in the configuration model. Preprint, 2017. [arXiv:1710.02644](https://arxiv.org/abs/1710.02644)
- [6] Edward A. Bender & E. Rodney Canfield: The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A* **24** (1978), no. 3, 296–307. MR 0505796
- [7] Béla Bollobás: A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** (1980), no. 4, 311–316. MR 0595929
- [8] Béla Bollobás & Oliver Riordan: An old approach to the giant component problem. *J. Combin. Theory Ser. B* **113** (2015), 236–260. MR 3343756
- [9] Tom Britton, Maria Deijfen & Anders Martin-Löf: Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.*, **124** (2006), no. 6, 1377–1397. MR 2266448
- [10] Pu Gao & Nicholas Wormald: Enumeration of graphs with a heavy-tailed degree sequence. *Adv. Math.* **287** (2016), 412–450. MR 3422681
- [11] Pu Gao & Nicholas Wormald: Uniform generation of random regular graphs. *SIAM J. Comput.* **46** (2017), no. 4, 1395–1427. MR 3686817
- [12] Pu Gao & Nicholas Wormald: Uniform generation of random graphs with power-law degree sequences. Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, 1741–1758, SIAM, Philadelphia, PA, 2018. MR 3775902
- [13] Allan Gut: *Probability: A Graduate Course*, 2nd ed. Springer, New York, 2013. MR 2977961
- [14] Hamed Hatami & Michael Molloy: The scaling window for a random graph with a given degree sequence. *Random Structures Algorithms* **41** (2012), no. 1, 99–123. MR 2943428
- [15] Remco van der Hofstad: *Random Graphs and Complex Networks. Vol. 1*. Cambridge University Press, Cambridge, 2017. MR 3617364
- [16] Svante Janson: The probability that a random multigraph is simple. *Combin. Probab. Comput.* **18** (2009), no. 1-2, 205–225. MR 2497380
- [17] Svante Janson: Asymptotic equivalence and contiguity of some random graphs. *Random Structures Algorithms* **36** (2010), 26–45. MR 2591045
- [18] Svante Janson: The probability that a random multigraph is simple. II. *J. Appl. Probab.* *51A* (2014), 123–137. MR 3317354
- [19] Svante Janson: Asymptotic normality in random graphs with given vertex degrees. Preprint, 2018. [arXiv:1812.08063](https://arxiv.org/abs/1812.08063)
- [20] Svante Janson, Remco van der Hofstad & Malwina Łuczak: Component structure of the configuration model: barely supercritical case. *Random Structures Algorithms*, to appear.
- [21] Svante Janson & Malwina Łuczak: Asymptotic normality of the  $k$ -core in random graphs. *Ann. Appl. Probab.* **18** (2008), no. 3, 1085–1137. MR 2418239

- [22] Svante Janson & Malwina Łuczak: A new approach to the giant component problem. *Random Structures Algorithms* **34** (2009), no. 2, 197–216. MR 2490288
- [23] Svante Janson, Tomasz Łuczak & Andrzej Ruciński: *Random Graphs*. Wiley, New York, 2000. MR 1782847
- [24] Felix Joos, Guillem Perarnau, Dieter Rautenbach & Bruce Reed: How to determine if a random graph with a fixed degree sequence has a giant component. *Probab. Theory Related Fields* **170** (2018), no. 1-2, 263–310. MR 3748325
- [25] Brendan D. McKay: Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combin.* **19A** (1985), 15–25. MR 0790916
- [26] Brendan D. McKay & Nicholas C. Wormald: Uniform generation of random regular graphs of moderate degree. *J. Algorithms* **11** (1990), no. 1, 52–67. MR 1041166
- [27] Michael Molloy & Bruce Reed: A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6** (1995), no. 2-3, 161–179. MR 1370952
- [28] Michael Molloy & Bruce Reed: The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.* **7** (1998), no. 3, 295–305. MR 1664335
- [29] Oliver Riordan: The phase transition in the configuration model. *Combin. Probab. Comput.* **21** (2012), no. 1-2, 265–299. MR 2900063
- [30] Jonas Sjöstrand: Making a multigraph simple by a sequence of double edge swaps. In preparation.
- [31] Nicholas C. Wormald: The asymptotic distribution of short cycles in random regular graphs. *J. Combin. Theory Ser. B* **31** (1981), no. 2, 168–182. MR 0630980
- [32] Nicholas C. Wormald: Generating random regular graphs. *J. Algorithms* **5** (1984), no. 2, 247–280. MR 0744493
- [33] Nicholas C. Wormald: Models of random regular graphs. *Surveys in Combinatorics 1999 (Canterbury)*, 239–298, London Math. Soc. Lecture Note Ser., 267, Cambridge Univ. Press, Cambridge, 1999. MR 1725006

DEPARTMENT OF MATHEMATICS, UPPSALA UNIVERSITY, PO BOX 480, SE-751 06 UPPSALA, SWEDEN

*E-mail address:* `svante.janson@math.uu.se`

*URL:* `http://www.math.uu.se/svante-janson`