THE PROBABILITY THAT A RANDOM MULTIGRAPH IS SIMPLE, II

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 $\frac{1}{2}\sum_i d_i \to \infty$, the probability that the multigraph is simple stays away from 0 if and only if $\sum_i d_i^2 = O\bigl(\sum_i d_i\bigr)$. 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.

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 there are any such graphs at all; in particular, $\sum_i d_i$ has to be even.) A standard method to study $G(n, (d_i)_1^n)$ 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 Bender and Canfield [2] and Wormald [17; 18] for related constructions.) Note that $G^*(n, (d_i)_1^n)$ is defined for all $n \geqslant 1$ 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

$$\liminf_{n\to\infty} \mathbb{P}\bigl(G^*(n, (d_i)_1^n) \text{ is simple}\bigr) > 0 \tag{1.1}$$

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 $n \to \infty$. 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

Date: 23 July, 2013; revised 5 March 2014.

2010 *Mathematics Subject Classification.* 05C80; 05C30, 60C05.

Partly supported by the Knut and Alice Wallenberg Foundation.

to 1 as $n \to \infty$ does so for $G(n, (d_i)_1^n)$ too. (However, note also that Bollobás and Riordan [6] have recently shown that the method may be applied even when (1.1) does not hold; in the problem they study, the probability that $G^*(n, (d_i)_1^n)$ is simple may be almost exponentially small, but they show that the error probability for the studied properties 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. We will throughout the paper let

$$N := \sum_i d_i, \tag{1.2}$$

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 $N/2$. (The reader that makes a detailed comparison with [10] should note that the notation differs slightly.)

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

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

Remark 1.2. 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. We could, with only notational changes, instead use an independent index ν as in [10], assuming that $n = n_\nu \to \infty$.

Note also that if we assume $n = O(N)$, which always may 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 X_{ij} 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. We define

$$Z := \sum_{i=1}^n X_i + \sum_{i<j} Y_{ij}; \tag{1.3}$$

thus $G^*(n, (d_i)_1^n)$ is simple $\iff Z = 0$.

As shown in [10], in the case $\max_i d_i = o(N^{1/2})$, it is not difficult to show Theorem 1.1 by the method used by Bollobás [4; 5], proving a Poisson approximation of Z by the method of moments. In general, however, $\max_i d_i$ may be of the order $N^{1/2}$ 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 1.3 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] are enough.

We define (with some hindsight)

$$\lambda_i := \binom{d_i}{2}\frac{1}{N} = \frac{d_i(d_i-1)}{2N} \tag{1.4}$$

and, for $i \neq j$,

$$\lambda_{ij} := \frac{\sqrt{d_i(d_i-1)d_j(d_j-1)}}{N}, \tag{1.5}$$

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

$$\widehat{X}_i \sim \mathrm{Po}(\lambda_i), \qquad \widehat{X}_{ij} \sim \mathrm{Po}(\lambda_{ij}). \tag{1.6}$$

In analogy with (1.3), we further define $\widehat{Y}_{ij} := \binom{\widehat{X}_{ij}}{2}$ and

$$\widehat{Z} := \sum_{i=1}^n \widehat{X}_i + \sum_{i<j} \widehat{Y}_{ij} = \sum_{i=1}^n \widehat{X}_i + \sum_{i<j}\binom{\widehat{X}_{ij}}{2}. \tag{1.7}$$

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 1.3 ([10])**.** *Assume that $n \to \infty$ and $N \to \infty$. Then*

$$\begin{aligned}\mathbb{P}\big(G^*(n,(d_i)_1^n) \textit{ is simple}\big) &= \mathbb{P}(Z=0) = \mathbb{P}(\widehat{Z}=0) + o(1)\\ &= \exp\Big(-\sum_i \lambda_i - \sum_{i<j}\big(\lambda_{ij} - \log(1+\lambda_{ij})\big)\Big) + o(1).\end{aligned}$$

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

Lemma 1.4. *Suppose that $\sum_i d_i^2 = O(N)$. Then, for every fixed $m \geqslant 1$,*

$$\mathbb{E}\, Z^m = \mathbb{E}\, \widehat{Z}^m + O\big(N^{-1/2}\big). \tag{1.8}$$

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

Remark 1.5. The proof shows that the error term $O(N^{-1/2})$ in (1.8) may be replaced by $O(\max_i d_i/N)$, which always is at least as good by (3.1).

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

Acknowledgement. I thank Malwina Luczak for helpful comments.

2. Preliminaries

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

Lemma 2.1. *Let $X \in \mathrm{Po}(\lambda)$ and let $Y := \binom{X}{2}$. Then, for every $m \geqslant 1$, $\mathbb{E}(Y)_m = h_m(\lambda)$ 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 $|\lambda| \leqslant 1$, and if $m \geqslant 2$, then h_m has a triple root at 0, so $h_m(\lambda) = O(|\lambda|^3)$ for $|\lambda| \leqslant 1$.*

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 λ of degree k for every $k \geqslant 0$.

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

$$h_m(\lambda) = \sum_{j=3}^{\infty} \left(\binom{j}{2}\right)_m \frac{\lambda^j}{j!} e^{-\lambda} = O(\lambda^3) \tag{2.1}$$

as $\lambda \to 0$, 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

$$h_1(\lambda) = \mathbb{E}\, Y = \mathbb{E}(X(X-1)/2) = \tfrac{1}{2}\lambda^2. \tag{2.2}$$

□

Lemma 2.2. *Let $\widehat{Z}$ be given by (1.7) and assume that $\lambda_{ij} = O(1)$.*

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

$$\mathbb{E} \exp\Bigl(t\sqrt{\widehat{Z}}\Bigr) = \exp\biggl(O\biggl(\sum_i \lambda_i + \sum_{i<j} \lambda_{ij}^2\biggr)\biggr). \tag{2.3}$$

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

$$\bigl(\mathbb{E}\, \widehat{Z}^m\bigr)^{1/m} = O(m^2), \tag{2.4}$$

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

Proof. (i): By (1.7),

$$\sqrt{\widehat{Z}} \leqslant \sum_i \sqrt{\widehat{X}_i} + \sum_{i<j} \sqrt{\widehat{Y}_{ij}} \leqslant \sum_i \widehat{X}_i + \sum_{i<j} \widehat{X}_{ij} \mathbf{1}\{\widehat{X}_{ij} \geqslant 2\}, \tag{2.5}$$

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

$$\mathbb{E}\, e^{t\widehat{X}_i} = \exp\bigl((e^t - 1)\lambda_i\bigr) = \exp\bigl(O(\lambda_i)\bigr) \tag{2.6}$$

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

$$\begin{aligned}\mathbb{E}\exp\Bigl(t\widehat{X}_{ij}\mathbf{1}\{\widehat{X}_{ij} \geqslant 2\}\Bigr) &= \mathbb{E}\, e^{t\widehat{X}_{ij}} - \mathbb{P}(\widehat{X}_{ij} = 1)(e^t - 1) \\ &= e^{(e^t-1)\lambda_{ij}} - (e^t - 1)\lambda_{ij}e^{-\lambda_{ij}} \\ &= 1 + (e^t - 1)\lambda_{ij}\bigl(1 - e^{-\lambda_{ij}}\bigr) + O\bigl(\lambda_{ij}^2\bigr) \\ &= 1 + O\bigl(\lambda_{ij}^2\bigr) \leqslant \exp\bigl(O(\lambda_{ij}^2)\bigr). \end{aligned} \tag{2.7}$$

Consequently, (2.3) follows from (2.5)–(2.7).

(ii): Taking $t = 1$, (i) yields $\exp\bigl(\sqrt{\widehat{Z}}\bigr) \leqslant C_1$ for some C_1. Since $\exp\bigl(\sqrt{\widehat{Z}}\bigr) \geqslant \widehat{Z}^m/(2m)!$, this implies

$$\mathbb{E}\,\widehat{Z}^m \leqslant (2m)!\;\mathbb{E}\exp\bigl(\sqrt{\widehat{Z}}\bigr) \leqslant C_1(2m)^{2m}, \tag{2.8}$$

and thus $\mathbb{E}\bigl(\widehat{Z}^m\bigr)^{1/m} \leqslant 4C_1m^2$ for all $m \geqslant 1$. □

3. Proof of Lemma 1.4

Our proof of Lemma 1.4 is rather long, although based on simple calculations, and we will formulate a couple of intermediate steps as separate lemmas. We begin by noting that the assumption $\sum_i d_i^2 = O(N)$ implies

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

and thus, see (1.4)–(1.5),

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

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

$$\sum_i \lambda_i^m = O\Bigl(\sum_i \lambda_i\Bigr) = O\biggl(\frac{\sum_i d_i^2}{N}\biggr) = O(1).$$

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

$$\sum_{i<j} \lambda_{ij}^m = O\Bigl(\sum_{i<j} \lambda_{ij}^2\Bigr) = O\biggl(\frac{\sum_{ij} d_i^2 d_j^2}{N^2}\biggr) = O(1).$$

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 $d_i = d \geqslant 2$ and all $\binom{n}{2}$ λ_{ij} equal to $d(d-1)/N = (d-1)/n$, so their sum is $(d-1)(n-1)/2$. This complicates the proof, since it forces us to obtain error estimates involving λ_{ij}^2.

Let $\mathcal{H}_i$ be the set of the half-edges at vertex i; thus $|\mathcal{H}_i| = d_i$. Further, 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 α and β are joined to an edge in our random pairing. (Thus $I_{\alpha\beta} = I_{\beta\alpha}$.) Note that

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

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

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

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

for any fixed ℓ 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 $\ell \geqslant 1$ and all i,*

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

Proof. We may assume that $d_i \geqslant 2$, since the case $d_i \leqslant 1$ 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

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

Similarly, for any fixed $\ell \geqslant 2$, there are $2^{-\ell}(d_i)_{2\ell}$ ways to select a sequence of ℓ disjoint (unordered) pairs of half-edges at i, and thus by (3.6), using (1.4), (3.1), (3.2) and (1.6),

$$\begin{aligned}
\mathbb{E}(X_i)_\ell &= \frac{(d_i)_{2\ell}}{2^\ell N^\ell}\bigl(1+O(N^{-1})\bigr) = \frac{(d_i(d_1-1))^\ell + O\bigl(d_i^{2\ell-1}\bigr)}{2^\ell N^\ell}\bigl(1+O(N^{-1})\bigr)\\
&= \lambda_i^\ell\bigl(1+O(N^{-1})\bigr) + O\Bigl(\frac{d_i}{N}\lambda_i^{\ell-1}\Bigr)\\
&= \lambda_i^\ell + O\bigl(N^{-1}\lambda_i^\ell\bigr) + O\bigl(N^{-1/2}\lambda_i^{\ell-1}\bigr)\\
&= \lambda_i^\ell + O\bigl(N^{-1/2}\lambda_i\bigr)\\
&= \mathbb{E}(\widehat{X}_i)_\ell + O\bigl(N^{-1/2}\lambda_i\bigr).
\end{aligned} \tag{3.9}$$

The conclusion (3.7) now follows from (3.8)–(3.9) and the standard relations between moments and factorial moments, together with (3.2). $\square$

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 $\ell \geqslant 1$ and all $i \neq j$,*

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

Proof. We may assume $d_i, d_j \geqslant 2$ 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 $\frac{1}{2} d_i d_j (d_i - 1)(d_j - 1)$ ways, and thus, by (3.6), (1.5) and (2.2),

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

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

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

where we only sum over terms such that the ℓ 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, $\tilde{I}_{\alpha\beta}$ are i.i.d. $\mathrm{Be}(1/N)$.) Let, in analogy with (3.5),

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

and let $\widetilde{Y}_{ij} := \binom{\widetilde{X}_{ij}}{2}$. Then, $(\widetilde{Y}_{ij})_\ell$ 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 2ℓ indicators; however, there may be repetitions among the indicators, so each term is a product of r distinct indicators where $r \leqslant 2\ell$. Since we assume $\ell \geqslant 2$, and the pairs $\{\{\alpha_k, \beta_k\}, \{\alpha'_k, \beta'_k\}\}$ are distinct, $r \geqslant 3$ for each term.

Taking expectations and using (3.6), we see that the terms in (3.12) where all occuring pairs $\{\alpha_k, \beta_k\}$ are distinct yield the same contributions to $\mathbb{E}(Y_{ij})_\ell$ and $\mathbb{E}(\widetilde{Y}_{ij})_\ell$, apart from a factor $\big(1 + O(N^{-1})\big)$.

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 $(\widetilde{Y}_{ij})_\ell$ do not. The number of such terms for a given $r \leqslant 2\ell$ is $O(d_i^{r-1} d_j^r + d_i^r d_j^{r-1})$ and thus their contribution to $\mathbb{E}(\widetilde{Y}_{ij})_\ell$ is, using (3.6) and (3.1),

$$O\left(\frac{d_i^{r-1} d_j^r + d_i^r d_j^{r-1}}{N^r} \right) = O\left(\frac{d_j + d_i}{N} \lambda_{ij}^{r-1} \right) = O\left(N^{-1/2} \lambda_{ij}^{r-1} \right). \tag{3.14}$$

Summing over $3 \leqslant r \leqslant 2\ell$, this yields, using (3.2), a total contribution $O\left(N^{-1/2} \lambda_{ij}^2 \right)$. Consequently, we have

$$\mathbb{E}(Y_{ij})_\ell = \mathbb{E}(\widetilde{Y}_{ij})_\ell \big(1 + O(N^{-1})\big) + O\left(N^{-1/2} \lambda_{ij}^2 \right). \tag{3.15}$$

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

$$\check{X}_{ij} := \sum_{\alpha\in\mathcal{H}_i,\beta\in\mathcal{H}_j} J_{\alpha\beta} \sim \mathrm{Po}\left(\frac{d_i d_j}{N}\right), \tag{3.16}$$

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 $J_{\alpha\beta}$. We take the expectation and note that the only difference from $\mathbb{E}(\widetilde{Y}_{ij})_\ell$ is for terms where some $J_{\alpha\beta}$ is repeated. We have, for any fixed $k \geqslant 1$,

$$\mathbb{E}\, J_{\alpha\beta}^k = \frac{1}{N} + O(N^{-2}) = \frac{1}{N}\bigl(1 + O(N^{-1})\bigr), \tag{3.17}$$

while

$$\mathbb{E}\, \tilde{I}_{\alpha\beta}^k = \mathbb{E}\, \tilde{I}_{\alpha\beta} = \frac{1}{N}. \tag{3.18}$$

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

$$\mathbb{E}(\check{Y}_{ij})_\ell = \mathbb{E}(\widetilde{Y}_{ij})_\ell\bigl(1 + O(N^{-1})\bigr). \tag{3.19}$$

Note that we here use $\check{Y}_{ij} = \binom{\check{X}_{ij}}{2}$, where $\check{X}_{ij} \sim \mathrm{Po}(d_i d_j/N)$ has a mean $\check{\lambda}_{ij} := d_i d_j/N$ that differs from $\mathbb{E}\,\widehat{X}_{ij} = \lambda_{ij}$ given by (1.5). We have

$$\check{\lambda}_{ij} \geqslant \lambda_{ij} \geqslant \frac{(d_i - 1)(d_j - 1)}{N} > \check{\lambda}_{ij} - \frac{d_i + d_j}{N}. \tag{3.20}$$

We use Lemma 2.1 and note that the lemma implies that $h'_\ell(\lambda) = O(\lambda^2)$ for each $\ell \geqslant 2$ and $\lambda = O(1)$. Hence, by (3.20) and (3.1)–(3.2),

$$\begin{aligned}\mathbb{E}(\check{Y}_{ij})_\ell - \mathbb{E}(\widehat{Y}_{ij})_\ell &= h_\ell(\check{\lambda}_{ij}) - h_\ell(\lambda_{ij}) = O\bigl(\check{\lambda}_{ij}^2(\check{\lambda}_{ij} - \lambda_{ij})\bigr)\\ &= O\Bigl(\frac{d_i + d_j}{N}\check{\lambda}_{ij}^2\Bigr) = O\Bigl(N^{-1/2}\lambda_{ij}^2\Bigr).\end{aligned} \tag{3.21}$$

Finally, (3.15), (3.19) and (3.21) yield, for each $\ell \geqslant 2$,

$$\mathbb{E}(Y_{ij})_\ell = \mathbb{E}(\widehat{Y}_{ij})_\ell\bigl(1 + O(N^{-1})\bigr) + O\Bigl(N^{-1/2}\lambda_{ij}^2\Bigr). \tag{3.22}$$

By (3.11), this holds for $\ell = 1$ too. By (3.2) and Lemma 2.1, for each $\ell \geqslant 1$, $\mathbb{E}(\widehat{Y}_{ij})_\ell = O(\lambda_{ij}^2)$, and thus (3.22) can be written

$$\mathbb{E}(Y_{ij})_\ell = \mathbb{E}(\widehat{Y}_{ij})_\ell + O\Bigl(N^{-1/2}\lambda_{ij}^2\Bigr), \tag{3.23}$$

for each fixed $\ell \geqslant 1$. 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 $\ell \geqslant 1$,

$$\mathbb{E}\, X_i^\ell + \mathbb{E}\, \widehat{X}_i^\ell = O\bigl(\lambda_i + \lambda_i^\ell + N^{-1/2}\lambda_i\bigr) = O(\lambda_i). \tag{3.24}$$

$$\mathbb{E}\, Y_{ij}^\ell + \mathbb{E}\, \widehat{Y}_{ij}^\ell = O\bigl(\lambda_{ij}^2 + \lambda_{ij}^{2\ell} + N^{-1/2}\lambda_{ij}^2\bigr) = O(\lambda_{ij}^2). \tag{3.25}$$

Proof of Lemma 1.4. We uncouple the terms in (1.3) by letting $(I_{\alpha\beta}^{(i,j)})_{\alpha,\beta}$ be independent copies of $(I_{\alpha\beta})_{\alpha,\beta}$, for $1 \leqslant i,j \leqslant n$, and defining, in analogy with (3.4)–(3.5) and (1.3),

$$\overline{X}_i := \sum_{\alpha,\beta\in\mathcal{H}_i\colon \alpha<\beta} I_{\alpha\beta}^{(i,i)}, \tag{3.26}$$

$$\overline{X}_{ij} := \sum_{\alpha\in\mathcal{H}_i,\beta\in\mathcal{H}_j} I_{\alpha\beta}^{(i,j)}, \tag{3.27}$$

$$\overline{Y}_{ij} := \binom{\overline{X}_{ij}}{2}, \tag{3.28}$$

$$\overline{Z} := \sum_{i=1}^{n} \overline{X}_i + \sum_{i<j} \overline{Y}_{ij}. \tag{3.29}$$

Note that the summands in (3.26) are not independent; they have the same structure as $(I_{\alpha\beta})_{\alpha,\beta\in\mathcal{H}_i}$ and thus $\overline{X}_i \overset{\mathrm{d}}{=} X_i$, and similarly $X_{ij} \overset{\mathrm{d}}{=} X_{ij}$. However, different sums $\overline{X}_i$ and $\overline{X}_{ij}$ are independent (unlike X_i and X_{ij}).

We begin by comparing $\mathbb{E}\,\overline{Z}^m$ and $\mathbb{E}\,\widehat{Z}^m$. Since the terms in (3.29) are independent, the moment $\mathbb{E}\,\overline{Z}^m$ can be written as a certain polynomial $g_m\big(\mathbb{E}\,X_i^\ell, \mathbb{E}\,Y_{ij}^\ell : i,j\in[n],\ \ell\leqslant m\big)$ in the moments $\mathbb{E}(\overline{X}_i)^\ell = \mathbb{E}\,X_i^\ell$ and $\mathbb{E}(\overline{X}_{ij})^\ell = \mathbb{E}\,X_{ij}^\ell$ for $1\leqslant \ell\leqslant m$ and $i,j\in[n]$.

By (1.7), $\mathbb{E}\,\widehat{Z}^m$ can be expressed in the same way as $g_m\big(\mathbb{E}\,\widehat{X}_i^\ell, \mathbb{E}\,\widehat{Y}_{ij}^\ell : i,j\in[n],\ \ell\leqslant m\big)$ for the same polynomial g_m. It follows that

$$\mathbb{E}\,\overline{Z}^m - \mathbb{E}\,\widehat{Z}^m = \sum_{\ell=1}^{m}\sum_i \big(\mathbb{E}\,X_i^\ell - \mathbb{E}\,\widehat{X}_i^\ell\big)R_{\ell i} + \sum_{\ell=1}^{m}\sum_{i<j}\big(\mathbb{E}\,Y_{ij}^\ell - \mathbb{E}\,\widehat{Y}_{ij}^\ell\big)R_{\ell ij} \tag{3.30}$$

for some polynomials $R_{\ell i}$ and $R_{\ell ij}$ in the moments $\mathbb{E}\,X_i^k$, $\mathbb{E}\,\widehat{X}_i^k$, $\mathbb{E}\,Y_{ij}^k$, $\mathbb{E}\,\widehat{Y}_{ij}^k$, for $k\leqslant m$; it is easily seen from (3.24)–(3.25) and (3.3) that

$$R_{\ell i}, R_{\ell ij} = O\left(\sum_{\nu=1}^{m}\Big(\sum_i \lambda_i + \sum_{i<j}\lambda_{ij}^2\Big)^{\nu}\right) = O(1) \tag{3.31}$$

uniformly in $i,j\in[n]$ and $\ell\leqslant m$. Hence, (3.30) yields, together with (3.7), (3.10) and (3.3),

$$\mathbb{E}\,\overline{Z}^m - \mathbb{E}\,\widehat{Z}^m = O\bigg(\sum_i N^{-1/2}\lambda_i + \sum_{i<j} N^{-1/2}\lambda_{ij}^2\bigg) = O\big(N^{-1/2}\big). \tag{3.32}$$

It remains to compare $\mathbb{E}\,Z^m$ and $\mathbb{E}\,\overline{Z}^m$. By (1.3) and (3.4)–(3.5), Z^m can be expanded as a sum of certain products

$$I_{\alpha_1\beta_1}\cdots I_{\alpha_\ell\beta_\ell} \tag{3.33}$$

where $1\leqslant \ell\leqslant 2m$ and we may assume that the pairs $\{\alpha_1,\beta_1\},\dots,\{\alpha_\ell,\beta_\ell\}$ are distinct. (Some products (3.33) may be repeated in Z^m, but only $O(1)$

times.) Furthermore, by (3.26)–(3.29), $\overline{Z}^m$ is the sum of the corresponding products

$$\overline{I}_{\alpha_1\beta_1}\cdots\overline{I}_{\alpha_\ell\beta_\ell}, \tag{3.34}$$

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

We say that a product (3.33) or (3.34) 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, β_ν and β_μ, 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 β_ν and β_μ belong to the same vertex.) It follows from (3.6) that for each good product, the corresponding contributions to $\mathbb{E}\,Z^m$ and $\mathbb{E}\,\overline{Z}^m$ differ only by a factor $\bigl(1+O(N^{-1})\bigr)$. For a bad product, however, the contribution to $\mathbb{E}\,Z^m$ is 0. We thus have to estimate the contribution to $\mathbb{E}\,\overline{Z}^m$ of the bad products.

We define the *support* of a product (3.34) as the multigraph with vertex set $[n]$ and edge set $\{\alpha_\nu\beta_\nu : 1\leqslant\nu\leqslant\ell\}$, 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 (3.34), then F thus has ℓ edges (possibly including loops). Furthermore, it follows from (3.26)–(3.29) 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 non-zero 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 $\delta_i\geqslant 2$. Moreover, if (3.34) 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 $\delta_i\geqslant 4$.

Let F be a multigraph with vertex set $[n]$ and ℓ edges, and denote again its vertex degrees by $\delta_1,\dots,\delta_n$. Thus $\sum_i\delta_i=2\ell$. Let S_F be the contribution to $\mathbb{E}\,\overline{Z}^m$ from bad products (3.33) 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. Furthermore, as just shown, this can only occur for i with $\delta_i\geqslant 4$. 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 $\delta_j\neq 1$,

$$\begin{aligned} S_F &= O\left(N^{-\ell}\sum_{i:\delta_i\geqslant 4} d_i^{\delta_i-1}\prod_{j\neq i} d_j^{\delta_j}\right)\\ &= O\left(N^{-1/2}\sum_{i:\delta_i\geqslant 4}\left(\frac{d_i}{N^{1/2}}\right)^{\delta_i-1}\prod_{j\neq i}\left(\frac{d_j}{N^{1/2}}\right)^{\delta_j}\right)\\ &= O\left(N^{-1/2}\prod_{j:\delta_j>0}\left(\frac{d_j}{N^{1/2}}\right)^{2}\right). \end{aligned} \tag{3.35}$$

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

$$\sum_F S_F = O\left(N^{-1/2}\sum_F \prod_{j:\delta_j>0} \frac{d_j^2}{N}\right). \tag{3.36}$$

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

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

and (3.36) yields

$$\sum_F S_F = O\bigl(N^{-1/2}\bigr). \tag{3.37}$$

Summarizing, the argument above yields

$$\mathbb{E}\,Z^m = \mathbb{E}\,\overline{Z}^m\bigl(1+O(N^{-1})\bigr) + O\bigl(N^{-1/2}\bigr) = \mathbb{E}\,\overline{Z}^m + O\bigl(N^{-1/2}\bigr), \tag{3.38}$$

since $\mathbb{E}\,\overline{Z}^m = O(1)$, e.g. by (3.32) and Lemma 2.2. (Or by arguing similarly as above, summing over supports.)

The lemma follows from (3.32) and (3.38). ◻

4. Proof of theorems 1.1 and 1.3

We first assume that $\sum_i d_i^2 = O(N)$ and prove the following, more precise, statement.

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

Proof. Note that the assumption $\sum_i d_i^2 = O(N)$ 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 $\widehat{Z} \xrightarrow{\mathrm{d}} Z_\infty$ for some random variable Z_∞. Furthermore, the estimate $\mathbb{E}\,\widehat{Z}^m = O(1)$ for each m implies that $\widehat{Z}^m$ is uniformly integrable for each $m \geqslant 1$, and thus

$$\mathbb{E}\,\widehat{Z}^m \to Z_\infty^m, \tag{4.1}$$

see e.g. [9, Theorems 5.4.2 and 5.5.9]. By Lemma 1.4, we thus also have

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

for each $m \geqslant 1$. Furthermore, by (2.4) and (4.1),

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

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

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

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

Hence $\widehat{Z} \xrightarrow{\mathrm{d}} Z_\infty$ and $Z \xrightarrow{\mathrm{d}} Z_\infty$, and thus

$$d_{\mathrm{TV}}(Z,\widehat{Z}) \leqslant d_{\mathrm{TV}}(Z,Z_\infty) + d_{\mathrm{TV}}(\widehat{Z},Z_\infty) \to 0. \tag{4.5}$$

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

Remark 4.2. 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 $d_1 = d_2 \sim N^{1/2}$ when $\lambda_{12} \to 1$ and $Z_\infty \geqslant \binom{\widehat{X}}{2}$ with $\widehat{X} \sim \mathrm{Po}(1)$. In this case, furthermore, by Minkowski's inequality,

$$\begin{aligned}\bigl(\mathbb{E}\, Z_\infty^m\bigr)^{1/m} &\geqslant \frac12\bigl(\mathbb{E}(\widehat{X}^2-\widehat{X})^m\bigr)^{1/m} \geqslant \frac12\bigl(\mathbb{E}\,\widehat{X}^{2m}\bigr)^{1/m} - \frac12\bigl(\mathbb{E}\,\widehat{X}^{m}\bigr)^{1/m}\\ &\sim \frac12\Bigl(\frac{2m}{e\log m}\Bigr)^2 = \frac{2m^2}{e^2\log^2 m}\end{aligned} \tag{4.6}$$

using simple estimates for the moments $\mathbb{E}\,\widehat{X}^m$ when $\widehat{X}\sim\mathrm{Po}(1)$, which are the Bell numbers. (Or by more precise asymptotics in e.g. [7, Proposition VIII.3] and [16, §26.7].) Hence, in this case, $\sum_m \bigl(\mathbb{E}\, Z_\infty^m\bigr)^{-1/m} < \infty$; in other words, Z_∞ does not satisfy the usual Carleman criterion $\sum_m \bigl(\mathbb{E}\, Z_\infty^m\bigr)^{-1/m} = \infty$ for the distribution of to be determined by its moments. However, since we here deal with non-negative 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_∞. Alternatively, we may observe that (4.2) implies $\mathbb{E}(\pm\sqrt{Z})^k \to \mathbb{E}(\pm\sqrt{Z_\infty})^k$ for all $k\geqslant 0$, where the moments trivially vanish when k is odd; since $\pm\sqrt{Z_\infty}$ has a finite moment generating function by (2.3) and Fatou's lemma, the usual sufficient condition for the method of moments yields $\pm\sqrt{Z} \xrightarrow{\mathrm{d}} \pm\sqrt{Z_\infty}$, and thus $Z \xrightarrow{\mathrm{d}} Z_\infty$.)

Proof of Theorems 1.1 and 1.3. In the case $\sum_i d_i^2 = O(N)$, Theorem 1.3 follows from Lemma 4.1, since

$$\begin{aligned}\mathbb{P}(\widehat{Z}=0) = \mathbb{P}\big(\widehat{X}_i = \widehat{Y}_{ij} = 0 \text{ for all } i,j\big) &= \prod_i \mathbb{P}(\widehat{X}_i = 0) \prod_{i<j} \mathbb{P}(\widehat{X}_{ij} \leqslant 1) \\ &= \prod_i e^{-\lambda_i} \prod_{i<j} (1+\lambda_{ij}) e^{-\lambda_{ij}}. \end{aligned} \tag{4.7}$$

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

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

$$\sum_i \lambda_i = \frac{\sum_i d_i^2 - \sum_i d_i}{2N} = \frac{\sum_i d_i^2}{2N} - \frac{1}{2} \to \infty, \tag{4.8}$$

it follows from (4.7) that $\mathbb{P}(\widehat{Z} = 0) \to 0$, and it remains to show that $\mathbb{P}(Z = 0) \to 0$. We do this by the method used in [10] for this case. We 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, $(\bar{d}_i)_1^{\bar{n}}$ say, satisfies $\sum_i \bar{d}_i^2 \leqslant AN$. (Note that the number N of half-edges is unchanged.) Then, as $N \to \infty$, see [10] for details, $\sum_i \bar{d}_i^2 \sim AN$ and, denoting the new random multigraph by $\overline{G}$ and using Lemma 4.1 together with (4.7) and (4.8) on $\overline{G}$,

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

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

Remark 4.3. The proof of Lemma 4.1 shows that if $\sum_i d_i^2 = O(N)$ and $N \to \infty$, and furthermore $\widehat{Z} \xrightarrow{\text{d}} Z_\infty$ for some random variable Z_∞ (which is a kind of regularity property of the degree sequences $(d_i)_1^n$), then also $Z \xrightarrow{\text{d}} Z_\infty$, 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, and we let $\mathcal{L}$ be the event that at most $\log n$ vertices will be infected. It is shown in [12] that for the multigraph $G^*(n,(d_i)_1^n)$, $\mathbb{P}(\mathcal{L}) \to \varkappa$

for some $\varkappa > 0$, and we want to conclude that, assuming $\sum_i d_i^2 = O(N)$, the same is true for the simple random graph $G(n, (d_i)_1^n)$, i.e., that

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

where as above Z is the number of loops and pairs of parallel edges. By considering a subsequence, we may assume that $Z \xrightarrow{\mathrm{d}} Z_\infty$ for some random variable Z_∞, see Remark 4.3. Then, using $\mathbb{P}(\mathcal{L}) \to \varkappa > 0$ and $\liminf \mathbb{P}(Z = 0) > 0$ (Theorem 1.1), (5.1) is equivalent to $\mathbb{P}(\mathcal{L}$ and $Z = 0) \to \varkappa \mathbb{P}(Z_\infty = 0)$ and thus to

$$\mathbb{P}(Z = 0 \mid \mathcal{L}) \to \mathbb{P}(Z_\infty = 0). \tag{5.2}$$

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

$$\mathbb{E}(Z^m \mid \mathcal{L}) \to \mathbb{E}\, Z_\infty^m. \tag{5.3}$$

Actually, for technical reasons, we show a modification of (5.3): we split $Z = Z_1 + Z_2$, where Z_2 is the number of loops and pairs of parallel edges that include an initially infected vertex. It is easily shown that $\mathbb{E}\, Z_2 \to 0$, and thus it suffices to show that

$$\mathbb{E}(Z_1^m \mid \mathcal{L}) \to \mathbb{E}\, Z_\infty^m. \tag{5.4}$$

In order to do this, we write $Z_1^m = \sum_\gamma I_\gamma$, where I_γ 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 γ, if we condition on $I_\gamma = 1$, we have another instance of the configuration model, with the degrees at the vertices involved in γ reduced, plus some extra edges giving γ, and it is easy to see that the result $\mathbb{P}(\mathcal{L}) \to \varkappa$ applies to this modification too, and thus

$$\mathbb{P}(\mathcal{L} \mid I_\gamma = 1) = \varkappa + o(1) \tag{5.5}$$

uniformly for all γ. We invert the conditioning again and obtain

$$\mathbb{E}(I_\gamma \mid \mathcal{L}) = \frac{\mathbb{P}(\mathcal{L} \mid I_\gamma = 1)\,\mathbb{P}(I_\gamma = 1)}{\mathbb{P}(\mathcal{L})} = \bigl(1 + o(1)\bigr)\,\mathbb{E}(I_\gamma). \tag{5.6}$$

Consequently,

$$\mathbb{E}\bigl(Z_1^m \mid \mathcal{L}\bigr) = \sum_\gamma \mathbb{E}(I_\gamma \mid \mathcal{L}) \sim \sum_\gamma \mathbb{E}(I_\gamma) = \mathbb{E}\, Z_1^m, \tag{5.7}$$

and since $\mathbb{E}\, Z_1^m \to \mathbb{E}\, Z_\infty^m$, this yields the desired (5.4).

6. Bipartite graphs

A similar result for bipartite graphs has been proved by Blanchet and Stauffer [3]; see e.g. [1], [13], [8] for earlier results. (These results are often stated in an equivalent form about 0-1 matrices.) We suppose we are given 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 $G\bigl(n, (s_i)_1^{n'}, (t_j)_1^{n''}\bigr)$ with these degree sequences as well as the corresponding random bipartite multigraph $G^* = G^*\big(n,(s_i)_1^{n'},(t_j)_1^{n''}\big)$ constructed by the configuration model. (These have N edges.) We order the two degree sequences in decreasing order as $s_{(1)} \geqslant \ldots \geqslant s_{(n')}$ and $t_{(1)} \geqslant \ldots \geqslant t_{(n'')}$, 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 (Blanchet and Stauffer [3])**.** *Assume that $N \to \infty$. Then $\liminf_{n\to\infty} \mathbb{P}\big(G^*\big(n,(s_i)_1^{n'},(t_j)_1^{n''}\big)$ is simple$\big) > 0$ if and only if the following two conditions hold:*

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

(ii) *For any fixed $m \geqslant 1$,*
$$\sum_{i=\min\{t,m\}}^{n'} s_{(i)} = \Omega(N), \tag{6.2}$$
$$\sum_{j=\min\{s,m\}}^{n''} t_{(j)} = \Omega(N). \tag{6.3}$$

(We have reformulated and simplified (ii) from [3]. Recall that $x = \Omega(N)$ means that $\liminf x/N > 0$.)

Remark 6.2. Here (i) corresponds to the condition $\sum_i d_i^2 = O(N)$ in Theorem 1.1, while (ii) is an additional complication. Note that if $s = o(N)$ then (6.2) holds, because the sum is $\geqslant N - (m-1)s$; similarly, if $t = o(N)$ then (6.3) holds. Hence (ii) is satisfied, and (i) 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 $\sum_j t_j(t_j-1) = O(1)$ and thus $t = O(1)$. On the other hand, in such cases, (i) 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 $j \geqslant 2$, then (i) holds but (6.2) fails for $m = 2$. Indeed, in this example, there is w.h.p. (i.e., with probability $1-o(1)$) a double edge v_1w_1, and thus G^* is w.h.p. not simple.

We can prove Theorem 6.1 too 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 $j \in [n'']$,
$$\lambda_{ij} := \frac{\sqrt{s_i(s_i-1)t_j(t_j-1)}}{N}, \tag{6.4}$$
let $\widehat{X}_{ij} \sim \mathrm{Po}(\lambda_{ij})$ 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 6.3. *Assume that $N \to \infty$ and that $s, t = o(N)$. Then*

$$\begin{aligned}\mathbb{P}\bigl(G^*\bigl(n, (s_i)_1^{n'}, (t_j)_1^{n''}\bigr) \textit{ is simple}\bigr) &= \mathbb{P}(Z = 0) = \mathbb{P}(\widehat{Z} = 0) + o(1) \\ &= \exp\Bigl(-\sum_{i,j}\bigl(\lambda_{ij} - \log(1 + \lambda_{ij})\bigr)\Bigr) + o(1).\end{aligned}$$

Proof (sketch). This is proved as Theorem 1.3, using analogues of Lemmas 1.4 and 4.1, with only minor differences. Instead of (3.1) we use the assumption $s, t = o(N)$, which leads to error terms of the order $O((s + t)/N)$, cf. Remark 1.5. Furthermore, (3.35) has to be modifed. Say that the vertex with a repeated half-edge is bad, and suppose that the bad vertex is in the first part. Let the non-zero 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 $\sum_\nu a_\nu = \sum_\mu b_\mu = \ell$. The contribution from all F with given (a_ν) and (b_μ) is, using Hölder's inequality and (6.1),

$$\begin{aligned}O\Bigl(N^{-\ell} &\sum_{i:s_i\geqslant 2} s_i^{a_1-1} \prod_{\nu\geqslant 2}\Bigl(\sum_{i:s_i\geqslant 2} s_i^{a_\nu}\Bigr)\prod_{\mu\geqslant 1}\Bigl(\sum_{j:t_j\geqslant 2} t_j^{b_\mu}\Bigr)\Bigr) \\ &= O\Bigl(N^{-\ell}\Bigl(\sum_{i:s_i\geqslant 2} s_i^2\Bigr)^{(a_1-1)/2+\sum_{\nu\geqslant 2} a_\nu/2}\Bigl(\sum_{j:t_j\geqslant 2} t_j^2\Bigr)^{\sum_\mu b_\mu/2}\Bigr) \\ &= O\Bigl(N^{-\ell}\Bigl(\sum_i s_i(s_i-1)\Bigr)^{(\ell-1)/2}\Bigl(\sum_j t_j(t_j-1)\Bigr)^{\ell/2}\Bigr) \\ &= O\Bigl(N^{-1}\Bigl(\sum_j t_j(t_j-1)\Bigr)^{1/2}\Bigr) = O\bigl(t^{1/2}/N^{1/2}\bigr).\end{aligned}$$

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

Proof of Theorem 6.1. The case $s, t = o(N)$ (when (ii) is automatic) follows from Theorem 6.3; note that

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

By considering subsequences, and symmetry, it remains only to consider the case $s = \Omega(N)$. It is easy to see that (i) is necessary in this case too so we may assume (i). As said above, this implies $t = O(1)$, and furthermore, that only $O(1)$ degrees t_j are > 1. By taking a further subsequence, we may assume that t is constant. Then (6.3) always holds, and it suffices to consider the case $m = t$ in (6.2), i.e.,

$$\sum_{i=t}^{n'} s_{(i)} = \Omega(N). \tag{6.5}$$

If (6.5) does not hold, then (at least for a subsequence), w.h.p. $\sum_{i=t}^{n'} s_{(i)} = o(N)$, and then w.h.p. 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 (6.5) 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 ε for some $\varepsilon > 0$ to not 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. □

References

[1] A. Békéssy, P. Békéssy & J. Komlós, Asymptotic enumeration of regular matrices. *Studia Sci. Math. Hungar.* **7** (1972), 343–353.

[2] E. A. Bender & E. R. Canfield, The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory Ser. A*, **24** (1978), no. 3, 296–307.

[3] J. Blanchet & A. Stauffer, Characterizing optimal sampling of binary contingency tables via the configuration model. *Random Structures Algorithms* **42** (2013), no. 2, 159–184.

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

[5] B. Bollobás, *Random Graphs*, 2nd ed. Cambridge Univ. Press, Cambridge, 2001.

[6] B. Bollobás & O. Riordan, An old approach to the giant component problem. Preprint, 2012. `arXiv:1209.3691.`

[7] P. Flajolet & R. Sedgewick, *Analytic Combinatorics.* Cambridge Univ. Press, Cambridge, UK, 2009.

[8] C. Greenhill, B. D. McKay & X. Wang, Asymptotic enumeration of sparse 0-1 matrices with irregular row and column sums. *J. Combin. Theory Ser. A* **113** (2006), no. 2, 291–324.

[9] A. Gut, *Probability: A Graduate Course*, 2nd ed. Springer, New York, 2013.

[10] S. Janson, The probability that a random multigraph is simple. *Combin. Probab. Comput.* **18** (2009), 205–225.

[11] S. Janson & M. Luczak, Asymptotic normality of the k-core in random graphs. *Ann. Appl. Probab.* **18** (2008), no. 3, 1085–1137.

[12] S. Janson, M. Luczak & P. Windridge, Law of large numbers for the SIR epidemic on a random graph with given degrees. Preprint, 2013. `arXiv:1308.5493.`

[13] B. D. McKay, Asymptotics for 0-1 matrices with prescribed line sums. *Enumeration and design (Waterloo, Ont., 1982)*, pp. 225–238, Academic Press, Toronto, ON, 1984.

[14] B. D. McKay, Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combin.* **19A** (1985), 15–25.

[15] B. D. McKay & N. C. Wormald, Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica* **11** (1991), no. 4, 369–382.

[16] F. W. J. Olver, D. W. Lozier, R. F. Boisvert & C. W. Clark, *NIST Handbook of Mathematical Functions*. Cambridge Univ. Press, 2010. Also available as *NIST Digital Library of Mathematical Functions*. `http://dlmf.nist.gov/`

[17] N. C. Wormald, Some problems in the enumeration of labelled graphs. Ph. D. thesis, University of Newcastle, 1978.

[18] N. C. Wormald, The asymptotic distribution of short cycles in random regular graphs. *J. Combin. Theory Ser. B* **31** (1981), no. 2, 168–182.

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

E-mail address: `svante.janson@math.uu.se` `http://www2.math.uu.se/~svante/`