

# CELEBRATING 50 YEARS OF THE APPLIED PROBABILITY TRUST

Edited by  
S. ASMUSSEN, P. JAGERS, I. MOLCHANOV and L. C. G. ROGERS

Part 4. Random graphs and particle systems

## THE PROBABILITY THAT A RANDOM MULTIGRAPH IS SIMPLE. II

SVANTE JANSON, *Uppsala University*

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden.

Email address: [svante.janson@math.uu.se](mailto:svante.janson@math.uu.se)



APPLIED PROBABILITY TRUST  
DECEMBER 2014

# THE PROBABILITY THAT A RANDOM MULTIGRAPH IS SIMPLE. II

BY SVANTE JANSON

---

## Abstract

Consider a random multigraph  $G^*$  with given vertex degrees  $d_1, \dots, d_n$ , constructed by the configuration model. We give a new proof of the fact that, asymptotically for a sequence of such multigraphs with the number of edges  $\frac{1}{2} \sum_i d_i \rightarrow \infty$ , the probability that the multigraph is simple stays away from 0 if and only if  $\sum_i d_i^2 = O(\sum_i d_i)$ . The new proof uses the method of moments, which makes it possible to use it in some applications concerning convergence in distribution. Corresponding results for bipartite graphs are included.

*Keywords:* Configuration model; random multigraph; random bipartite graph

2010 Mathematics Subject Classification: Primary 05C80; 05C30; 60C05

---

## 1. Introduction

Let  $G(n, (d_i)_1^n)$  be the random (simple) graph with vertex set  $[n] := \{1, \dots, n\}$  and vertex degrees  $d_1, \dots, d_n$ , uniformly chosen among all such graphs. (We assume that such graphs exist; in particular,  $\sum_i d_i$  has to be even.) A standard method of studying  $G(n, (d_i)_1^n)$  is to consider the related random labelled multigraph  $G^*(n, (d_i)_1^n)$  defined by taking a set of  $d_i$  half-edges at each vertex  $i$  and then joining the half-edges into edges by taking a random partition of the set of all half-edges into pairs. This is known as the configuration model, and was introduced by Bollobás [4]; see also [5, Section II.4]. (See [2] and Wormald [17, 18] for related constructions.) Note that  $G^*(n, (d_i)_1^n)$  is defined for all  $n \geq 1$  and all sequences  $(d_i)_1^n$  such that  $\sum_i d_i$  is even (we tacitly assume this throughout the paper), and that we obtain  $G(n, (d_i)_1^n)$  if we condition  $G^*(n, (d_i)_1^n)$  on being a simple graph.

It is then important to estimate the probability that  $G^*(n, (d_i)_1^n)$  is simple, and in particular to decide whether

$$\liminf_{n \rightarrow \infty} \mathbb{P}\{G^*(n, (d_i)_1^n) \text{ is simple}\} > 0 \quad (1.1)$$

for given sequences  $(d_i)_1^n =: (d_i^{(n)})_1^n$ . (We assume throughout that we consider a sequence of instances, and consider asymptotics as  $n \rightarrow \infty$ . Thus, our degree sequence  $(d_i)_1^n$  depends on  $n$ , and so do other quantities introduced below; for simplicity, we omit this from the notation.) Note that (1.1) implies that any statement holding for  $G^*(n, (d_i)_1^n)$  with probability tending to 1 as  $n \rightarrow \infty$  does so for  $G(n, (d_i)_1^n)$  too. (However, note also that Bollobás and Riordan [6] recently showed that the method may be applied even when (1.1) does not hold; in the problem they studied, the probability that  $G^*(n, (d_i)_1^n)$  is simple may be almost exponentially small, but they showed that the error probabilities for the properties they studied are even smaller.)

Various sufficient conditions for (1.1) have been given by several authors; see Bender and Canfield [2], Bollobás [4, 5], McKay [14], and McKay and Wormald [15]. The final result was proved in [10], where, in particular, the following was shown (but any reader making a detailed

comparison with that paper should note that our notation here differs slightly). Throughout this paper, let

$$N := \sum_i d_i$$

denote the total number of half-edges; thus,  $N$  is even and the number of edges in  $G(n, (d_i)_1^n)$  or  $G^*(n, (d_i)_1^n)$  is  $N/2$ .

**Theorem 1.1.** ([10].) *Assume that  $N \rightarrow \infty$ . Then*

$$\liminf_{n \rightarrow \infty} \mathbb{P}\{G^*(n, (d_i)_1^n) \text{ is simple}\} > 0 \iff \sum_i d_i^2 = O(N).$$

**Remark 1.1.** For simplicity, the graphs and the degree sequences  $(d_i)_1^n = (d_i^{(n)})_1^n$  in Theorem 1.1 are indexed by  $n$ , and, thus,  $N = N(n)$  depends on  $n$  too. With only notational changes, we could instead use an independent index  $\nu$  as in [10], assuming that  $n = n_\nu \rightarrow \infty$ .

Note also that if we assume that  $n = O(N)$ —and this can always be achieved by ignoring all isolated vertices—then the condition  $\sum_i d_i^2 = O(N)$  is equivalent to  $\sum_i d_i^2 = O(n)$  (see [10, Remark 1]).

Let  $X_i$  be the number of loops at vertex  $i$  in  $G^*(n, (d_i)_1^n)$ , and let  $X_{ij}$  be the number of edges between  $i$  and  $j$ . Moreover, let  $Y_{ij} = \binom{X_{ij}}{2}$  be the number of pairs of parallel edges between  $i$  and  $j$ . Define

$$Z := \sum_{i=1}^n X_i + \sum_{i < j} Y_{ij}; \tag{1.2}$$

thus,  $G^*(n, (d_i)_1^n)$  is simple, which is equivalent to  $Z = 0$ .

As shown in [10], in the case that  $\max_i d_i = o(N^{1/2})$ , it is not difficult to prove Theorem 1.1 by Bollobás’s [4, 5] approach of proving a Poisson approximation of  $Z$  by the method of moments. In general, however, we can have  $\max_i d_i = O(N^{1/2})$  even when  $\sum_i d_i^2 = O(N)$ , and in this case,  $Z$  may have a non-Poisson asymptotic distribution. The proof in [10] therefore used a more complicated method with switchings.

The purpose of this paper is to give a new proof of Theorem 1.1, and of the more precise Theorem 1.2 below, using Poisson approximations of  $X_i$  and  $X_{ij}$  to find the asymptotic distribution of  $Z$ . The new proof uses the method of moments. (In [10], we were pessimistic about the possibility of this; our pessimism was thus unfounded.) The new proof presented here is conceptually simpler than the proof in [10], but it is not much shorter. The main reason for the new proof is that it enables us to transfer not only results on convergence in probability but also some results on convergence in distribution from the random multigraph  $G(n, (d_i)_1^n)$  to the simple graph  $G^*(n, (d_i)_1^n)$  by conditioning on the existence of specific loops or pairs of parallel edges; see Section 5 and [12] for an application (which was the motivation for the present paper), and [11] for an earlier example of this method in a case where  $\sum_i d_i^2 = o(N)$  and the results of [10] suffice.

Define (with some hindsight)

$$\lambda_i := \binom{d_i}{2} \frac{1}{N} = \frac{d_i(d_i - 1)}{2N} \tag{1.3}$$

and, for  $i \neq j$ ,

$$\lambda_{ij} := \frac{\sqrt{d_i(d_i - 1)d_j(d_j - 1)}}{N} = 2\sqrt{\lambda_i \lambda_j}, \tag{1.4}$$

and let  $\widehat{X}_i$  and  $\widehat{X}_{ij}$  be independent Poisson random variables with

$$\widehat{X}_i \sim \text{Poi}(\lambda_i), \quad \widehat{X}_{ij} \sim \text{Poi}(\lambda_{ij}). \tag{1.5}$$

By analogy with (1.2), we further define  $\widehat{Y}_{ij} := \binom{\widehat{X}_{ij}}{2}$  and

$$\widehat{Z} := \sum_{i=1}^n \widehat{X}_i + \sum_{i<j} \widehat{Y}_{ij} = \sum_{i=1}^n \widehat{X}_i + \sum_{i<j} \binom{\widehat{X}_{ij}}{2}. \tag{1.6}$$

We shall show that the distribution of  $Z$  is well approximated by  $\widehat{Z}$ , see Lemma 4.1, which yields our new proof of the following estimate. Theorem 1.1 is a simple corollary.

**Theorem 1.2.** ([10].) *Assume that  $n \rightarrow \infty$  and  $N \rightarrow \infty$ . Then*

$$\begin{aligned} \mathbb{P}\{G^*(n, (d_i)_i^n) \text{ is simple}\} &= \mathbb{P}\{Z = 0\} \\ &= \mathbb{P}\{\widehat{Z} = 0\} + o(1) \\ &= \exp\left(-\sum_i \lambda_i - \sum_{i<j} (\lambda_{ij} - \log(1 + \lambda_{ij}))\right) + o(1). \end{aligned}$$

As noted earlier, our proof uses the method of moments, and most of our work lies in deriving the following estimate, in Section 3. This is done by combinatorial calculations that are straightforward in principle, but nevertheless rather long.

**Lemma 1.1.** *Suppose that  $\sum_i d_i^2 = O(N)$ . Then, for every fixed integer  $m \geq 1$ ,*

$$\mathbb{E} Z^m = \mathbb{E} \widehat{Z}^m + O(N^{-1/2}). \tag{1.7}$$

Explicitly, the statement means that, for every  $C < \infty$  and  $m \geq 1$ , there is a constant  $C' = C'(C, m)$  such that if  $\sum_i d_i^2 \leq CN$ , then  $|\mathbb{E} Z^m - \mathbb{E} \widehat{Z}^m| \leq C' N^{-1/2}$ .

**Remark 1.2.** The proof of Lemma 1.1 shows that the error term  $O(N^{-1/2})$  in (1.7) can be replaced by  $O(\max\{d_i\}/N)$ , which by (3.1) is always at least as good.

In Section 6 we give some remarks on the corresponding, but somewhat different, result for bipartite graphs due to Blanchet and Stauffer [3].

## 2. Preliminaries

We denote falling factorials by  $(n)_k := n(n - 1) \cdots (n - k + 1)$ .

**Lemma 2.1.** *Let  $X \in \text{Poi}(\lambda)$ , and let  $Y := \binom{X}{2}$ . Then, for every  $m \geq 1$ ,  $\mathbb{E}[(Y)_m] = h_m(\lambda)$  for a polynomial  $h_m(\lambda)$  of degree  $2m$ . Furthermore,  $h_m$  has a double root at 0, so  $h_m(\lambda) = O(|\lambda|^2)$  for  $|\lambda| \leq 1$ , and if  $m \geq 2$  then  $h_m$  has a triple root at 0, so  $h_m(\lambda) = O(|\lambda|^3)$  for  $|\lambda| \leq 1$ .*

*Proof.*  $(Y)_m$  is a polynomial in  $X$  of degree  $2m$ , and it is well known (and easy to see from the moment generating function) that  $\mathbb{E} X^k$  is a polynomial in  $\lambda$  of degree  $k$  for every  $k \geq 0$ .

Suppose that  $m \geq 2$ . If  $X \leq 2$  then  $Y \leq 1$  and, thus,  $(Y)_m = 0$ . Hence,

$$h_m(\lambda) = \sum_{j=3}^{\infty} \binom{j}{2}_m \frac{\lambda^j}{j!} e^{-\lambda} = O(\lambda^3) \quad \text{as } \lambda \rightarrow 0,$$

and, thus,  $h_m$  has a triple root at 0. The same argument shows that  $h_1$  has a double root at 0; this is also seen from the explicit formula

$$h_1(\lambda) = \mathbb{E} Y = \mathbb{E} \left[ \frac{1}{2} X(X - 1) \right] = \frac{1}{2} \lambda^2. \tag{2.1}$$

**Lemma 2.2.** *Let  $\widehat{Z}$  be given by (1.6), and assume that  $\lambda_{ij} = O(1)$ .*

(i) *For every fixed  $t \geq 0$ ,*

$$\mathbb{E}[\exp(t\sqrt{\widehat{Z}})] = \exp\left(O\left(\sum_i \lambda_i + \sum_{i < j} \lambda_{ij}^2\right)\right). \tag{2.2}$$

(ii) *For every  $C < \infty$ , if  $\sum_i \lambda_i + \sum_{i < j} \lambda_{ij}^2 \leq C$  then*

$$(\mathbb{E} \widehat{Z}^m)^{1/m} = O(m^2), \tag{2.3}$$

*uniformly in all such  $\widehat{Z}$  and  $m \geq 1$ .*

*Proof.* (i) By (1.6),

$$\sqrt{\widehat{Z}} \leq \sum_i \sqrt{\widehat{X}_i} + \sum_{i < j} \sqrt{\widehat{Y}_{ij}} \leq \sum_i \widehat{X}_i + \sum_{i < j} \widehat{X}_{ij} \mathbf{1}\{\widehat{X}_{ij} \geq 2\}, \tag{2.4}$$

where the terms on the right-hand side are independent. Furthermore,

$$\mathbb{E} e^{t\widehat{X}_i} = \exp((e^t - 1)\lambda_i) = \exp(O(\lambda_i)) \tag{2.5}$$

and, since  $t$  is fixed and  $\lambda_{ij} = O(1)$ ,

$$\begin{aligned} \mathbb{E} \exp(t\widehat{X}_{ij} \mathbf{1}\{\widehat{X}_{ij} \geq 2\}) &= \mathbb{E} e^{t\widehat{X}_{ij}} - \mathbb{P}\{\widehat{X}_{ij} = 1\}(e^t - 1) \\ &= e^{(e^t - 1)\lambda_{ij}} - (e^t - 1)\lambda_{ij}e^{-\lambda_{ij}} \\ &= 1 + (e^t - 1)\lambda_{ij}(1 - e^{-\lambda_{ij}}) + O(\lambda_{ij}^2) \\ &= 1 + O(\lambda_{ij}^2) \\ &\leq \exp(O(\lambda_{ij}^2)). \end{aligned} \tag{2.6}$$

Consequently, (2.2) follows from (2.4)–(2.6).

(ii) Taking  $t = 1$ , (i) yields  $\exp(\sqrt{\widehat{Z}}) \leq C_1$  for some  $C_1$ . Since  $\exp(\sqrt{\widehat{Z}}) \geq \widehat{Z}^m / (2m)!$ , this implies that

$$\mathbb{E} \widehat{Z}^m \leq (2m)! \mathbb{E} \exp(\sqrt{\widehat{Z}}) \leq C_1 (2m)^{2m},$$

and, thus,  $\mathbb{E}(\widehat{Z}^m)^{1/m} \leq 4C_1 m^2$  for all  $m \geq 1$ .

### 3. Proof of Lemma 1.1

Our proof of Lemma 1.1 is rather long in spite of its being based on simple calculations, so we formulate a couple of intermediate steps as separate lemmas.

Note first that the assumption that  $\sum_i d_i^2 = O(N)$  implies that

$$\max_i d_i = O(N^{1/2}) \tag{3.1}$$

and, thus (see (1.3) and (1.4)),

$$\lambda_i = O(1) \quad \text{and} \quad \lambda_{ij} = O(1), \tag{3.2}$$

uniformly in all  $i$  and  $j$ . Furthermore, for any fixed  $m \geq 1$ ,

$$\sum_i \lambda_i^m = O\left(\sum_i \lambda_i\right) = O\left(\frac{\sum_i d_i^2}{N}\right) = O(1).$$

Similarly, for any fixed  $m \geq 2$ ,

$$\sum_{i < j} \lambda_{ij}^m = O\left(\sum_{i < j} \lambda_{ij}^2\right) = O\left(\frac{\sum_{ij} d_i^2 d_j^2}{N^2}\right) = O(1).$$

In particular,

$$\sum_i \lambda_i + \sum_{i < j} \lambda_{ij}^2 = O(1). \tag{3.3}$$

However, note that there is no general bound on  $\sum_{i < j} \lambda_{ij}$ , as is shown by the case of regular graphs with all  $d_i = d \geq 2$  and  $\lambda_{ij} = d(d - 1)/N = (d - 1)/n$  for all  $\binom{n}{2}$  of the  $\lambda_{ij}$ , so their sum is  $(d - 1)(n - 1)/2$ . This complicates the proof because it forces us to obtain error estimates involving  $\lambda_{ij}^2$ .

Let  $\mathcal{H}_i$  be the set of half-edges at vertex  $i$ ; thus,  $|\mathcal{H}_i| = d_i$ . Furthermore, let  $\mathcal{H} := \bigcup_i \mathcal{H}_i$  be the set of all half-edges. For convenience, we order  $\mathcal{H}$  (by any linear order).

For  $\alpha, \beta \in \mathcal{H}$ , let  $I_{\alpha\beta}$  be the indicator that the half-edges  $\alpha$  and  $\beta$  are joined to an edge in our random pairing. (Thus,  $I_{\alpha\beta} = I_{\beta\alpha}$ .) Note that

$$X_i = \sum_{\{\alpha, \beta \in \mathcal{H}_i : \alpha < \beta\}} I_{\alpha\beta}, \tag{3.4}$$

$$X_{ij} = \sum_{\alpha \in \mathcal{H}_i, \beta \in \mathcal{H}_j} I_{\alpha\beta}. \tag{3.5}$$

We have  $\mathbb{E} I_{\alpha\beta} = 1/(N - 1)$  for any distinct  $\alpha, \beta \in \mathcal{H}$ . More generally,

$$\mathbb{E}(I_{\alpha_1\beta_1} \cdots I_{\alpha_\ell\beta_\ell}) = \frac{1}{(N - 1)(N - 3) \cdots (N - 2\ell + 1)} = N^{-\ell}(1 + O(N^{-1})) \tag{3.6}$$

for any fixed  $\ell$  and any distinct half-edges  $\alpha_1, \beta_1, \dots, \alpha_\ell, \beta_\ell$ . Furthermore, the expectation in (3.6) vanishes if two pairs  $\{\alpha_i, \beta_i\}$  and  $\{\alpha_j, \beta_j\}$  have exactly one common half-edge.

We consider first, as a warm-up,  $\mathbb{E} X_i^\ell$  for a single vertex  $i$ .

**Lemma 3.1.** *Suppose that  $\sum_i d_i^2 = O(N)$ . Then, for every fixed  $\ell \geq 1$  and all  $i$ ,*

$$\mathbb{E} X_i^\ell = \mathbb{E} \widehat{X}_i^\ell + O(N^{-1/2}\lambda_i). \tag{3.7}$$

*Proof.* We assume that  $d_i \geq 2$  because the case  $d_i \leq 1$  is trivial with  $\lambda_i = 0$  and  $X_i = \widehat{X}_i = 0$ .

Since there are  $\binom{d_i}{2}$  possible loops at  $i$ , (3.6) yields

$$\mathbb{E} X_i = \binom{d_i}{2} \frac{1}{N - 1} = \lambda_i(1 + O(N^{-1})). \tag{3.8}$$

Similarly, for any fixed  $\ell \geq 2$ , there are  $2^{-\ell} \binom{d_i}{2} 2^\ell$  ways to select a sequence of  $\ell$  disjoint (unordered) pairs of half-edges at  $i$ , and, thus, by (3.6), using (1.3), (3.1), (3.2), and (1.5),

$$\begin{aligned} \mathbb{E}[(X_i)_\ell] &= \frac{\binom{d_i}{2} 2^\ell}{2^\ell N^\ell} (1 + O(N^{-1})) \\ &= \frac{(d_i(d_i - 1))^\ell + O(d_i^{2\ell-1})}{2^\ell N^\ell} (1 + O(N^{-1})) \\ &= \lambda_i^\ell (1 + O(N^{-1})) + O\left(\frac{d_i}{N} \lambda_i^{\ell-1}\right) \\ &= \lambda_i^\ell + O(N^{-1}\lambda_i^\ell) + O(N^{-1/2}\lambda_i^{\ell-1}) \end{aligned}$$

$$\begin{aligned}
 &= \lambda_i^\ell + O(N^{-1/2}\lambda_i) \\
 &= \mathbb{E}[(\widehat{X}_i)_\ell] + O(N^{-1/2}\lambda_i).
 \end{aligned}
 \tag{3.9}$$

Conclusion (3.7) now follows from (3.8)–(3.9) and the standard relations between moments and factorial moments, together with (3.2).

We next consider moments of  $Y_{ij}$ , where  $i \neq j$ .

**Lemma 3.2.** *Suppose that  $\sum_i d_i^2 = O(N)$ . Then, for every fixed  $\ell \geq 1$  and all  $i \neq j$ ,*

$$\mathbb{E} Y_{ij}^\ell = \mathbb{E} \widehat{Y}_{ij}^\ell + O(N^{-1/2}\lambda_{ij}^2).
 \tag{3.10}$$

*Proof.* We may assume that  $d_i, d_j \geq 2$  since otherwise  $\lambda_{ij} = 0$  and  $Y_{ij} = \widehat{Y}_{ij} = 0$ .

An unordered pair of two disjoint pairs from  $\mathcal{H}_i \times \mathcal{H}_j$  can be chosen in  $\frac{1}{2}d_i(d_i - 1)d_j(d_j - 1)$  ways, and, thus, by (3.6), (1.4), and (2.1),

$$\mathbb{E} Y_{ij} = \frac{d_i d_j (d_i - 1)(d_j - 1)}{2(N - 1)(N - 3)} = \frac{\lambda_{ij}^2}{2}(1 + O(N^{-1})) = \mathbb{E} \widehat{Y}_{ij}(1 + O(N^{-1})).
 \tag{3.11}$$

Let  $\ell \geq 2$ . Then  $(Y_{ij})_\ell$  is a sum

$$\sum_{\{\alpha_k, \alpha'_k \in \mathcal{H}_i, \beta_k, \beta'_k \in \mathcal{H}_j : \alpha_k < \alpha'_k\}} \prod_{k=1}^\ell (I_{\alpha_k \beta_k} I_{\alpha'_k \beta'_k}),
 \tag{3.12}$$

where we sum only over terms such that the  $\ell$  pairs of pairs  $\{\{\alpha_k, \beta_k\}, \{\alpha'_k, \beta'_k\}\}$  are distinct.

We approximate  $\mathbb{E}[(Y_{ij})_\ell]$  in several steps. First, let  $\tilde{I}_{\alpha\beta}$  for  $\alpha \in \mathcal{H}_i$  and  $\beta \in \mathcal{H}_j$  be independent indicator variables with  $\mathbb{P}\{\tilde{I}_{\alpha\beta} = 1\} = 1/N$ . (In other words, the  $\tilde{I}_{\alpha\beta}$  are independent and identically distributed (i.i.d.) Bernoulli random variables  $\text{Ber}(1/N)$ .) By analogy with (3.5), let

$$\tilde{X}_{ij} := \sum_{\alpha \in \mathcal{H}_i, \beta \in \mathcal{H}_j} \tilde{I}_{\alpha\beta},
 \tag{3.13}$$

and let  $\tilde{Y}_{ij} := \binom{\tilde{X}_{ij}}{2}$ . Then  $(\tilde{Y}_{ij})_\ell$  is a sum similar to (3.12), with  $I_{\alpha\beta}$  replaced by  $\tilde{I}_{\alpha\beta}$ . Note that (3.12) is a sum of terms that are products of  $2\ell$  indicators; however, there may be repetitions among the indicators, so each term is a product of  $r$  distinct indicators where  $r \leq 2\ell$ . Since we assume that  $\ell \geq 2$ , and the pairs  $\{\alpha_k, \beta_k\}$  and  $\{\alpha'_k, \beta'_k\}$  are distinct,  $r \geq 3$  for each term.

Taking expectations and using (3.6), we see that the terms in (3.12), where all pairs  $\{\alpha_k, \beta_k\}$  that occur are distinct, yield the same contributions to  $\mathbb{E}[(Y_{ij})_\ell]$  and  $\mathbb{E}[(\tilde{Y}_{ij})_\ell]$ , apart from a factor  $(1 + O(N^{-1}))$ . However, there are also terms containing factors  $I_{\alpha\beta}$  and  $I_{\alpha'\beta'}$ , where  $\alpha = \alpha'$  or  $\beta = \beta'$  (but not both). Such terms vanish identically for  $(Y_{ij})_\ell$ , but the corresponding terms for  $(\tilde{Y}_{ij})_\ell$  do not. The number of such terms for any given  $r \leq 2\ell$  is  $O(d_i^{r-1}d_j^r + d_i^r d_j^{r-1})$  and, thus, using (3.6) and (3.1), their contribution to  $\mathbb{E}[(\tilde{Y}_{ij})_\ell]$  is

$$O\left(\frac{d_i^{r-1}d_j^r + d_i^r d_j^{r-1}}{N^r}\right) = O\left(\frac{d_j + d_i}{N} \lambda_{ij}^{r-1}\right) = O(N^{-1/2}\lambda_{ij}^{r-1}).$$

Summing over  $3 \leq r \leq 2\ell$  and using (3.2) yields a total contribution  $O(N^{-1/2}\lambda_{ij}^2)$ . Consequently,

$$\mathbb{E}[(Y_{ij})_\ell] = \mathbb{E}[(\tilde{Y}_{ij})_\ell](1 + O(N^{-1})) + O(N^{-1/2}\lambda_{ij}^2).
 \tag{3.14}$$

Next, replace the i.i.d. indicators  $\tilde{I}_{\alpha\beta}$  by i.i.d. Poisson variables  $J_{\alpha\beta} \sim \text{Poi}(1/N)$  with the same mean, and by analogy with (3.5) and (3.13), let

$$\check{X}_{ij} := \sum_{\alpha \in \mathcal{H}_i, \beta \in \mathcal{H}_j} J_{\alpha\beta} \sim \text{Poi}\left(\frac{d_i d_j}{N}\right),$$

and let  $\check{Y}_{ij} = \binom{\check{X}_{ij}}{2}$ . Then  $(\check{Y}_{ij})_\ell$  can be expanded as a sum similar to (3.12), with  $I_{\alpha\beta}$  replaced by  $J_{\alpha\beta}$ . Take the expectation and note that the only difference from  $\mathbb{E}[(\tilde{Y}_{ij})_\ell]$  is for terms where some  $J_{\alpha\beta}$  is repeated. We have, for any fixed  $k \geq 1$ ,

$$\mathbb{E} J_{\alpha\beta}^k = \frac{1}{N} + O(N^{-2}) = \frac{1}{N}(1 + O(N^{-1})),$$

while

$$\mathbb{E} \tilde{I}_{\alpha\beta}^k = \mathbb{E} \tilde{I}_{\alpha\beta} = \frac{1}{N}.$$

Hence, for each term, the difference, if any, is a multiplicative factor  $1 + O(N^{-1})$ , and, thus,

$$\mathbb{E}[(\check{Y}_{ij})_\ell] = \mathbb{E}[(\tilde{Y}_{ij})_\ell](1 + O(N^{-1})). \tag{3.15}$$

Note that here we use  $\check{Y}_{ij} = \binom{\check{X}_{ij}}{2}$ , where  $\check{X}_{ij} \sim \text{Poi}(d_i d_j / N)$  has a mean  $\check{\lambda}_{ij} = d_i d_j / N$  that differs from  $\mathbb{E}[\tilde{X}_{ij}] = \lambda_{ij}$  given by (1.4). We have

$$\check{\lambda}_{ij} \geq \lambda_{ij} \geq \frac{(d_i - 1)(d_j - 1)}{N} > \check{\lambda}_{ij} - \frac{d_i + d_j}{N}. \tag{3.16}$$

We use Lemma 2.1 and note that the lemma implies that  $h'_\ell(\lambda) = O(\lambda^2)$  for each  $\ell \geq 2$  and  $\lambda = O(1)$ . Hence, by (3.16) and (3.1)–(3.2),

$$\begin{aligned} \mathbb{E}[(\check{Y}_{ij})_\ell] - \mathbb{E}[(\widehat{Y}_{ij})_\ell] &= h_\ell(\check{\lambda}_{ij}) - h_\ell(\lambda_{ij}) \\ &= O(\check{\lambda}_{ij}^2(\check{\lambda}_{ij} - \lambda_{ij})) \\ &= O\left(\frac{d_i + d_j}{N} \check{\lambda}_{ij}^2\right) \\ &= O(N^{-1/2} \lambda_{ij}^2). \end{aligned} \tag{3.17}$$

Finally, (3.14), (3.15), and (3.17) yield, for each  $\ell \geq 2$ ,

$$\mathbb{E}[(Y_{ij})_\ell] = \mathbb{E}[(\widehat{Y}_{ij})_\ell](1 + O(N^{-1})) + O(N^{-1/2} \lambda_{ij}^2). \tag{3.18}$$

By (3.11), this holds for  $\ell = 1$  too. By (3.2) and Lemma 2.1, for each  $\ell \geq 1$ ,  $\mathbb{E}[(\widehat{Y}_{ij})_\ell] = O(\lambda_{ij}^2)$ , and, thus, (3.18) can be written as

$$\mathbb{E}[(Y_{ij})_\ell] = \mathbb{E}[(\widehat{Y}_{ij})_\ell] + O(N^{-1/2} \lambda_{ij}^2)$$

for each fixed  $\ell \geq 1$ . The conclusion now follows, as in Lemma 3.1, by the relations between moments and factorial moments, again using the bound (3.2).

In particular, note that Lemmas 3.1 and 3.2 together with Lemma 2.1 and (3.2) imply the bounds, for every fixed  $\ell \geq 1$ ,

$$\mathbb{E} X_i^\ell + \mathbb{E} \widehat{X}_i^\ell = O(\lambda_i + \lambda_i^\ell + N^{-1/2} \lambda_i) = O(\lambda_i), \tag{3.19}$$

$$\mathbb{E} Y_{ij}^\ell + \mathbb{E} \widehat{Y}_{ij}^\ell = O(\lambda_{ij}^2 + \lambda_{ij}^{2\ell} + N^{-1/2} \lambda_{ij}^2) = O(\lambda_{ij}^2). \tag{3.20}$$



*Proof of Lemma 1.1.* We uncouple the terms in (1.2) by letting  $\{I_{\alpha\beta}^{(i,j)}\}_{\alpha,\beta}$  be independent copies of  $(I_{\alpha\beta})_{\alpha,\beta}$  for  $1 \leq i, j \leq n$ , and defining, by analogy with (3.4)–(3.5) and (1.2),

$$\bar{X}_i := \sum_{\{\alpha,\beta \in \mathcal{H}_i: \alpha < \beta\}} I_{\alpha\beta}^{(i,i)}, \tag{3.21}$$

$$\bar{X}_{ij} := \sum_{\alpha \in \mathcal{H}_i, \beta \in \mathcal{H}_j} I_{\alpha\beta}^{(i,j)}, \tag{3.22}$$

$$\bar{Y}_{ij} := \binom{\bar{X}_{ij}}{2}, \tag{3.23}$$

$$\bar{Z} := \sum_{i=1}^n \bar{X}_i + \sum_{i < j} \bar{Y}_{ij}. \tag{3.24}$$

Note that the summands in (3.21) are not independent; they have the same structure as  $(I_{\alpha\beta})_{\alpha,\beta \in \mathcal{H}_i}$  and, thus,  $\bar{X}_i \stackrel{D}{=} X_i$ , and similarly  $X_{ij} \stackrel{D}{=} X_{ij}$ . However, different sums  $\bar{X}_i$  and  $\bar{X}_{ij}$  are independent (unlike  $X_i$  and  $X_{ij}$ ).

We begin by comparing  $\mathbb{E} \bar{Z}^m$  and  $\mathbb{E} \widehat{Z}^m$ . Since the terms in (3.24) are independent, the moment  $\mathbb{E} \bar{Z}^m$  can be written as a certain polynomial  $g_m(\mathbb{E} X_i^\ell, \mathbb{E} Y_{ij}^\ell : i, j \in [n], \ell \leq m)$  in the moments  $\mathbb{E}(\bar{X}_i)^\ell = \mathbb{E} X_i^\ell$  and  $\mathbb{E}(\bar{X}_{ij})^\ell = \mathbb{E} X_{ij}^\ell$  for  $1 \leq \ell \leq m$  and  $i, j \in [n]$ .

By (1.6),  $\mathbb{E} \widehat{Z}^m$  can be expressed in the same way as  $g_m(\mathbb{E} \widehat{X}_i^\ell, \mathbb{E} \widehat{Y}_{ij}^\ell : i, j \in [n], \ell \leq m)$  for the same polynomial  $g_m$ . It follows that

$$\mathbb{E} \bar{Z}^m - \mathbb{E} \widehat{Z}^m = \sum_{\ell=1}^m \sum_i (\mathbb{E} X_i^\ell - \mathbb{E} \widehat{X}_i^\ell) R_{\ell i} + \sum_{\ell=1}^m \sum_{i < j} (\mathbb{E} Y_{ij}^\ell - \mathbb{E} \widehat{Y}_{ij}^\ell) R_{\ell ij} \tag{3.25}$$

for some polynomials  $R_{\ell i}$  and  $R_{\ell ij}$  in the moments  $\mathbb{E} X_i^k, \mathbb{E} \widehat{X}_i^k, \mathbb{E} Y_{ij}^k$ , and  $\mathbb{E} \widehat{Y}_{ij}^k$  for  $k \leq m$ ; it is easily seen from (3.19)–(3.20) and (3.3) that

$$R_{\ell i}, R_{\ell ij} = O\left(\sum_{v=1}^m \left(\sum_i \lambda_i + \sum_{i < j} \lambda_{ij}^2\right)^v\right) = O(1),$$

uniformly in  $i, j \in [n]$  and  $\ell \leq m$ . Hence, (3.25) with (3.7), (3.10), and (3.3) yields

$$\mathbb{E} \bar{Z}^m - \mathbb{E} \widehat{Z}^m = O\left(\sum_i N^{-1/2} \lambda_i + \sum_{i < j} N^{-1/2} \lambda_{ij}^2\right) = O(N^{-1/2}). \tag{3.26}$$

It remains to compare  $\mathbb{E} Z^m$  and  $\mathbb{E} \bar{Z}^m$ . By (1.2) and (3.4)–(3.5),  $Z^m$  can be expanded as a sum of certain products

$$I_{\alpha_1 \beta_1} \cdots I_{\alpha_\ell \beta_\ell}, \tag{3.27}$$

where  $1 \leq \ell \leq 2m$ , and we may assume that the pairs  $\{\alpha_1, \beta_1\}, \dots, \{\alpha_\ell, \beta_\ell\}$  are distinct. (Some products (3.27) may be repeated in  $Z^m$ , but only  $O(1)$  times.) Moreover, by (3.21)–(3.24),  $\bar{Z}^m$  is the sum of the corresponding products

$$\bar{I}_{\alpha_1 \beta_1} \cdots \bar{I}_{\alpha_\ell \beta_\ell}, \tag{3.28}$$

where  $\bar{I}_{\alpha\beta} := I_{\alpha\beta}^{(i,j)}$  when  $\alpha \in \mathcal{H}_i$  and  $\beta \in \mathcal{H}_j$ .

We say that a product (3.27) or (3.28) is *bad* if it contains two factors  $I_{\alpha_\nu \beta_\nu}$  and  $I_{\alpha_\mu \beta_\mu}$  such that the pairs  $\{\alpha_\nu, \beta_\nu\}$  and  $\{\alpha_\mu, \beta_\mu\}$  contain a common index, say  $\alpha_\nu = \alpha_\mu$ , and, furthermore, the two remaining indices,  $\beta_\nu$  and  $\beta_\mu$ , say, are half-edges belonging to different vertices,

i.e.  $\beta_\nu \in \mathcal{H}_i$  and  $\beta_\mu \in \mathcal{H}_j$  with  $i \neq j$ . Otherwise, we say that the product is *good*. (Note that a good product may contain factors  $I_{\alpha_\nu\beta_\nu}$  and  $I_{\alpha_\mu\beta_\mu}$  with  $\alpha_\nu = \alpha_\mu$  as long as  $\beta_\nu$  and  $\beta_\mu$  belong to the same vertex.) It follows from (3.6) that, for each good product, the corresponding contributions to  $\mathbb{E} Z^m$  and  $\mathbb{E} \bar{Z}^m$  differ only by a factor  $(1 + O(N^{-1}))$ . For a bad product, however, the contribution to  $\mathbb{E} Z^m$  is 0. We thus have to estimate the contribution to  $\mathbb{E} \bar{Z}^m$  of the bad products.

Define the *support* of a product (3.28) as the multigraph with vertex set  $[n]$  and edge set  $\{\alpha_\nu\beta_\nu : 1 \leq \nu \leq \ell\}$ , i.e. the multigraph obtained by forming edges from the pairs of half-edges appearing as indices in the product. If  $F$  is the support of (3.28) then  $F$  thus has  $\ell$  edges (possibly including loops). Furthermore, it follows from (3.21)–(3.24) that every edge in  $F$  that is not a loop has at least one edge parallel to it. Hence, a vertex  $i$  in  $F$  with nonzero degree has degree at least 2. In other words, if we denote the vertex degrees in  $F$  by  $\delta_1, \dots, \delta_n$  then  $\delta_i = 0$  or  $\delta_i \geq 2$ . Moreover, if (3.28) is bad with, say,  $\alpha_\nu = \alpha_\mu \in \mathcal{H}_i$ , then there are edges in  $F$  from  $i$  to at least two vertices  $j$  and  $k$  (one of which may equal  $i$ ), and, thus, the degree  $\delta_i \geq 4$ .

Let  $F$  be a multigraph with vertex set  $[n]$  and  $\ell$  edges, and again denote its vertex degrees by  $\delta_1, \dots, \delta_n$ , so  $\sum_i \delta_i = 2\ell$ . Let  $S_F$  be the contribution to  $\mathbb{E} \bar{Z}^m$  from bad products (3.27) with support  $F$ . A bad product has some half-edge repeated, and if this belongs to  $\mathcal{H}_i$ , there are  $O(d_i^{\delta_i-1} \prod_{j \neq i} d_j^{\delta_j})$  choices for the product. Also, as just shown, this can only occur for  $i$  with  $\delta_i \geq 4$ . Since each product yields a contribution  $O(N^{-\ell})$  by (3.6), we have, using  $2\ell = \sum_i \delta_i$  and (3.1) together with the fact that  $\delta_j \neq 1$ ,

$$S_F = O\left(N^{-\ell} \sum_{\{i: \delta_i \geq 4\}} d_i^{\delta_i-1} \prod_{j \neq i} d_j^{\delta_j}\right) \tag{3.29}$$

$$= O\left(N^{-1/2} \sum_{\{i: \delta_i \geq 4\}} \left(\frac{d_i}{N^{1/2}}\right)^{\delta_i-1} \prod_{j \neq i} \left(\frac{d_j}{N^{1/2}}\right)^{\delta_j}\right) \tag{3.30}$$

$$= O\left(N^{-1/2} \prod_{\{j: \delta_j > 0\}} \left(\frac{d_j}{N^{1/2}}\right)^2\right). \tag{3.31}$$

Summing over all possible  $F$ , and recalling that  $\ell \leq 2m$ , it follows that the total contribution to  $\mathbb{E} \bar{Z}^m$  from bad products is

$$\sum_F S_F = O\left(N^{-1/2} \sum_F \prod_{\{j: \delta_j > 0\}} \frac{d_j^2}{N}\right). \tag{3.32}$$

For each support  $F$ , the set  $\{j: \delta_j > 0\} = \{j: \delta_j \geq 2\}$  has size at most  $\ell \leq 2m$ , and, for each choice of this set, there are  $O(1)$  possible  $F$ . Hence,

$$\sum_F \prod_{\{j: \delta_j > 0\}} \frac{d_j^2}{N} = O\left(\sum_{k=1}^{2m} \sum_{j_1 < \dots < j_k} \prod_{i=1}^k \frac{d_{j_i}^2}{N}\right) = O\left(\sum_{k=1}^{2m} \left(\sum_{j=1}^n \frac{d_j^2}{N}\right)^k\right) = O(1),$$

and (3.32) yields

$$\sum_F S_F = O(N^{-1/2}).$$

Summarizing, the argument above yields

$$\mathbb{E} Z^m = \mathbb{E} \bar{Z}^m (1 + O(N^{-1})) + O(N^{-1/2}) = \mathbb{E} \bar{Z}^m + O(N^{-1/2}), \tag{3.33}$$

since  $\mathbb{E} \bar{Z}^m = O(1)$ , e.g. by (3.26) and Lemma 2.2 (or, using an argument similar to the above, by summing over supports).

The lemma follows from (3.26) and (3.33).

#### 4. Proof of Theorems 1.1 and 1.2

Assume first that  $\sum_i d_i^2 = O(N)$ ; we prove the following more precise statement.

**Lemma 4.1.** *Suppose that  $\sum_i d_i^2 = O(N)$  and  $N \rightarrow \infty$ . Then  $d_{TV}(Z, \widehat{Z}) \rightarrow 0$ .*

*Proof.* Note that the assumption that  $\sum_i d_i^2 = O(N)$  implies that (3.3) holds.

By Lemma 2.2(ii) and (3.3),  $\mathbb{E} \widehat{Z}^m = O(1)$  for each  $m$ . In particular, the sequence  $\widehat{Z}$  is tight, and by considering a suitable subsequence we may assume that  $\widehat{Z} \xrightarrow{D} Z_\infty$  for some random variable  $Z_\infty$ . Furthermore, the estimate  $\mathbb{E} \widehat{Z}^m = O(1)$  for each  $m$  implies that  $\widehat{Z}^m$  is uniformly integrable for each  $m \geq 1$ , and, thus,

$$\mathbb{E} \widehat{Z}^m \rightarrow \mathbb{E} Z_\infty^m; \tag{4.1}$$

see, e.g. [9, Theorems 5.4.2 and 5.5.9]. By Lemma 1.1, we thus also have

$$\mathbb{E} Z^m \rightarrow \mathbb{E} Z_\infty^m \tag{4.2}$$

for each  $m \geq 1$ . Furthermore, by (2.3) and (4.1),

$$(\mathbb{E} Z_\infty^m)^{1/m} = O(m^2). \tag{4.3}$$

We can now apply the method of moments and conclude from (4.2) that  $Z \xrightarrow{D} Z_\infty$ . We justify the use of the method of moments by (4.3), which implies that

$$\sum_m (\mathbb{E} Z_\infty^m)^{-1/2m} = \infty; \tag{4.4}$$

since  $Z_\infty \geq 0$ , this weaker form of the usual Carleman criterion shows that the distribution of  $Z_\infty$  is determined (among all distributions on  $[0, \infty)$ ) by its moments, and, thus (since also  $Z \geq 0$ ), the method of moments applies; see, e.g. [9, Section 4.10]. Hence,  $\widehat{Z} \xrightarrow{D} Z_\infty$  and  $Z \xrightarrow{D} Z_\infty$ , and, thus,

$$d_{TV}(Z, \widehat{Z}) \leq d_{TV}(Z, Z_\infty) + d_{TV}(\widehat{Z}, Z_\infty) \rightarrow 0. \tag{4.5}$$

This shows that the desired result (4.5) holds for some subsequence. The same argument shows that, for every subsequence of  $n \rightarrow \infty$ , (4.5) holds for some subsubsequence; as is well known, this implies that (4.5) holds for the original sequence.

**Remark 4.1.** Note that  $\mathbb{E} e^{t\widehat{Y}_{ij}} = \infty$  for every  $t > 0$  when  $\lambda_{ij} > 0$ . Hence,  $\widehat{Z}$  does not have a finite moment generating function. Similarly, it is possible that  $\mathbb{E} e^{tZ_\infty} = \infty$ ; consider, for example, the case in which  $d_1 = d_2 \sim N^{1/2}$  when  $\lambda_{12} \rightarrow 1$  and  $Z_\infty \geq \binom{\widehat{X}}{2}$  with  $\widehat{X} \sim \text{Poi}(1)$ . In this case, by Minkowski’s inequality,

$$\begin{aligned} (\mathbb{E} Z_\infty^m)^{1/m} &\geq \frac{1}{2} (\mathbb{E} (\widehat{X}^2 - \widehat{X})^m)^{1/m} \\ &\geq \frac{1}{2} (\mathbb{E} \widehat{X}^{2m})^{1/m} - \frac{1}{2} (\mathbb{E} \widehat{X}^m)^{1/m} \\ &\sim \frac{1}{2} \left( \frac{2m}{e \log m} \right)^2 \\ &= \frac{2m^2}{e^2 \log^2 m} \end{aligned}$$

using simple estimates for the moments  $\mathbb{E} \widehat{X}^m$  when  $\widehat{X} \sim \text{Poi}(1)$ , which are the Bell numbers. (Or by more precise asymptotics in, e.g. [7, Proposition VIII.3] and [16, Section 26.7].) Hence, in this case,  $\sum_m (\mathbb{E} Z_\infty^m)^{-1/m} < \infty$ ; in other words,  $Z_\infty$  does not satisfy the usual Carleman criterion  $\sum_m (\mathbb{E} Z_\infty^m)^{-1/m} = \infty$  for its distribution to be determined by its moments. However, since we here deal with nonnegative random variables, we can use the weaker condition (4.4). (This weaker version is well known, and follows from the standard version by considering the square root  $\pm Z_\infty$  with random sign, independent of  $Z_\infty$ . Alternatively, observe that (4.2) implies that  $\mathbb{E}(\pm\sqrt{Z})^k \rightarrow \mathbb{E}(\pm\sqrt{Z_\infty})^k$  for all  $k \geq 0$ , where the moments trivially vanish when  $k$  is odd; since  $\pm\sqrt{Z_\infty}$  has a finite moment generating function by (2.2) and Fatou's lemma, the usual sufficient condition for the method of moments yields  $\pm\sqrt{Z} \xrightarrow{D} \pm\sqrt{Z_\infty}$ , and, thus,  $Z \xrightarrow{D} Z_\infty$ .)

*Proof of Theorems 1.1 and 1.2.* In the case  $\sum_i d_i^2 = O(N)$ , Theorem 1.2 follows from Lemma 4.1, since

$$\begin{aligned} \mathbb{P}\{\widehat{Z} = 0\} &= \mathbb{P}\{\widehat{X}_i = \widehat{Y}_{ij} = 0 \text{ for all } i, j\} \\ &= \prod_i \mathbb{P}\{\widehat{X}_i = 0\} \prod_{i < j} \mathbb{P}\{\widehat{X}_{ij} \leq 1\} \\ &= \prod_i e^{-\lambda_i} \prod_{i < j} (1 + \lambda_{ij})e^{-\lambda_{ij}}. \end{aligned} \tag{4.6}$$

Furthermore, we have  $\lambda_{ij} - \log(1 + \lambda_{ij}) = O(\lambda_{ij}^2)$ , so it follows from this and (3.3) that  $\liminf_{n \rightarrow \infty} \mathbb{P}\{G^*(n, (d_i)_1^n) \text{ is simple}\} > 0$ , verifying Theorem 1.1 in this case.

It remains (by considering subsequences) only to consider the case when  $\sum_i d_i^2/N \rightarrow \infty$ . Since then

$$\sum_i \lambda_i = \frac{\sum_i d_i^2 - \sum_i d_i}{2N} = \frac{\sum_i d_i^2}{2N} - \frac{1}{2} \rightarrow \infty, \tag{4.7}$$

it follows from (4.6) that  $\mathbb{P}\{\widehat{Z} = 0\} \rightarrow 0$ , and it remains to show that  $\mathbb{P}\{Z = 0\} \rightarrow 0$ . We do this by the method used in [10] for this case. Fix  $A > 1$ , and split vertices by replacing some  $d_j$  by  $d_j - 1$  and a new vertex  $n + 1$  with  $d_{n+1} = 1$ , repeating until the new degree sequence,  $(\bar{d}_i)_1^n$  say, satisfies  $\sum_i \bar{d}_i^2 \leq AN$ . (Note that the number  $N$  of half-edges is unchanged.) Then, as  $N \rightarrow \infty$ , see [10] for details,  $\sum_i \bar{d}_i^2 \sim AN$  and, denoting the new random multigraph by  $\bar{G}$  and using Lemma 4.1 together with (4.6) and (4.7) on  $\bar{G}$ ,

$$\begin{aligned} \mathbb{P}\{G(n, (d_i)_1^n) \text{ is simple}\} &\leq \mathbb{P}\{\bar{G} \text{ is simple}\} \\ &\leq \exp\left(-\sum_i \frac{\bar{d}_i(\bar{d}_i - 1)}{2N}\right) + o(1) \\ &= \exp\left(-\frac{\sum_i \bar{d}_i^2}{2N} + \frac{1}{2}\right) + o(1) \\ &\rightarrow \exp\left(-\frac{A - 1}{2}\right). \end{aligned}$$

Since  $A$  is arbitrary, it follows that  $\mathbb{P}\{G(n, (d_i)_1^n) \text{ is simple}\} = \mathbb{P}\{Z = 0\} \rightarrow 0$  in this case, which completes the proof.

**Remark 4.2.** The proof of Lemma 4.1 shows that if  $\sum_i d_i^2 = O(N)$  for  $N \rightarrow \infty$  and also that  $\widehat{Z} \xrightarrow{D} Z_\infty$  for some random variable  $Z_\infty$  (which is a kind of regularity property of the degree sequences  $(d_i)_1^n$ ), then  $Z \xrightarrow{D} Z_\infty$ , with convergence of all moments.

### 5. An application

We sketch here an application of our results (see [12] for details). We believe that similar arguments can be used for other problems too.

We consider a certain random infection process on the (multi)graph, under certain assumptions; let  $\mathcal{L}$  be the event that at most  $\log n$  vertices will be infected. It was shown in [12] that, for the multigraph  $G^*(n, (d_i)_1^n)$ ,  $\mathbb{P}(\mathcal{L}) \rightarrow \kappa$  for some  $\kappa > 0$ ; we want to conclude that, assuming that  $\sum_i d_i^2 = O(N)$ , the same is true for the simple random graph  $G(n, (d_i)_1^n)$ , i.e. that

$$\mathbb{P}(\mathcal{L} \mid \{Z = 0\}) \rightarrow \kappa, \tag{5.1}$$

where, as above,  $Z$  is the number of loops and pairs of parallel edges. By considering a subsequence, we can assume that  $Z \xrightarrow{D} Z_\infty$  for some random variable  $Z_\infty$  (see Remark 4.2). Then, using  $\mathbb{P}\{\mathcal{L}\} \rightarrow \kappa > 0$  and  $\liminf \mathbb{P}\{Z = 0\} > 0$  (Theorem 1.1), (5.1) is equivalent to  $\mathbb{P}\{\mathcal{L} \cap \{Z = 0\}\} \rightarrow \kappa \mathbb{P}\{Z_\infty = 0\}$  and, thus, to

$$\mathbb{P}\{\{Z = 0\} \mid \mathcal{L}\} \rightarrow \mathbb{P}\{Z_\infty = 0\}.$$

Furthermore, the distribution of  $Z_\infty$  is determined by its moments, at least among nonnegative distributions (see the proof of Lemma 4.1). Consequently, it suffices to show that, for every fixed  $m \geq 0$ ,

$$\mathbb{E}[Z^m \mid \mathcal{L}] \rightarrow \mathbb{E} Z_\infty^m. \tag{5.2}$$

Actually, for technical reasons, we show a modification of (5.2): we let  $Z = Z_1 + Z_2$ , where  $Z_2$  is the number of loops and pairs of parallel edges that include an initially infected vertex. It is easily shown that  $\mathbb{E} Z_2 \rightarrow 0$ , and, thus, it suffices to show that

$$\mathbb{E}[Z_1^m \mid \mathcal{L}] \rightarrow \mathbb{E} Z_\infty^m. \tag{5.3}$$

To this end, we write  $Z_1^m = \sum_\gamma I_\gamma$ , where  $I_\gamma$  is the indicator that a certain  $m$ -tuple of loops and pairs of parallel edges exists in the configuration model yielding  $G^*(n, (d_i)_1^n)$ . For each  $\gamma$ , if we condition on  $I_\gamma = 1$ , we have another instance of the configuration model, with the degrees at the vertices involved in  $\gamma$  reduced, plus some extra edges giving  $\gamma$ , and it is easy to see that the result  $\mathbb{P}\{\mathcal{L}\} \rightarrow \kappa$  applies to this modification too, and, thus,

$$\mathbb{P}\{\mathcal{L} \mid \{I_\gamma = 1\}\} = \kappa + o(1),$$

uniformly for all  $\gamma$ . We invert the conditioning again and obtain

$$\mathbb{E}[I_\gamma \mid \mathcal{L}] = \frac{\mathbb{P}\{\mathcal{L} \mid \{I_\gamma = 1\}\} \mathbb{P}\{I_\gamma = 1\}}{\mathbb{P}\{\mathcal{L}\}} = (1 + o(1)) \mathbb{E}(I_\gamma).$$

Consequently,

$$\mathbb{E}[Z_1^m \mid \mathcal{L}] = \sum_\gamma \mathbb{E}[I_\gamma \mid \mathcal{L}] \sim \sum_\gamma \mathbb{E}(I_\gamma) = \mathbb{E} Z_1^m,$$

and since  $\mathbb{E} Z_1^m \rightarrow \mathbb{E} Z_\infty^m$ , this yields (5.3), as desired.

### 6. Bipartite graphs

A similar result for bipartite graphs has been proved by Blanchet and Stauffer [3]; see, e.g. [1, 8, 13] for earlier results that are often stated in an equivalent form about matrices with  $\{0, 1\}$ -valued elements. We suppose as given the degree sequences  $(s_i)_1^n$  and  $(t_j)_1^n$

for the two parts, with  $N := \sum_i s_i = \sum_j t_j$ , and consider a random bipartite simple graph  $G(n, (s_i)_{i=1}^{n'}, (t_j)_{j=1}^{n''})$  with these degree sequences as well as the corresponding random bipartite multigraph  $G^* = G^*(n, (s_i)_{i=1}^{n'}, (t_j)_{j=1}^{n''})$  constructed by the configuration model. (These have  $N$  edges.) We order the two degree sequences in decreasing order as  $s_{(1)} \geq \dots \geq s_{(n')}$  and  $t_{(1)} \geq \dots \geq t_{(n'')}$ , and let  $s := s_{(1)} = \max_i s_i$  and  $t := t_{(1)} = \max_j t_j$ . Label the vertices in the two parts  $v_1, \dots, v_{n'}$  and  $w_1, \dots, w_{n''}$ , in order of decreasing degrees; thus,  $v_i[w_j]$  has degree  $s_{(i)}[t_{(j)}]$ .

**Theorem 6.1.** ([3].) *Assume that  $N \rightarrow \infty$ . Then*

$$\liminf_{n \rightarrow \infty} \mathbb{P}\{G^*(n, (s_i)_{i=1}^{n'}, (t_j)_{j=1}^{n''}) \text{ is simple}\} > 0$$

*if and only if the following two conditions hold:*

$$\sum_i \sum_j s_i(s_i - 1)t_j(t_j - 1) = O(N^2), \tag{6.1}$$

$$\sum_{i=\min\{t,m\}}^{n'} s_{(i)} = \Omega(N) \quad \text{and} \quad \sum_{j=\min\{s,m\}}^{n''} t_{(j)} = \Omega(N). \tag{6.2}$$

(Equation (6.2) is reformulated and simplified from [3]. Recall that  $x = \Omega(N)$  means that  $\liminf x/N > 0$ .)

**Remark 6.1.** Here (6.1) corresponds to the condition  $\sum_i d_i^2 = O(N)$  in Theorem 1.1, while (6.2) is an additional complication. Note that if  $s = o(N)$  then the first part of (6.2) holds, because the sum is greater than or equal to  $N - (m - 1)s$ ; similarly, if  $t = o(N)$  then the second part of (6.2) holds. Hence, (6.2) is satisfied, and (6.1) is sufficient, unless for some subsequence either  $s = \Omega(N)$  or  $t = \Omega(N)$ . Note also that both these cannot occur when (6.1) holds; in fact, if  $s = \Omega(N)$  then (6.1) implies that  $\sum_j t_j(t_j - 1) = O(1)$  and, thus,  $t = O(1)$ . On the other hand, in such cases, (6.1) is not enough, as pointed out by Blanchet and Stauffer [3]. For example, if  $s_1 = N - o(N)$ ,  $t_1 = 2$ , and  $t_j = 1$  for  $j \geq 2$ , then (6.1) holds but (6.2) fails for  $m = 2$ . Indeed, in this example, there is with high probability (i.e. with probability  $1 - o(1)$ ) a double edge  $v_1 w_1$ , and, thus,  $G^*$  is with high probability not simple.

We can also prove Theorem 6.1 by the methods of this paper (the proof by Blanchet and Stauffer [3] is different). There are no loops, and, thus, no  $X_i$ , but we define  $X_{ij}$  and  $Y_{ij}$  as above (with the original labelling), and let  $Z := \sum_{i=1}^{n'} \sum_{j=1}^{n''} Y_{ij}$ . Similarly, we define, for  $i \in [n']$  and  $j \in [n'']$ ,

$$\lambda_{ij} := \frac{\sqrt{s_i(s_i - 1)t_j(t_j - 1)}}{N},$$

let  $\widehat{X}_{ij} \sim \text{Poi}(\lambda_{ij})$  and  $\widehat{Y}_{ij} := \binom{\widehat{X}_{ij}}{2}$  be as above, and let  $\widehat{Z} := \sum_{i=1}^{n'} \sum_{j=1}^{n''} \widehat{Y}_{ij}$ . Note that (6.1) is  $\sum_{i,j} \lambda_{ij}^2 = O(1)$ .

**Theorem 6.2.** *Assume that  $N \rightarrow \infty$  and that  $s, t = o(N)$ . Then*

$$\begin{aligned} \mathbb{P}\{G^*(n, (s_i)_{i=1}^{n'}, (t_j)_{j=1}^{n''}) \text{ is simple}\} &= \mathbb{P}\{Z = 0\} \\ &= \mathbb{P}\{\widehat{Z} = 0\} + o(1) \\ &= \exp\left(-\sum_{i,j} (\lambda_{ij} - \log(1 + \lambda_{ij}))\right) + o(1). \end{aligned}$$

*Proof (sketch).* The proof is similar to that of Theorem 1.2, using analogues of Lemmas 1.1 and 4.1 with only minor differences. Instead of (3.1) use the assumption that  $s, t = o(N)$ , which leads to error terms of the order  $O((s + t)/N)$ ; cf. Remark 1.2. Furthermore, (3.31) has to be modified. Say that the vertex with a repeated half-edge is bad, and suppose that the bad vertex is in the first part. Let the nonzero vertex degrees in  $F$  be  $a_1, a_2, \dots$  in the first part and  $b_1, b_2, \dots$  in the second part, in any order with the bad vertex having degree  $a_1$ . Thus,  $\sum_v a_v = \sum_\mu b_\mu = \ell$ . Using Hölder’s inequality and (6.1), the contribution from all  $F$  with given  $(a_v)$  and  $(b_\mu)$  is

$$\begin{aligned} & O\left(N^{-\ell} \sum_{\{i: s_i \geq 2\}} s_i^{a_1-1} \prod_{v \geq 2} \left(\sum_{\{i: s_i \geq 2\}} s_i^{a_v}\right) \prod_{\mu \geq 1} \left(\sum_{\{j: t_j \geq 2\}} t_j^{b_\mu}\right)\right) \\ &= O\left(N^{-\ell} \left(\sum_{\{i: s_i \geq 2\}} s_i^2\right)^{(a_1-1)/2 + \sum_{v \geq 2} a_v/2} \left(\sum_{\{j: t_j \geq 2\}} t_j^2\right)^{\sum_\mu b_\mu/2}\right) \\ &= O\left(N^{-\ell} \left(\sum_i s_i(s_i - 1)\right)^{(\ell-1)/2} \left(\sum_j t_j(t_j - 1)\right)^{\ell/2}\right) \\ &= O\left(N^{-1} \left(\sum_j t_j(t_j - 1)\right)^{1/2}\right) \\ &= O\left(\frac{t^{1/2}}{N^{1/2}}\right). \end{aligned}$$

Summing over the finitely many  $(a_v)$  and  $(b_\mu)$ , and adding the case with the bad vertex in the second part, we obtain  $O((s + t)^{1/2}/N^{1/2}) = o(1)$ .

*Proof of Theorem 6.1.* The case  $s, t = o(N)$  (when (6.2) is automatic) follows from Theorem 6.2; note that

$$-\sum_{i,j} (\lambda_{ij} - \log(1 + \lambda_{ij})) = O(1) \iff \sum_{i,j} \lambda_{ij}^2 = O(1) \iff (6.1).$$

By symmetry and considering subsequences, it remains to consider only the case  $s = \Omega(N)$ . It is easy to see that (6.1) is necessary in this case too, so we may assume that (6.1) holds. As noted above, this implies that  $t = O(1)$ , and, furthermore, that only  $O(1)$  degrees  $t_j$  are greater than 1. By taking a further subsequence, we may assume that  $t$  is constant. Then the second part of (6.2) always holds, and it suffices to consider the case  $m = t$  in the first part of (6.2), i.e.

$$\sum_{i=t}^{n'} s(i) = \Omega(N). \tag{6.3}$$

If (6.3) does not hold then (at least for a subsequence), with high probability,  $\sum_{i=t}^{n'} s(i) = o(N)$ , and then, with high probability, the  $t$  edges from  $w_1$  go only to  $\{v_i : i < t\}$ , so, by the pigeonhole principle, there is a double edge.

Conversely, if (6.3) holds, it is easy to see that if we first match the half-edges from  $w_1, w_2, \dots$ , in this order, there is (for large  $n$ ) for each half-edge a probability at least  $\varepsilon$  for some  $\varepsilon > 0$  not to create a double edge; since there are only  $O(1)$  such vertices with  $t_j > 1$ , it follows that  $\mathbb{P}\{G^* \text{ is simple}\}$  is bounded below.

### Acknowledgements

I thank Malwina Luczak for helpful comments. This research was partly supported by the Knut and Alice Wallenberg Foundation.

### References

- [1] BÉKÉSSY, A., BÉKÉSSY, P. AND KOMLÓS, J. (1972). Asymptotic enumeration of regular matrices. *Studia Sci. Math. Hungar.* **7**, 343–353.
- [2] BENDER, E. A. AND CANFIELD, E. R. (1978). The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory A* **24**, 296–307.
- [3] BLANCHET, J. AND STAUFFER, A. (2013). Characterizing optimal sampling of binary contingency tables via the configuration model. *Random Structures Algorithms* **42**, 159–184.
- [4] BOLLOBÁS, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Europ. J. Combinatorics* **1**, 311–316.
- [5] BOLLOBÁS, B. (2001). *Random Graphs*, 2nd edn. Cambridge University Press.
- [6] BOLLOBÁS, B. AND RIORDAN, O. (2012). An old approach to the giant component problem. Preprint. Available at <http://arxiv.org/abs/1209.3691v1>.
- [7] FLAJOLET, P. AND SEDGEWICK, R. (2009). *Analytic Combinatorics*. Cambridge University Press.
- [8] GREENHILL, C., MCKAY, B. D. AND WANG, X. (2006). Asymptotic enumeration of sparse 0-1 matrices with irregular row and column sums. *J. Combinatorial Theory A* **113**, 291–324.
- [9] GUT, A. (2013). *Probability: A Graduate Course*, 2nd edn. Springer, New York.
- [10] JANSON, S. (2009). The probability that a random multigraph is simple. *Combinatorics Prob. Comput.* **18**, 205–225.
- [11] JANSON, S. AND LUCZAK, M. J. (2008). Asymptotic normality of the  $k$ -core in random graphs. *Ann. Appl. Prob.* **18**, 1085–1137.
- [12] JANSON, S., LUCZAK, M. AND WINDRIDGE, P. (2014). Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures Algorithms* **45**, 724–761.
- [13] MCKAY, B. D. (1984). Asymptotics for 0-1 matrices with prescribed line sums. In *Enumeration and Design* (Waterloo, Ontario, 1982), Academic Press, Toronto, ON, pp. 225–238.
- [14] MCKAY, B. D. (1985). Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combinatoria* **19A**, 15–25.
- [15] MCKAY, B. D. AND WORMALD, N. C. (1991). Asymptotic enumeration by degree sequence of graphs with degrees  $o(n^{1/2})$ . *Combinatorica* **11**, 369–382.
- [16] OLVER, F. W. J., LOZIER, D. W., BOISVERT, R. F. AND CLARK, C. W. (2010). *NIST Handbook of Mathematical Functions*. Cambridge University Press.
- [17] WORMALD, N. C. (1978). Some problems in the enumeration of labelled graphs. Doctoral Thesis, University of Newcastle.
- [18] WORMALD, N. C. (1981). The asymptotic distribution of short cycles in random regular graphs. *J. Combinatorial Theory B* **31**, 168–182.

SVANTE JANSON, *Uppsala University*

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden.

Email address: [svante.janson@math.uu.se](mailto:svante.janson@math.uu.se)