Degree sequences of random digraphs and bipartite graphs

BRENDAN D. MCKAY[*] AND FIONA SKERMAN[*]

We investigate the joint distribution of the vertex degrees in three models of random bipartite graphs. Namely, we can choose each edge with a specified probability, choose a specified number of edges, or specify the vertex degrees in one of the two colour classes.

This problem can alternatively be described in terms of the row and sum columns of a random binary matrix or the in-degrees and out-degrees of a random digraph, in which case we can optionally forbid loops. It can also be cast as a problem in random hypergraphs, or as a classical occupancy, allocation, or coupon collection problem.

In each case, provided the two colour classes are not too different in size nor the number of edges too low, we define a probability space based on independent binomial variables and show that its probability masses asymptotically equal those of the degrees in the graph model almost everywhere. The accuracy is sufficient to asymptotically determine the expectation of any joint function of the degrees whose maximum is at most polynomially greater than its expectation.

1. Introduction

We will study the joint distributions of the vertex degrees for three different models of random bipartite graphs. In each case, we construct simpler probability spaces which match these distributions to high precision. The new probability spaces are based on independent binomial distributions and allow asymptotic calculations of any random variable which is a function of the degrees and has maximum at most polynomially greater than its expectation. In Section 2.1 we will show an example of such a calculation. Note that our results are much stronger than contiguity or decreasing total variation distance. These results are similar to those obtained by McKay and Wormald [29, 30] for the case of ordinary (not necessarily bipartite) graphs.

We prefer to use graph terminology, but will also describe the problem in the matrix and other settings. Consider a probability space of $m \times n$ matrices over $\{0,1\}$. Three probability spaces will be considered. In the first case, which we call $\mathcal{G}_p$, some number $p \in (0,1)$ is specified and each entry of the matrix is independently equal to 1 with probability p and equal to 0 otherwise. In the second case, which we call $\mathcal{G}_k$, some integer k is specified, and all $m \times n$ binary matrices with exactly k ones have the same probability, and no other matrices are allowed. In the third case, which we call $\mathcal{G}_t$, a list of n integers $t_1, \ldots, t_n$ is specified, and all $m \times n$ binary matrices with column sums $t_1, \ldots, t_n$, respectively, are equally likely and no others are allowed.

We can interpret the matrix as a bipartite graph in the standard fashion. Associate distinct vertices $U = \{u_1, \ldots, u_m\}$ with the rows, and $V = \{v_1, \ldots, v_n\}$ with the columns, and place an edge between u_i and v_j exactly when the matrix entry in position (i,j) equals 1. The row and column sums of the matrix correspond to the degrees of the vertices.

These probability models have also appeared in other settings. Given m bins, at each stage $j = 1, \ldots, n$ throw t_j balls into distinct bins with all $\binom{m}{t_j}$ possible placings equally likely. Then the distribution of the number of balls in each bin $\boldsymbol{S} = (S_1, \ldots, S_m)$ can be studied. This model is referred to as *allocation by complexes* and is precisely our $\mathcal{G}_{\boldsymbol{t}}$ model. If we allow the number of balls thrown to be a random variable T_j, binomially distributed with parameters (m, p), we attain the $\mathcal{G}_p$ model.

Similarly, in the *coupon collection* problem a customer repeatedly buys a random number, T, of distinct coupons from a set of m possible different coupons. This covers both our $\mathcal{G}_p$ case when T is binomially distributed with parameters (m, p) and our $\mathcal{G}_{\boldsymbol{t}}$ case where $T_j = t_j$ with probability 1. (Here, our vector $\boldsymbol{s}$ describes the number of each coupon collected and $\boldsymbol{t}$ the number of coupons collected at each stage.)

Finally, consider a hypergraph on m vertices. At each stage $j = 1, \ldots, n$, choose at random a hyperedge of size t_j, allowing multi-edges. Then if we set S_i to be the number of hyperedges which contain the ith vertex, we obtain the $\mathcal{G}_{\boldsymbol{t}}$ model.

If $m = n$, we can also associate the matrix with a directed graph. There are n vertices $\{w_1, \ldots, w_n\}$. A matrix entry equal to 1 in position (i, j) corresponds to a directed edge from w_i to w_j. The case $i = j$ is permitted, so these directed graphs can have loops. The row and column sums of the matrix correspond to the out-degrees and in-degrees, respectively, of the directed graph. We will also treat the case of loop-free digraphs, which correspond to square matrices with zero diagonal. Our methods would also work if some other limited set of matrix entries are required to be zero, but we have not applied them in that case.

We now continue using the bipartite graph formulation. For each of the three probability spaces of random bipartite graphs, we seek to examine the $(m+n)$-dimensional joint distribution of the vertex degrees. If G is a bipartite graph on $U \cup V$ (respecting the partition into U and V), then $\boldsymbol{s} = \boldsymbol{s}(G) = (s_1, \ldots, s_m)$ is the list of degrees of $u_1, \ldots, u_m$, and $\boldsymbol{t} = \boldsymbol{t}(G) = (t_1, \ldots, t_n)$ is the list of degrees of $v_1, \ldots, v_n$. We call the pair $(\boldsymbol{s}, \boldsymbol{t})$ the *degree sequence* of G.

Define $I_n = \{0, 1, \ldots, n\}$ and $I_{m,n} = I_n^m \times I_m^n$. Also let $G(\boldsymbol{s}, \boldsymbol{t})$ be the number of (labelled) bipartite graphs on $U \cup V$ with degree sequence $(\boldsymbol{s}, \boldsymbol{t})$. In the case of $m = n$, we also define $\vec{G}(\boldsymbol{s}, \boldsymbol{t})$ to be the number of loop-free digraphs with in-degrees $\boldsymbol{s}$ and out-degrees $\boldsymbol{t}$.

For precision we need to distinguish between random variables (written in uppercase) and the values they may take (written in lowercase). For each probability space of random graphs, as determined by the context,

$\boldsymbol{S} = (S_1, \ldots, S_m)$ will denote the random variable given by the degrees in U and $\boldsymbol{T} = (T_1, \ldots, T_n)$ will denote the random variable given by the the degrees in V. We will take $\boldsymbol{S}$ to have range I_n^m and $\boldsymbol{T}$ to have range I_m^n. Also define random variables

$$K = \sum_{i=1}^{m} S_i \quad \text{and} \quad \Lambda = \frac{K}{mn}.$$

As usual, q is an abbreviation for $1 - p$.

1.1. Historical notes

The $\mathcal{G}_t$ model has received wide-ranging attention, in particular the distribution of the number of isolated vertices. This is also a natural question in the alternative (non-graph) wordings of the model. It corresponds to the number of empty bins in the allocation model [3, 13, 19, 31, 44, 38, 12], the number of uncollected coupons in the collector's problem [43, 25], the number of isolated vertices in the hypergraph model and the number of zero rows in the binary matrix model [14]. More generally, the number of vertices with a particular degree (or range of degrees) in $\mathcal{G}_t$ has been studied in allocation [32, 39, 40], graph [7, 23] and matrix models [8]. A different extension on this theme is to study the distribution of the number of draws required to go from i to j non-empty bins [1, 21, 32, 41, 42]. In a similar direction, Khakimullin and Enatskaya studied the distribution of the number of draws to exceed a particular lineup in the bins in the $\mathcal{G}_t$ model [18] and in the i.i.d. case which includes the $\mathcal{G}_p$ model as well [20]. The monograph by Kolchin gives many results on $\mathcal{G}_t$ phrased as the balls and bins model [22].

We are interested in asymptotic results as we take m, n roughly equal as they tend to infinity, but another natural option is to fix m, the number of vertices in one part, and let n, the number of vertices in the other part, tend to infinity. There seems to be a consistent divide in the literature that when considered as a graph the asymptotics of $\mathcal{G}_t$ are studied with m, n both tending towards infinity while the balls and bins and coupon collection articles (including those cited above) fix m and take n tending toward infinity. The latter corresponds to fixing the number of bins and taking the number of balls to infinity or having a fixed number of coupons and letting the number of sampling rounds tend to infinity.

In the other two probability models on bipartite graphs, $\mathcal{G}_p$ and $\mathcal{G}_k$, two types of results are known: those on the minimum and maximum degrees [7, 5, 36] and those on the number of vertices with a given degree [23, 33, 34].

For results in the digraph counterpart $\vec{\mathcal{G}}_p$ see [37] (and below). The model $\mathcal{G}_p$ also appears in papers on ball and bin models. Sometimes the numbers of balls thrown at each stage are allowed to be i.i.d. random variables [16]. If we then take these random variables to be binomially distributed with parameters m, p we recover the $\mathcal{G}_p$ model. Godbole et al. [9] study the number of sets of r mutually threatening rooks. This corresponds to the number of vertices with degree $h \geq r$ weighted by $\binom{h}{r}$ in our $\mathcal{G}_p$ and $\mathcal{G}_k$ models.

Of the papers cited, we highlight some which concern the minimum and maximum degrees, a fixed number of the smallest and largest degrees and the distribution of the hth largest degree.

Khakimullin determined the asymptotic distribution of the hth largest degree when the average degree increases faster than $\log m$ [16]. The model used allowed the numbers of balls allocated at each step to be i.i.d. random variables and so includes both our $\mathcal{G}_p$ and uniform $\mathcal{G}_t$ cases. This extends an earlier result by the same author which gave the asymptotic distribution of the largest degree [17].

Palka and Sperling showed that if we fix p such that $np = w(n) \log n = o(n)$, then any fixed number of the smallest and largest degrees are unique in $\vec{\mathcal{G}}_p$ and in the uniform $\mathcal{G}_t$ model [37]. A similar result for the $\vec{\mathcal{G}}_t$ model is shown by Palka in [35], where $\boldsymbol{t} = (d, d, \ldots, d)$ and $d = w(n) \log n = o(n)$. There is also some work on the degrees in random digraphs by Jaworski and Karoński [15] who showed, in the case that $\boldsymbol{t} = (d, d, \ldots, d)$ and $d = o(n)$, that the minimum vertex degree in $\mathcal{G}_t$ is almost surely the same as that in $\vec{\mathcal{G}}_t$.

1.2. Asymptotic notation

As we are dealing with asymptotics of functions of many variables, we must be careful to define our asymptotic notation.

We will tacitly assume that all variables not declared to be constant are functions of a single underlying index ℓ that takes values $1, 2, \ldots$, and that all asymptotic statements refer to $\ell \to \infty$. Thus, the size parameters m, n are in reality functions $m(\ell)$ and $n(\ell)$, and a statement like $f(m, n) = O(g(m, n))$ means that there is a constant $A > 0$ such that $|f(m, n)| \leq A|g(m, n)|$ when ℓ is large enough. This should not be cause for alarm, because we will invariably impose conditions implying that $m, n \to \infty$ as $\ell \to \infty$.

The expression $\tilde{o}(1)$ represents any function of ℓ of magnitude $O(e^{-n^c})$ for some constant $c > 0$. The constant c might be different for different appearances of the notation. The class $\tilde{o}(1)$ is closed under addition, multiplication, taking positive powers, and multiplication by polynomials in n.

1.3. Graph models

We now define a sequence of finite probability spaces that we call "models", with sample space either $I(m,n) = I_n^m \times I_m^n$ or I_n^m. The probability measure for each model will be defined using random variables $(\boldsymbol{S}, \boldsymbol{T})$ or $\boldsymbol{S}$, respectively, whose distribution equals the respective probability measure. In general our notation will not distinguish between each probability space and its probability measure.

We first consider six models whose probability measures are derived from the degrees of a random bipartite graph or digraph G.

1. (*p-models* $\mathcal{G}_p, \vec{\mathcal{G}}_p$, *for* $0 < p < 1$) Generate G by choosing each of the mn possible edges $u_i v_j$ with probability p, such choices being independent. The probability distribution $\mathcal{G}_p = \mathcal{G}_p(m,n)$ on $I_{m,n}$ is that of the degree sequence $(\boldsymbol{S}, \boldsymbol{T})$ of G. If $m = n$ and the edges $\{u_i v_i\}$ are forbidden, we obtain the probability distribution $\vec{\mathcal{G}}_p$ instead, corresponding to the degree sequences of a loop-free digraph where each possible directed edge is chosen independently with probability p. Note that $G(\boldsymbol{s}, \boldsymbol{t}) = 0$ for many pairs $(\boldsymbol{s}, \boldsymbol{t})$. We have

$$\text{Prob}_{\mathcal{G}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) = p^k q^{mn-k} G(\boldsymbol{s}, \boldsymbol{t}),$$
$$\text{Prob}_{\vec{\mathcal{G}}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) = p^k q^{n^2-n-k} \vec{G}(\boldsymbol{s}, \boldsymbol{t}),$$

   where $q = 1 - p$ and $k = \sum_{i=1}^m s_i$.

2. (*k-models* $\mathcal{G}_k, \vec{\mathcal{G}}_k$, *for integer* $k \geq 0$) Generate G by choosing each of the bipartite graphs on $U \cup V$ having k edges, with equal probability. The probability distribution $\mathcal{G}_k = \mathcal{G}_k(m,n)$ on $I_{m,n}$ is that of the degree sequence $(\boldsymbol{S}, \boldsymbol{T})$ of G. If $m = n$ and the edges $\{u_i v_i\}$ are forbidden, we obtain the distribution $\vec{\mathcal{G}}_k = \vec{\mathcal{G}}_k(n)$ of the degree-sequences for the uniform probability space of all loop-free digraphs with k edges. We have

$$\text{Prob}_{\mathcal{G}_k}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$$
$$= \begin{cases} \binom{mn}{k}^{-1} G(\boldsymbol{s}, \boldsymbol{t}), & \text{if } \sum_{i=1}^m s_i = \sum_{j=1}^n t_j = k; \\ 0, & \text{otherwise,} \end{cases}$$
$$\text{Prob}_{\vec{\mathcal{G}}_k}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$$
$$= \begin{cases} \binom{n^2 - n}{k}^{-1} \vec{G}(\boldsymbol{s}, \boldsymbol{t}), & \text{if } \sum_{i=1}^n s_i = \sum_{j=1}^n t_j = k; \\ 0, & \text{otherwise.} \end{cases}$$

3. (*t-models* $\mathcal{G}_t, \vec{\mathcal{G}}_t$, *for* $t \in I_m^n$) Generate G by choosing each of the bipartite graphs on $U \cup V$ having $t(G) = t$, with equal probability. For consistency we can define the random variable T to have the value t, but since this is constant we will define our probability spaces using S only. The probability distribution $\mathcal{G}_t = \mathcal{G}_t(m)$ on I_n^m is that of the degree sequence S of G in U. If $m = n$ and the edges $\{u_i v_i\}$ are forbidden, we obtain the distribution $\vec{\mathcal{G}}_t = \vec{\mathcal{G}}_t(n)$ of the in-degrees for the uniform probability distribution of all loop-free digraphs with fixed out-degrees t. For a given $t \in I_m^n$, we have

$$\mathrm{Prob}_{\mathcal{G}_t}(S = s) = \prod_{j=1}^n \binom{m}{t_j}^{-1} G(s, t),$$

$$\mathrm{Prob}_{\vec{\mathcal{G}}_t}(S = s) = \prod_{j=1}^n \binom{n-1}{t_j}^{-1} \vec{G}(s, t).$$

The probability spaces $\mathcal{G}_p$, $\mathcal{G}_k$ and $\mathcal{G}_t$ are clearly related, by mixing and conditioning. In particular, for any event $E \subseteq I_{m,n}$ or $E' \subset I_n^m$, the following hold. Note that the first relationships on lines (2) and (3) are independent of p and assume $0 < p < 1$.

(1) $$\mathrm{Prob}_{\mathcal{G}_p}(E) = \sum_{k=0}^{mn} \binom{mn}{k} p^k q^{mn-k} \mathrm{Prob}_{\mathcal{G}_k}(E)$$

$$= \sum_{t \in I_m^n} \left(\prod_{j=1}^n \binom{m}{t_j} p^{t_j} q^{m-t_j} \right) \mathrm{Prob}_{\mathcal{G}_t}(E),$$

(2) $$\mathrm{Prob}_{\mathcal{G}_k}(E) = \mathrm{Prob}_{\mathcal{G}_p}(E \mid K = k)$$

$$= \sum_{t: \sum_{j=1}^n t_j = k} \binom{mn}{k}^{-1} \prod_{j=1}^n \binom{m}{t_j} \mathrm{Prob}_{\mathcal{G}_t}(E),$$

(3) $$\mathrm{Prob}_{\mathcal{G}_t}(E') = \mathrm{Prob}_{\mathcal{G}_p}(E' \times \{t\} \mid T = t)$$

$$= \mathrm{Prob}_{\mathcal{G}_{k=\sum_j t_j}}(E' \times \{t\} \mid T = t),$$

with similar relations between $\vec{\mathcal{G}}_p$, $\vec{\mathcal{G}}_k$ and $\vec{\mathcal{G}}_t$.

Note that the separate distributions of S and T in $\mathcal{G}_p$ and $\mathcal{G}_k$ are elementary. In $\mathcal{G}_p$, the components of S have independent binomial distributions, while in the $\mathcal{G}_k$ model S has a multivariate hypergeometric distribution. The difficulty is in quantifying the dependence between S and T when all $m+n$ components are considered together.

1.4. Binomial models

Our aim is to compare the degree sequence distributions defined above to some distributions derived from independent binomials. Our motivating observation is the known marginal distributions of $\boldsymbol{S}$ and $\boldsymbol{T}$ in the models $\mathcal{G}_p$ and $\mathcal{G}_k$.

1. (*Independent models* $\mathcal{I}_p, \vec{\mathcal{I}}_p$, *for* $0 < p < 1$) Generate m components distributed $\mathrm{Bin}(n, p)$ and n components distributed $\mathrm{Bin}(m, p)$, all $m+n$ components being independent. The joint distribution on $I_{m,n}$ is $\mathcal{I}_p = \mathcal{I}_p(m, n)$. If instead we have $m = n$ and the $2n$ components are all distributed $\mathrm{Bin}(n-1, p)$, the joint distribution on $I_{n,n}$ is $\vec{\mathcal{I}}_p = \vec{\mathcal{I}}_p(n)$. We have

$$\mathrm{Prob}_{\mathcal{I}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$$

$$= p^{\sum_i s_i + \sum_j t_j} q^{2mn - \sum_i s_i - \sum_j t_j} \prod_{i=1}^{m} \binom{n}{s_i} \prod_{j=1}^{n} \binom{m}{t_j},$$

$$\mathrm{Prob}_{\vec{\mathcal{I}}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$$

$$= p^{\sum_i s_i + \sum_j t_j} q^{2n^2 - 2n - \sum_i s_i - \sum_j t_j} \prod_{i=1}^{n} \binom{n-1}{s_i} \prod_{j=1}^{n} \binom{n-1}{t_j}.$$

2. (*Binomial p-models* $\mathcal{B}_p, \vec{\mathcal{B}}_p$, *for* $0 < p < 1$) The distribution $\mathcal{B}_p = \mathcal{B}_p(m, n)$ on $I_{m,n}$ is the conditional distribution of $\mathcal{I}_p$ subject to $\sum_{i=1}^{m} S_i = \sum_{j=1}^{n} T_j$. For $m = n$, the distribution $\vec{\mathcal{B}}_p = \vec{\mathcal{B}}_p(n)$ on $I_{n,n}$ is obtained from $\vec{\mathcal{I}}_p$ by the same conditioning. We have

$$\mathrm{Prob}_{\mathcal{B}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$$

$$= \begin{cases} \dfrac{\mathrm{Prob}_{\mathcal{I}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})}{\mathrm{Prob}_{\mathcal{I}_p}\left(\sum_{i=1}^{m} S_i = \sum_{j=1}^{n} T_j\right)}, & \text{if } \sum_{i=1}^{m} s_i = \sum_{j=1}^{n} t_j; \\ 0, & \text{otherwise,} \end{cases}$$

and similarly for $\vec{\mathcal{B}}_p$.

3. (*Binomial k-models* $\mathcal{B}_k, \vec{\mathcal{B}}_k$, *for integer* $k \geq 0$) The distribution $\mathcal{B}_k = \mathcal{B}_k(m, n)$ on $I_{m,n}$ is the conditional distribution of $\mathcal{I}_p$ subject to $\sum_{i=1}^{m} S_i = \sum_{j=1}^{n} T_j = k$. For $m = n$, $\vec{\mathcal{B}}_k = \vec{\mathcal{B}}_k(n)$ is derived from $\vec{\mathcal{I}}_p$ in the same way. In both cases, the distribution doesn't depend on p.

We have

$$\text{Prob}_{\mathcal{B}_k}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$$

$$= \begin{cases} \binom{mn}{k}^{-2} \prod_{i=1}^{m} \binom{n}{s_i} \prod_{j=1}^{n} \binom{m}{t_j}, & \text{if } \sum_{i=1}^{m} s_i = \sum_{j=1}^{n} t_j = k; \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{Prob}_{\vec{\mathcal{B}}_k}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$$

$$= \begin{cases} \binom{n^2 - n}{k}^{-2} \prod_{i=1}^{n} \binom{n-1}{s_i} \prod_{j=1}^{n} \binom{n-1}{t_j}, & \text{if } \sum_{i=1}^{n} s_i = \sum_{j=1}^{n} t_j = k; \\ 0, & \text{otherwise.} \end{cases}$$

In each case $\boldsymbol{S}$ and $\boldsymbol{T}$ have independent multivariate hypergeometric distributions.

4. (*Binomial $\boldsymbol{t}$-models $\mathcal{B}_{\boldsymbol{t}}, \vec{\mathcal{B}}_{\boldsymbol{t}}$, for $\boldsymbol{t} \in I_m^n$*) The distribution $\mathcal{B}_{\boldsymbol{t}} = \mathcal{B}_{\boldsymbol{t}}(m, n)$ on I_n^m is the distribution of $\boldsymbol{S}$ when $(\boldsymbol{S}, \boldsymbol{T})$ has distribution $\mathcal{B}_k$ for $k = \sum_{j=1}^{n} t_j$. For $m = n$, $\vec{\mathcal{B}}_{\boldsymbol{t}} = \vec{\mathcal{B}}_{\boldsymbol{t}}(n)$ is derived from $\vec{\mathcal{B}}_k$ in the same way. For a given $\boldsymbol{t} \in I_m^n$, we have

$$\text{Prob}_{\mathcal{B}_{\boldsymbol{t}}}(\boldsymbol{S} = \boldsymbol{s}) = \begin{cases} \binom{mn}{k}^{-1} \prod_{i=1}^{m} \binom{n}{s_i} & \text{if } \sum_{i=1}^{m} s_i = \sum_{j=1}^{n} t_j; \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{Prob}_{\vec{\mathcal{B}}_{\boldsymbol{t}}}(\boldsymbol{S} = \boldsymbol{s}) = \begin{cases} \binom{n^2 - n}{k}^{-1} \prod_{i=1}^{n} \binom{n-1}{s_i} & \text{if } \sum_{i=1}^{n} s_i = \sum_{j=1}^{n} t_j; \\ 0, & \text{otherwise.} \end{cases}$$

In each case, $\boldsymbol{S}$ has a multivariate hypergeometric distribution.

5. (*Integrated p-models $\mathcal{V}_p, \vec{\mathcal{V}}_p$, for $0 < p < 1$*) The distribution $\mathcal{V}_p = \mathcal{V}_p(m, n)$ on $I_{m,n}$ is a mixture of $\mathcal{B}_{p'}$ distributions, while for $m = n$ the distribution $\vec{\mathcal{V}}_p = \vec{\mathcal{V}}_p(n)$ on $I_{n,n}$ is a mixture of $\vec{\mathcal{B}}_{p'}$ distributions. Let

$$K_p(p') = \left(\frac{mn}{\pi pq}\right)^{1/2} \exp\left(-\frac{mn}{pq}(p' - p)^2\right),$$

$$V(p) = \int_0^1 K_p(p')\, dp'.$$

Then we define

$$\mathrm{Prob}_{\mathcal{V}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) = V(p)^{-1} \int_0^1 K_p(p') \, \mathrm{Prob}_{\mathcal{B}_{p'}}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) \, dp',$$

$$\mathrm{Prob}_{\vec{\mathcal{V}}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) = V(p)^{-1} \int_0^1 K_p(p') \, \mathrm{Prob}_{\vec{\mathcal{B}}_{p'}}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) \, dp'.$$

Our main theorems will show that, under certain conditions, $\mathcal{G}_p$ is very close to $\mathcal{V}_p$, $\mathcal{G}_k$ to $\mathcal{B}_k$, and $\mathcal{G}_{\boldsymbol{t}}$ to $\mathcal{B}_{\boldsymbol{t}}$. Similar relationships hold for the digraph models.

1.5. The main theorems

Consider positive integers m, n and a real variable $x \in (0, 1)$. (As explained in Section 1.2, these variables are actually functions of a background index ℓ.) For constants $a, \varepsilon > 0$, we say that (m, n, x) is (a, ε)-*acceptable* if

$$m, n \to \infty \text{ with } m = o(n^{1+\varepsilon}), n = o(m^{1+\varepsilon}), \text{ and}$$

(4)
$$\frac{(1 - 2x)^2}{4x(1 - x)}\left(1 + \frac{5m}{6n} + \frac{5n}{6m}\right) < a \log n.$$

Note that (4) implies $x(1 - x) = \Omega\big((\log n)^{-1}\big)$.

For $\varepsilon > 0$, a vector $(x_1, x_2, \ldots, x_N)$ will be called ε-*regular* if

$$x_i - \frac{1}{N} \sum_{j=1}^N x_j = O(N^{1/2+\varepsilon})$$

uniformly for $i = 1, \ldots, N$. We say that $(\boldsymbol{s}, \boldsymbol{t})$ is ε-*regular* if $\sum_{i=1}^m s_i = \sum_{j=1}^n t_j$ and $\boldsymbol{s}, \boldsymbol{t}$ are both ε-regular.

Finally, define $\lambda_m(\boldsymbol{t}) = (mn)^{-1} \sum_{j=1}^n t_j$. If $\sum_{i=1}^m s_i = \sum_{j=1}^n t_j$, the common value of $\lambda_n(\boldsymbol{s})$ and $\lambda_m(\boldsymbol{t})$ will be denoted by λ. Note that λ is the value in $[0, 1]$ that gives the density of a bipartite graph with degrees $(\boldsymbol{s}, \boldsymbol{t})$, relative to $K_{m,n}$. In the case of loop-free digraphs, $\lambda \in [0, 1 - 1/n]$.

We now state the theorems that are the main contribution of this paper. Their proofs will be given in Section 4, after some preliminary lemmas are proved in Section 3.

Theorem 1.1. *Let constants $a, b > 0$ satisfy $a + b < \frac{1}{2}$. Then there is a constant $\varepsilon = \varepsilon(a, b) > 0$ such that the following holds. Let $\mathcal{D}$ and $\mathcal{D}'$ be probability spaces on $I_{m,n}$ in one of the following cases.*

(a) (m, n, p) *is* (a, ε)-*acceptable and* $(\mathcal{D}, \mathcal{D}') = (\mathcal{G}_p, \mathcal{V}_p)$,
(b) $m = n$, (n, n, p) *is* (a, ε)-*acceptable and* $(\mathcal{D}, \mathcal{D}') = (\vec{\mathcal{G}}_p, \vec{\mathcal{V}}_p)$,
(c) $(m, n, k/mn)$ *is* (a, ε)-*acceptable and* $(\mathcal{D}, \mathcal{D}') = (\mathcal{G}_k, \mathcal{B}_k)$,
(d) $m = n$, $(n, n, k/n^2)$ *is* (a, ε)-*acceptable and* $(\mathcal{D}, \mathcal{D}') = (\vec{\mathcal{G}}_k, \vec{\mathcal{B}}_k)$.

Then there is an event $B = B(\mathcal{D}) \subseteq I_{m,n}$ *such that* $\mathrm{Prob}_{\mathcal{D}}(B) = \tilde{o}(1)$, *and uniformly for* $(\boldsymbol{s}, \boldsymbol{t}) \in I_{m,n} \setminus B$,

$$\mathrm{Prob}_{\mathcal{D}}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) = \big(1 + O(n^{-b})\big) \mathrm{Prob}_{\mathcal{D}'}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}).$$

Moreover, let $X : I_{m,n} \to \mathbb{R}$ *be a random variable and let* $E \subseteq I_{m,n}$ *be an event. Then,*

$$\mathrm{Prob}_{\mathcal{D}}(E) = \big(1 + O(n^{-b})\big) \mathrm{Prob}_{\mathcal{D}'}(E) + \tilde{o}(1),$$

$$\mathbb{E}_{\mathcal{D}}(X) = \mathbb{E}_{\mathcal{D}'}(X) + O(n^{-b}) \mathbb{E}_{\mathcal{D}'}(|X|) + \tilde{o}(1) \max_{(\boldsymbol{s},\boldsymbol{t}) \in I_{m,n}} |X|,$$

$$\mathrm{Var}_{\mathcal{D}}(X) = \big(1 + O(n^{-b})\big) \mathrm{Var}_{\mathcal{D}'}(X) + \tilde{o}(1) \max_{(\boldsymbol{s},\boldsymbol{t}) \in I_{m,n}} X^2.$$

Theorem 1.2. *Let* $E \subseteq I_{m,n}$ *be an event. Then, under the conditions of Theorem 1.1,*

if $\mathrm{Prob}_{\mathcal{B}_p}(E) \to 0$ *then* $\mathrm{Prob}_{\mathcal{G}_p}(E) = \tilde{o}(1) + o(1)\sqrt{\mathrm{Prob}_{\mathcal{B}_p}(E)}$, *and*

if $\mathrm{Prob}_{\mathcal{G}_p}(E) \to 0$ *then* $\mathrm{Prob}_{\mathcal{B}_p}(E) = \tilde{o}(1) + o(1)\sqrt{\mathrm{Prob}_{\mathcal{G}_p}(E)}$.

Similarly, for $m = n$,

if $\mathrm{Prob}_{\vec{\mathcal{B}}_p}(E) \to 0$ *then* $\mathrm{Prob}_{\vec{\mathcal{G}}_p}(E) = \tilde{o}(1) + o(1)\sqrt{\mathrm{Prob}_{\vec{\mathcal{B}}_p}(E)}$, *and*

if $\mathrm{Prob}_{\vec{\mathcal{G}}_p}(E) \to 0$ *then* $\mathrm{Prob}_{\vec{\mathcal{B}}_p}(E) = \tilde{o}(1) + o(1)\sqrt{\mathrm{Prob}_{\vec{\mathcal{G}}_p}(E)}$.

In particular, $\mathcal{G}_p$ *and* $\mathcal{B}_p$ *are contiguous; i.e.,* $\mathrm{Prob}_{\mathcal{G}_p}(E) \to 0$ *if and only if* $\mathrm{Prob}_{\mathcal{B}_p}(E) \to 0$, *and similarly for* $\vec{\mathcal{G}}_p$ *and* $\vec{\mathcal{B}}_p$.

Theorem 1.3. *Let constants* $a, b > 0$ *satisfy* $a + b < \frac{1}{2}$. *Then there is a constant* $\varepsilon = \varepsilon(a, b) > 0$ *such that the following holds whenever* $(m, n, \lambda_m(\boldsymbol{t}))$ *is* (a, ε)-*acceptable and* $\boldsymbol{t}$ *is* ε-*regular. Let* $\mathcal{D}$ *and* $\mathcal{D}'$ *be probability spaces on* I_n^m *in one of the following cases.*

(a) $(\mathcal{D}, \mathcal{D}') = (\mathcal{G}_{\boldsymbol{t}}, \mathcal{B}_{\boldsymbol{t}})$,

(b) $m = n$ and $(\mathcal{D}, \mathcal{D}') = (\vec{\mathcal{G}_t}, \vec{\mathcal{B}_t})$.

Then there is an event $B = B(\mathcal{D}) \subseteq I_n^m$ such that $\mathrm{Prob}_{\mathcal{D}}(B) = \tilde{o}(1)$, and uniformly for $s \in I_n^m \setminus B$,

$$\mathrm{Prob}_{\mathcal{D}}(\boldsymbol{S} = \boldsymbol{s}) = \big(1 + O(n^{-b})\big)\,\mathrm{Prob}_{\mathcal{D}'}(\boldsymbol{S} = \boldsymbol{s}).$$

Moreover, let $X : I_n^m \to \mathbb{R}$ be a random variable and let $E \subseteq I_n^m$ be an event. Then,

$$\mathrm{Prob}_{\mathcal{D}}(E) = \big(1 + O(n^{-b})\big)\,\mathrm{Prob}_{\mathcal{D}'}(E) + \tilde{o}(1),$$

$$\mathbb{E}_{\mathcal{D}}(X) = \mathbb{E}_{\mathcal{D}'}(X) + O(n^{-b})\,\mathbb{E}_{\mathcal{D}'}(|X|) + \tilde{o}(1)\max_{\boldsymbol{s} \in I_n^m}|X|,$$

$$\mathrm{Var}_{\mathcal{D}}(X) = \big(1 + O(n^{-b})\big)\,\mathrm{Var}_{\mathcal{D}'}(X) + \tilde{o}(1)\max_{\boldsymbol{s} \in I_n^m}X^2.$$

A weak corollary of these theorems is that each of the distribution pairs $(\mathcal{G}_p, \mathcal{V}_p)$, $(\vec{\mathcal{G}_p}, \vec{\mathcal{V}_p})$, $(\mathcal{G}_k, \mathcal{B}_k)$, $(\vec{\mathcal{G}_k}, \vec{\mathcal{B}_k})$, $(\mathcal{G}_t, \mathcal{B}_t)$ and $(\vec{\mathcal{G}_t}, \vec{\mathcal{B}_t})$ have total variation distance $O(n^{-b})$ under the stated conditions.

The proofs of the theorems will be presented in Sections 3 and 4. Meanwhile, we will give an example that illustrates how the theorems can be applied.

2. Some useful lemmas and an example

We first record a few elementary properties.

Lemma 2.1. If $\sum_{i=1}^m s_i = \sum_{j=1}^n t_j = k$ and $pqmn \to \infty$, then

$$\mathrm{Prob}_{\mathcal{B}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$$
$$= \big(2 + O((pqmn)^{-1})\big)\,p^{2k}q^{2mn-2k}\sqrt{\pi pqmn}\,\prod_{i=1}^m\binom{n}{s_i}\prod_{j=1}^n\binom{m}{t_j},$$

$$\mathrm{Prob}_{\vec{\mathcal{B}_p}}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$$
$$= \big(2 + O((pqn^2))^{-1})\big)\,p^{2k}q^{2n^2-2n-2k}$$
$$\times \sqrt{\pi pqn(n-1)}\,\prod_{i=1}^n\binom{n-1}{s_i}\prod_{j=1}^n\binom{n-1}{t_j},$$

uniformly over $\boldsymbol{s}, \boldsymbol{t}$.

Proof. In $\mathcal{I}_p$, both $\sum_{i=1}^m S_i$ and $\sum_{j=1}^n T_j$ have the distribution $\mathrm{Bin}(mn, p)$. Therefore

$$(5) \quad \mathrm{Prob}_{\mathcal{I}_p}\left(\sum_{i=1}^m S_i = \sum_{j=1}^n T_j\right) = \sum_{k=0}^{mn} \binom{mn}{k}^2 p^{2k} q^{2mn-2k}$$

$$= \frac{1}{2\sqrt{\pi pqmn}}\left(1 + O((pqmn)^{-1})\right).$$

For the last step we use that the central part of the sum is approximately normal and sum it with the Euler-Maclaurin formula, while the two tails of the sum are negligible in comparison. The first claim now follows from the formulas for $\mathrm{Prob}_{\mathcal{B}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$ and $\mathrm{Prob}_{\mathcal{I}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$. The second claim is proved in the same manner. □

Lemma 2.2. *If $pqmn \to \infty$, then*

$$V(p) = 1 - o\big(e^{-pqmn}\big).$$

Proof. $K_p(p')$ is a normal density with mean p and variance $pq/(2mn)$, so we just need to apply standard normal tail bounds to the definition of $V(p)$. □

The next lemma demonstrates how statistics of variables in $\mathcal{B}_p$ can be converted into statistics in $\mathcal{V}_p$. Note that X can be the indicator variable of an event, so the lemma applies to probabilities as well.

Lemma 2.3 ([29]). *Let $X : I_{m,n} \to \mathbb{R}$ be a random variable. Then*

$$\mathbb{E}_{\mathcal{V}_p}(X) = V(p)^{-1} \int_0^1 K_p(p')\, \mathbb{E}_{\mathcal{B}_{p'}}(X)\, dp',$$

$$\mathrm{Var}_{\mathcal{V}_p}(X) = V(p)^{-1} \int_0^1 K_p(p')\big(\mathrm{Var}_{\mathcal{B}_{p'}}(X) + (\mathbb{E}_{\mathcal{V}_p}(X) - \mathbb{E}_{\mathcal{B}_{p'}}(X))^2\big)\, dp'.$$

2.1. Vertices of low degree in random digraphs

We now provide an example of how Theorem 1.1(b) can be applied to random digraphs. Since this is only an illustration, we will not attempt to treat all values of the parameters or to obtain the best possible error terms.

Let G be a random loop-free digraph on n vertices and edge probability $p = \frac{1}{2}$. For convenience we will assume that n is even, though treatment of the odd case would be much the same. As usual, $S_1, \ldots, S_n$ are the out-degrees of the vertices, and $T_1, \ldots, T_n$ are the in-degrees. Let X, Y be random variables which count the vertices with out-degree at most $\frac{n}{2} - 1$, and the

vertices with in-degree at most $\frac{n}{2} - 1$, respectively. It is easy to see that each of X and Y has a distribution exactly $\mathrm{Bin}(n, \frac{1}{2})$, but that X and Y are not independent. Our aim will be to find their asymptotic joint distribution.

We will first calculate some properties of binomial distributions truncated at the centre. Application of model $\vec{\mathcal{V}}_{1/2}$ requires us to consider probabilities close to $\frac{1}{2}$.

Lemma 2.4. *The following hold when $\varepsilon > 0$ is sufficiently small. Let n be even and let $p = \frac{1}{2} + \delta$ where $\delta = O(n^{-1+\varepsilon})$. For $0 \leq k \leq \frac{n}{2} - 1$ define*

$$b(p, k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

and $P(\delta) = \sum_{k=0}^{n/2-1} b(p, k)$. Then

$$(6) \qquad P(\delta) = \frac{1}{2} - \delta \sqrt{\frac{2n}{\pi}} + O(n^{-1-\varepsilon}).$$

Define two random variables, Z_δ^- by truncating $\mathrm{Bin}(n-1, p)$ to $[0, \frac{1}{2}n - 1]$, and Z_δ^+ by truncating $\mathrm{Bin}(n-1, p)$ to $[\frac{1}{2}n, n-1]$. Then

$$(7)$$
$$\mathbb{E}(Z_\delta^-) = \frac{1}{2}n - \sqrt{\frac{n}{2\pi}} + O(n^{1/2-\varepsilon}), \qquad \mathbb{E}(Z_\delta^+) = \frac{1}{2}n + \sqrt{\frac{n}{2\pi}} + O(n^{1/2-\varepsilon}),$$

$$(8)$$
$$\mathrm{Var}(Z_\delta^-) = \frac{(\pi-2)n}{4\pi} + O(n^{1-\varepsilon}), \qquad \mathrm{Var}(Z_\delta^+) = \frac{(\pi-2)n}{4\pi} + O(n^{1-\varepsilon}).$$

Proof. Define $s_j = b(p, \frac{1}{2}n - 1 - j)$. From [28] we have for $j = O(n^{1/2+\varepsilon})$ that

$$s_0 = \left(1 - 2\delta - 2\delta^2 n - 1/(4n) + O(n^{-1-\varepsilon})\right) \sqrt{\frac{2}{\pi n}}, \text{ and}$$
$$\frac{s_j}{s_0} = \left(1 - 4j^4/(3n^3) + O(n^{-1-\varepsilon})\right) \exp\left(-4\delta j - 2j(j+1)/n\right).$$

By summing s_j/s_0 using the Euler-Maclaurin method, as in [28], we obtain the formula for $P(\delta)$. Similarly summing $j s_j/s_0$ and $j^2 s_j/s_0$, we obtain the formulas for $\mathbb{E}(Z_\delta^-)$ and $\mathrm{Var}(Z_\delta^-)$.

Finally, note that the truncations divide the range exactly in half, and so we have $1 - P(\delta) = P(-\delta)$, $\mathbb{E}(Z_\delta^-) + \mathbb{E}(Z_{-\delta}^+) = n - 1$ and $\mathrm{Var}(Z_\delta^-) = \mathrm{Var}(Z_{-\delta}^+)$. This proves the statistics for Z_δ^+. $\qquad\square$

Theorem 2.5. *Suppose n is even and x, y are integers with $x, y = O(n^{-1/2+\varepsilon})$ for sufficiently small $\varepsilon > 0$. Then*

$$\text{Prob}_{\vec{\mathcal{G}}_{1/2}}\big((X = \tfrac{1}{2}n + x) \wedge (Y = \tfrac{1}{2}n + y)\big)$$

$$= \frac{2 + o(1)}{n\sqrt{\pi^2 - 4}} \exp\left(-\frac{2\pi(\pi x^2 + \pi y^2 - 4xy)}{(\pi^2 - 4)n}\right).$$

Proof. Theorem 1.1(b) tells us to calculate the probability in $\vec{\mathcal{V}}_{1/2}$, for which we need the probability in $\vec{\mathcal{B}}_p$ when $p \approx \tfrac{1}{2}$. For integers x, y, define events

$$E(x, y) = \big\{(\boldsymbol{S}, \boldsymbol{T}) \mid X = \tfrac{1}{2}n + x \wedge Y = \tfrac{1}{2}n + y\big\}, \text{ and}$$
$$E_\Sigma = \big\{(\boldsymbol{S}, \boldsymbol{T}) \mid \textstyle\sum_{i=0}^n S_i = \sum_{j=0}^n T_j\big\}.$$

Recall that $\vec{\mathcal{B}}_p$ is $\vec{\mathcal{I}}_p$ conditioned on event E_Σ so, applying Bayes' rule twice,

$$(9) \qquad \text{Prob}_{\vec{\mathcal{B}}_p}(E(x, y)) = \text{Prob}_{\vec{\mathcal{I}}_p}(E(x, y)) \frac{\text{Prob}_{\vec{\mathcal{I}}_p}(E_\Sigma \mid E(x, y))}{\text{Prob}_{\vec{\mathcal{I}}_p}(E_\Sigma)}.$$

We have already computed $\text{Prob}_{\vec{\mathcal{I}}_p}(E_\Sigma)$ in (5); for $p = \tfrac{1}{2} + o(1)$ it is

$$(10) \qquad \text{Prob}_{\vec{\mathcal{I}}_p}(E_\Sigma) = \frac{1 + o(1)}{n\sqrt{\pi}}.$$

Now consider $\text{Prob}_{\vec{\mathcal{I}}_p}(E_\Sigma \mid E(x, y))$. Under this conditioning, symmetry implies that $\sum_i S_i$ has the same distribution as the sum of $\tfrac{1}{2}n + x$ copies of Z_δ^- and $\tfrac{1}{2}n - x$ copies of Z_δ^+, all of these being independent. A similar fact holds for $\sum_j T_j$, which is in addition independent of $\sum_i S_i$ since we are operating in $\vec{\mathcal{I}}_p$. Also recall that the binomial distribution and therefore its truncations and their convolutions are log-concave, so we know from [4] that $\Delta = \sum_i S_i - \sum_j T_j$, in $\vec{\mathcal{I}}_p$ conditioned on $E(x, y)$, satisfies a local central-limit theorem. Using Lemma 2.4, we calculate

$$\mathbb{E}_{\vec{\mathcal{I}}_p}(\Delta \mid E(x, y)) = (y - x)\sqrt{\frac{2n}{\pi}} + O(n^{1-\varepsilon})$$

$$\text{Var}_{\vec{\mathcal{I}}_p}(\Delta \mid E(x, y)) = \frac{(\pi - 2)n^2}{2\pi} + O(n^{2-\varepsilon}),$$

and so

$$(11) \qquad \text{Prob}_{\vec{\mathcal{I}}_p}(E_\Sigma \mid E(x, y)) = \frac{1 + o(1)}{n\sqrt{\pi - 2}} \exp\left(-\frac{2(x - y)^2}{(\pi - 2)n}\right).$$

Finally, consider $\text{Prob}_{\vec{\mathcal{I}}_p}(E(x,y))$. Since the $2n$ events $S_i \leq \frac{1}{2}n - 1, T_j \leq \frac{1}{2}n - 1$ are independent in $\vec{\mathcal{I}}_p$, X and Y are independent Binomial variables $\text{Bin}(n, P(\delta))$. Using (6) and the normal approximation for the binomial distribution, we have

$$
\text{Prob}_{\vec{\mathcal{I}}_p}(E(x,y))
$$
$$
= \frac{2 + o(1)}{\pi n} \exp\left(-\frac{2(x\sqrt{\pi} + \delta\sqrt{2n^3})^2}{\pi n} - \frac{2(y\sqrt{\pi} + \delta\sqrt{2n^3})^2}{\pi n} \right).
$$

Applying this to (9) together with (10) and (11), we find that

$$
\text{Prob}_{\vec{\mathcal{B}}_p}(E(x,y)) = \frac{2 + o(1)}{n\sqrt{\pi(\pi - 2)}}
$$
$$
\times \exp\left(-\frac{2(x\sqrt{\pi} + \delta\sqrt{2n^3})^2}{\pi n} - \frac{2(y\sqrt{\pi} + \delta\sqrt{2n^3})^2}{\pi n} - \frac{2(x - y)^2}{(\pi - 2)n} \right).
$$

Now we apply Lemma 2.3 to pass the result to $\vec{\mathcal{V}}_{1/2}$. Multiplying by $K_{1/2}(\frac{1}{2} + \delta)$ and integrating, we obtain the formula in the theorem, which holds for $\vec{\mathcal{G}}_{1/2}$ on account of Theorem 1.1(b).   □

A corollary of the theorem is that $X - Y$ and $X + Y$ have asymptotically independent normal distributions, apart from necessarily having the same parity.

Theorem 2.6. *Under the conditions of the theorem, let α, β be integers of the same parity such that $\alpha, \beta = O(n^{1/2+\varepsilon})$. Then*

$$
\text{Prob}_{\vec{\mathcal{G}}_{1/2}}\big((X + Y = n + \alpha) \wedge (X - Y = \beta) \big)
$$
$$
= \frac{2 + o(1)}{n\sqrt{\pi^2 - 4}} \exp\left(-\frac{\pi\alpha^2}{(\pi + 2)n} - \frac{\pi\beta^2}{(\pi - 2)n} \right).
$$

More complex information could also be obtained, such as the distributions of all the order statistics of the degrees, but the calculations would be considerably more intricate. See [30] for similar calculations for ordinary graphs.

3. Properties of likely degree sequences

To prove our theorems, our first task will be to investigate the bulk behaviour of our various probability spaces in order to identify some behaviour that

has probability $\tilde{o}(1)$. We will apply a few concentration inequalities, which we now give.

Theorem 3.1 ([26, Lemma 1.2])**.** *Let* $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)$ *be a family of independent random variables, with* X_i *taking values in a set* A_i *for each* i. *Suppose that for each* j *the function* $f : \prod_{i=1}^{N} A_i \to \mathbb{R}$ *satisfies* $|f(\boldsymbol{x}) - f(\boldsymbol{x}')| \leq c_j$ *whenever* $\boldsymbol{x}, \boldsymbol{x}' \in \prod_{i=1}^{N} A_i$ *differ only in the* j-*th component. Then, for any* z,

$$\mathrm{Prob}\big(\big|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))\big| \geq z\big) \leq 2\exp\big(-2z^2 / \textstyle\sum_{i=1}^{N} c_i^2\big).$$

Theorem 3.2. *Let* $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)$ *be a family of independent real random variables such that* $|X_i - \mathbb{E}(X_i)| \leq c_i$ *for each* i. *Define* $X = \sum_{i=1}^{N} X_i$. *Then, for any* z,

$$\mathrm{Prob}\big(|X - \mathbb{E}(X)| \geq z\big) \leq 2\exp\big(-\tfrac{1}{2}z^2 / \textstyle\sum_{i=1}^{N} c_i^2\big)$$

Another consequence of Theorem 3.1 is the following.

Theorem 3.3. *Let* $A_1, \ldots, A_N$ *be finite sets, and let* $a_1, \ldots, a_N$ *be integers such that* $0 \leq a_i \leq |A_i|$ *for each* i. *Let* $\binom{A_i}{a_i}$ *denote the uniform probability space of* a_i-*element subsets of* A_i. *Suppose that for each* j *the function* $f : \prod_{i=1}^{N} \binom{A_i}{a_i} \to \mathbb{R}$ *satisfies* $|f(\boldsymbol{x}) - f(\boldsymbol{x}')| \leq c_j$ *whenever* $\boldsymbol{x}, \boldsymbol{x}' \in \prod_{i=1}^{N} \binom{A_i}{a_i}$ *are the same except that their* j-*th components* x_j, x_j' *have* $|x_j \cap x_j'| = a_j - 1$ *(i.e., the* a_j-*element subsets* x_j, x_j' *are minimally different). If* $\boldsymbol{X} = (X_1, \ldots, X_N)$ *is a family of independent set-valued random variables with distributions* $\binom{A_1}{a_1}, \ldots, \binom{A_N}{a_N}$, *then for any* z,

$$\mathrm{Prob}\big(\big|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))\big| \geq z\big) \leq 2\exp\left(\frac{-2z^2}{\sum_{i=1}^{N} c_i^2 \min\{a_i, |A_i| - a_i\}}\right).$$

Proof. We start by reminding the reader of a classical algorithm called "reservoir sampling", attributed by Knuth to Alan G. Waterman [24, p. 144]. Let $Y_{a_i+1}^{(i)}, \ldots, Y_{|A_i|}^{(i)}$ be independent random variables, where $Y_j^{(i)}$ has the discrete uniform distribution on $\{1, 2, \ldots, j\}$. Now suppose $A_i = \{w_1, \ldots, w_{|A_i|}\}$. Execute the following algorithm:

For $j = 1, \ldots, a_i$ set $x_j := w_j$;
For $j = a_i + 1, \ldots, |A_i|$, if $Y_j^{(i)} \leq a_i$ then set $x_{Y_j^{(i)}} := w_j$.

Define $X_i = X_i(Y_{a_i+1}^{(i)}, \ldots, Y_{|A_i|}^{(i)})$ to be the value of $\{x_1, \ldots, x_{a_i}\}$ when the algorithm finishes. The *raison d'être* of the algorithm, which is easy to check,

is that X_i has distribution $\binom{A_i}{a_i}$; i.e., it is uniform. It is also easy to check that the maximum change to X_i resulting from a change in a single $Y_j^{(i)}$ is that one element is replaced by another.

Therefore, we can apply Theorem 3.1 if we consider $f(\boldsymbol{X})$ as a function of all the independent variables $\{Y_j^{(i)}\}$. If $a_i < |A_i|/2$, we can represent X_i by its complement; this justifies the term $\min\{a_i, |A_i| - a_i\}$ in the theorem statement. $\qquad\square$

We next apply these concentration inequalities to show that certain events are very likely in our probability spaces.

Theorem 3.4. *The following are true for sufficiently small $\varepsilon > 0$.*

(a) *Suppose that (m, n, p) and $(m, n, k/mn)$ are (a, ε)-acceptable. Then*

$$\mathrm{Prob}_{\mathcal{D}}\big((\boldsymbol{S}, \boldsymbol{T}) \text{ is } \varepsilon\text{-regular}\big) = 1 - \tilde{o}(1)$$

for $\mathcal{D}$ being any of $\mathcal{G}_p$, $\mathcal{G}_k$, $\mathcal{I}_p$, $\mathcal{B}_p$, $\mathcal{B}_k$, or $\mathcal{V}_p$. The same is true for $m = n$ when $\mathcal{D}$ is any of $\vec{\mathcal{G}}_p$, $\vec{\mathcal{G}}_k$, $\vec{\mathcal{I}}_p$, $\vec{\mathcal{B}}_p$, $\vec{\mathcal{B}}_k$, or $\vec{\mathcal{V}}_p$.

(b) *If $\boldsymbol{t} \in I_m^n$ is ε-regular, and $(m, n, \lambda_m(\boldsymbol{t}))$ is (a, ε)-acceptable, then*

$$\mathrm{Prob}_{\mathcal{D}}\big(\boldsymbol{S} \text{ is } \varepsilon\text{-regular}\big) = 1 - \tilde{o}(1).$$

for $\mathcal{D}$ being $\mathcal{G}_{\boldsymbol{t}}$ or $\mathcal{B}_{\boldsymbol{t}}$. The same is true for $m = n$ when $\mathcal{D}$ is either of $\vec{\mathcal{G}}_{\boldsymbol{t}}$ or $\vec{\mathcal{B}}_{\boldsymbol{t}}$.

Proof. By symmetry, we need only show that $\boldsymbol{S}$ is almost always ε-regular.

In the case that $\mathcal{D}$ is $\mathcal{G}_p$ or $\mathcal{I}_p$, each S_i has the binomial distribution $\mathrm{Bin}(n, p)$, and K has the distribution $\mathrm{Bin}(mn, p)$. Therefore, by Corollary 3.2,

$$\mathrm{Prob}_{\mathcal{D}}\big(|S_i - pn| \geq n^{1/2 + \varepsilon/2}\big) = \tilde{o}(1), \quad i = 1, \ldots, m,$$
(12)
$$\mathrm{Prob}_{\mathcal{D}}\big(|\Lambda - p| \geq n^{-1 + 2\varepsilon}\big) = \tilde{o}(1),$$

from which it follows that

$$\mathrm{Prob}_{\mathcal{D}}\big(\boldsymbol{S} \text{ is } \varepsilon\text{-regular}\big) = 1 - \tilde{o}(1).$$

The cases that $\mathcal{D}$ is $\mathcal{G}_k$, $\mathcal{B}_p$, or $\mathcal{B}_k$ follow, since these are the same as slices of $\mathcal{G}_p$ or $\mathcal{I}_p$ of size $n^{-O(1)}$, using $p = k/mn$. Also, the distribution of $\boldsymbol{S}$ in $\mathcal{B}_{\boldsymbol{t}}$ is the same as in $\mathcal{B}_k$ for $k = \sum_{j=1}^n t_j$, so that case follows too.

For $\mathcal{D} = \mathcal{G}_t$, note that each S_i is the sum of independent variables $X_1, \ldots, X_n$, where X_j is a Bernoulli random variable with mean t_j/m. The theorem thus follows using the same argument as we used for $\mathcal{G}_p$.

Finally consider $\mathcal{D} = \mathcal{V}_p$. Taking X to be the indicator of the event that $\boldsymbol{S}$ is not ε-regular, Lemmas 2.2–2.3 give

$$\mathbb{E}_{\mathcal{V}_p}(X) = O(1) \int_0^1 K_p(p') \, \mathbb{E}_{\mathcal{B}_{p'}}(X) \, dp'$$

$$= O(1) \left(\int_0^{p-n^{-1+\varepsilon}} + \int_{p-n^{-1+\varepsilon}}^{p+n^{-1+\varepsilon}} + \int_{p+n^{-1+\varepsilon}}^1 \right) K_p(p') \, \mathbb{E}_{\mathcal{B}_{p'}}(X) \, dp'.$$

The first and third integrals are $\tilde{o}(1)$ since the tails of $K_p(p')$ are small (recall that it is a normal density with mean p and variance $O((mn)^{-1})$), while the second integral is $\tilde{o}(1)$ by the present theorem in the case $\mathcal{D} = \mathcal{B}_{p'}$. (Note that if (m,n,p) is (a,ε)-acceptable, then all $p' \in [p - n^{-1+\varepsilon}, p + n^{-1+\varepsilon}]$ are (a',ε)-acceptable for slightly different a'.)

For the digraph models, the proofs are essentially the same. $\qquad\square$

The following concentration results will form a key part of the proof of Theorem 1.1.

Theorem 3.5. *The following are true for sufficiently small $\varepsilon > 0$.*

(a) Suppose that (m,n,p) and $(m,n,k/mn)$ are (a,ε)-acceptable. Then

$$(13) \quad \mathrm{Prob}_{\mathcal{D}} \Big(\sum_{i=1}^m (S_i - n\Lambda)^2$$

$$= \big(1 + O(n^{-1/2+2\varepsilon})\big)\Lambda(1-\Lambda)mn \Big) = 1 - \tilde{o}(1),$$

$$(14) \quad \mathrm{Prob}_{\mathcal{D}} \Big(\sum_{j=1}^n (T_j - m\Lambda)^2$$

$$= \big(1 + O(m^{-1/2+2\varepsilon})\big)\Lambda(1-\Lambda)mn \Big) = 1 - \tilde{o}(1),$$

when $\mathcal{D}$ is $\mathcal{G}_p$ or $\mathcal{G}_k$. When $m = n$, the same bounds hold when $\mathcal{D}$ is $\vec{\mathcal{G}}_p$ or $\vec{\mathcal{G}}_k$.

(b) If $\boldsymbol{t} \in I_m^n$ is ε-regular, and $(m,n,\lambda_m(\boldsymbol{t}))$ is (a,ε)-acceptable, then (13) holds when $\mathcal{D}$ is $\mathcal{G}_t$, and when $m = n$ and $\mathcal{D}$ is $\vec{\mathcal{G}}_t$.

(c) If $m = n$, (n,n,p) and $(n,n,k/n^2)$ are (a,ε)-acceptable, then

$$(15) \quad \mathrm{Prob}_{\mathcal{D}} \Big(\sum_{i=1}^n (S_i - n\Lambda)(T_i - n\Lambda)$$

$$= O(n^{-1/2+2\varepsilon})\Lambda(1-\Lambda)n^2\Big) = 1 - \tilde{o}(1)$$

when $\mathcal{D}$ is $\vec{\mathcal{G}}_p$ or $\vec{\mathcal{G}}_k$.

(d) If $m = n$, $(n, n, \lambda_n(\boldsymbol{t}))$ is (a, ε)-acceptable and $\boldsymbol{t} \in I_n^n$ is ε-regular, then (15) holds when $\boldsymbol{T} = \boldsymbol{t}$ and $\mathcal{D}$ is $\vec{\mathcal{G}}_{\boldsymbol{t}}$.

Proof. Write $R = \sum_{i=1}^m (S_i - n\Lambda)^2$. For $i = 1, \ldots, m$ and $j = 1, \ldots, n$, let X_{ij} be the indicator for an edge from u_i to v_j. Define $\Delta_{ii'jj'} = (X_{ij} - X_{i'j})(X_{ij'} - X_{i'j'})$. Then we have

$$\frac{1}{2m} \sum_{i,i'=1}^m \sum_{j,j'=1}^n \Delta_{ii'jj'} = \frac{1}{m} \sum_{i,i'=1}^m \sum_{j,j'=1}^n X_{ij}X_{ij'} - \frac{1}{m} \sum_{i,i'=1}^m \sum_{j,j'=1}^n X_{ij}X_{i'j'}$$

(16)
$$= \sum_{i=1}^m S_i^2 - \frac{1}{m}\Big(\sum_{i=1}^m S_i\Big)^2 = R.$$

When $\mathcal{D}$ is either $\mathcal{G}_p$ or $\mathcal{G}_{\boldsymbol{t}}$, X_{ij} is independent of $X_{i'j'}$ if $j \neq j'$, and $\mathbb{E}_\mathcal{D}(X_{ij})$ is independent of i. This shows that $\mathbb{E}_\mathcal{D}(\Delta_{ii'jj'}) = 0$ for $j \neq j'$, leaving us with

$$\mathbb{E}_\mathcal{D}(R) = \frac{1}{2m} \sum_{i,i'=1}^m \sum_{j=1}^n \mathrm{Prob}_D(X_{ij} \neq X_{i'j}).$$

This gives

$$\mathbb{E}_{\mathcal{G}_p}(R) = pq(m-1)n,$$

$$\mathbb{E}_{\mathcal{G}_t}(R) = \frac{1}{m} \sum_{j=1}^n t_j(m - t_j).$$

Now define $R^* = \sum_{i=1}^m \min\{(S_i - n\Lambda)^2, m^{1+2\varepsilon}\}$. If $\boldsymbol{S}$ is ε-regular and S_j is changed by 1 for some j, which changes Λ by $1/mn$, then $\min\{(S_i - n\Lambda)^2, m^{1+2\varepsilon}\}$ changes by $O(m^{1/2+\varepsilon})$ for $i = j$ and by $O(m^{-1/2+\varepsilon})$ for $i \neq j$. Consequently, R^* changes by $O(m^{1/2+\varepsilon})$. Applying Theorem 3.1, we find that

$$\mathrm{Prob}_\mathcal{D}\big(|R^* - \mathbb{E}_\mathcal{D}(R^*)| \geq \tfrac{1}{2}m^{1+\varepsilon}n^{1/2+\varepsilon/2}\big) = \tilde{o}(1)$$

for $\mathcal{D} = \mathcal{G}_p$. It also holds for $\mathcal{D} = \mathcal{G}_{\boldsymbol{t}}$, using Theorem 3.3 in the same way.

Now Theorem 3.4 shows that $\mathrm{Prob}_\mathcal{D}(R \neq R^*) = \tilde{o}(1)$, which implies that $\mathbb{E}_D(R^*) = \mathbb{E}_D(R) + \tilde{o}(1)$. Therefore we can argue

$$\mathrm{Prob}_\mathcal{D}\big(|R - \mathbb{E}_\mathcal{D}(R)| \geq m^{1+\varepsilon}n^{1/2+\varepsilon/2}\big)$$

$$\leq \operatorname{Prob}_{\mathcal{D}}(R \neq R^*) + \operatorname{Prob}_{\mathcal{D}}\left(|R^* - \mathbb{E}_{\mathcal{D}}(R)| \geq m^{1+\varepsilon} n^{1/2+\varepsilon/2}\right)$$
$$\leq \tilde{o}(1) + \operatorname{Prob}_{\mathcal{D}}\left(|R^* - \mathbb{E}_{\mathcal{D}}(R^*)| \geq m^{1+\varepsilon} n^{1/2+\varepsilon/2} + \tilde{o}(1)\right)$$
$$= \tilde{o}(1).$$

We also have that Λ is fixed at the value $\lambda_m(\boldsymbol{t}) = (mn)^{-1} \sum_{j=1}^{n} t_j$ in $\mathcal{G}_{\boldsymbol{t}}$ and that

$$\operatorname{Prob}_{\mathcal{G}_p}\left(|\Lambda - p| \geq n^{-1+2\varepsilon}\right) = \tilde{o}(1),$$

by (12). From these bounds, inequality (13) follows for $\mathcal{G}_p$ and $\mathcal{G}_{\boldsymbol{t}}$, and (14) follows for $\mathcal{G}_p$ by symmetry. By choosing $p = k/mn$ and noting that $\mathcal{G}_k$ is a slice of size $n^{-O(1)}$ of $\mathcal{G}_p$, the theorem is proved for $\mathcal{G}_k$ too.

For $\mathcal{D} = \vec{\mathcal{G}}_p, \vec{\mathcal{G}}_k, \vec{\mathcal{G}}_{\boldsymbol{t}}$, the proofs of (13) and (14) follow the same pattern. Since (16) still holds, we can note that $\mathbb{E}_{\vec{\mathcal{G}}_p}(\Delta_{ii'jj'}) = \mathbb{E}_{\mathcal{G}_p}(\Delta_{ii'jj'})$ and $\mathbb{E}_{\vec{\mathcal{G}}_t}(\Delta_{ii'jj'}) = \mathbb{E}_{\mathcal{G}_t}(\Delta_{ii'jj'})$ unless $\{j, j'\} \subseteq \{i, i'\}$, to infer that $\mathbb{E}_{\vec{\mathcal{G}}_p}(R) = \mathbb{E}_{\mathcal{G}_p}(R) + O(n)$ and $\mathbb{E}_{\vec{\mathcal{G}}_t}(R) = \mathbb{E}_{\mathcal{G}_t}(R) + O(n)$. This is enough to ensure that the rest of the proof continues in the same way. (For the record, $\mathbb{E}_{\vec{\mathcal{G}}_p}(R) = pq(n-1)^2$.)

We now prove part (d); take $\mathcal{D} = \vec{\mathcal{G}}_{\boldsymbol{t}}$, with $\boldsymbol{t}$ being ε-regular and $(n, n, \lambda_n(\boldsymbol{t}))$ being (a, ε)-acceptable. We have

$$\mathbb{E}_{\vec{\mathcal{G}}_t}(S_i) = \sum_{j \neq i} \frac{t_j}{n-1} = \frac{\lambda n^2}{n-1} - \frac{t_i}{n-1},$$

from which it follows that

$$\mathbb{E}_{\vec{\mathcal{G}}_t}\left(\sum_{i=1}^{n}(S_i - \lambda n)(t_i - \lambda n)\right) = -\frac{\sum_{j=1}^{n}(t_j - \lambda n)^2}{n-1} = O(n^{1+2\varepsilon}).$$

In the notation of Theorem 3.3 set $A_j = [n] \setminus \{j\}$ and $a_j = t_j$ for each $j \in [n]$. Then in the probability space $\boldsymbol{X}$, X_j is the set of indices of vertices incident with v_j in $\vec{\mathcal{G}}_{\boldsymbol{t}}$. Note $S_i = |\{j : u_i \in X_j\}|$ and two sets being minimally different in the j-th component corresponds to two graphs in which one of the t_j edges incident with vertex v_j is incident with different vertices in U. This means, as $\boldsymbol{t}$ is ε-regular, $c_j = O(n^{1/2+\varepsilon})$ for each j and we can apply Theorem 3.3 to conclude that (d) holds.

In the case of $\mathcal{D} = \vec{\mathcal{G}}_p$, Theorem 3.4 says that $\boldsymbol{T}$ is ε-regular with probability $1 - \tilde{o}(1)$, so (c) is true for $\vec{\mathcal{G}}_p$. Finally, $\vec{\mathcal{G}}_k$ is a substantial slice of $\vec{\mathcal{G}}_p$ if $p = k/n^2$, so (c) holds for $\vec{\mathcal{G}}_k$ too. $\qquad\square$

4. Proofs of the main theorems

In this section we will give the proofs of the theorems and corollary stated in Section 1.5. The bases for our analysis are the following enumerative results of Canfield, Greenhill and McKay [6, 10]. Also see Barvinok and Hartigan [2] for an overlapping result.

Theorem 4.1 ([6, 10]). *Let $a, b > 0$ be constants such that $a + b < \frac{1}{2}$. Then there is a constant $\varepsilon_0 = \varepsilon_0(a, b) > 0$ such that the following is true for any fixed ε with $0 < \varepsilon \leq \varepsilon_0$. If $(\boldsymbol{s}, \boldsymbol{t})$ is ε-regular, then*

$$G(\boldsymbol{s}, \boldsymbol{t}) = \binom{mn}{\lambda mn}^{-1} \prod_{i=1}^{m} \binom{n}{s_i} \prod_{j=1}^{n} \binom{m}{t_j}$$

$$\times \exp\left(-\frac{1}{2}\left(1 - \frac{\sum_{i=1}^{m}(s_i - \lambda n)^2}{\lambda(1-\lambda)mn}\right)\left(1 - \frac{\sum_{j=1}^{n}(t_j - \lambda m)^2}{\lambda(1-\lambda)mn}\right) + O(n^{-b})\right).$$

Moreover, if $m = n$, then

$$\vec{G}(\boldsymbol{s}, \boldsymbol{t}) = \binom{n^2 - n}{\lambda n^2}^{-1} \prod_{i=1}^{n} \binom{n-1}{s_i} \prod_{j=1}^{n} \binom{n-1}{t_j}$$

$$\times \exp\left(-\frac{1}{2}\left(1 - \frac{\sum_{i=1}^{n}(s_i - \lambda n)^2}{\lambda(1-\lambda)n^2}\right)\left(1 - \frac{\sum_{j=1}^{n}(t_j - \lambda n)^2}{\lambda(1-\lambda)n^2}\right)\right.$$

$$\left. - \frac{\sum_{i=1}^{n}(s_i - \lambda n)(t_i - \lambda n)}{\lambda(1-\lambda)n^2} + O(n^{-b})\right).$$

We first consider $\mathcal{G}_p$. Suppose that $a, b > 0$ are constants with $a + b < \frac{1}{2}$, and that (m, n, p) is (a, ε)-acceptable. According to Theorems 3.4–4.1 and equation (12), there is an event $B \subseteq I_{m,n}$ such that $\mathrm{Prob}_{\mathcal{G}_p}(B) = \tilde{o}(1)$ and, for $(\boldsymbol{s}, \boldsymbol{t}) \notin B$,

$$(17) \qquad |K - pmn| \leq mn^{2\varepsilon},$$

$$\mathrm{Prob}_{\mathcal{G}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) = p^k q^{mn-k} \exp\left(O(n^{-b})\right) \binom{mn}{k}^{-1} \prod_{i=1}^{m}\binom{n}{s_i}\prod_{j=1}^{n}\binom{m}{t_j},$$

$$(18) \qquad\qquad = p^{2k} q^{2mn-2k} \sqrt{2\pi pqmn} \prod_{i=1}^{m}\binom{n}{s_i}\prod_{j=1}^{n}\binom{m}{t_j}$$

$$\times \exp\left(\frac{(k - pmn)^2}{2pqmn} + O(n^{-b})\right)$$

for $\sum_i s_i = \sum_j t_j = k$, where the last step follows by Stirling's formula and, as always, we are assuming that ε is sufficiently small.

We wish to show that (18) closely matches the probability in $\mathcal{V}_p$. Define $P(p, \boldsymbol{s}, \boldsymbol{t}) = \mathrm{Prob}_{\mathcal{B}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$. By the definition of $\mathcal{V}_p$, we have

$$\mathrm{Prob}_{\mathcal{V}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) = V(p)^{-1} \int_0^1 K_p(p') P(p', \boldsymbol{s}, \boldsymbol{t}) \, dp'.$$

By Section 1.4 item 2, we have

$$(19) \quad \frac{P(p', \boldsymbol{s}, \boldsymbol{t})}{P(p, \boldsymbol{s}, \boldsymbol{t})} = \frac{\mathrm{Prob}_{\mathcal{I}_p}\left(\sum_{i=1}^m S_i = \sum_{j=1}^n T_j\right)}{\mathrm{Prob}_{\mathcal{I}_{p'}}\left(\sum_{i=1}^m S_i = \sum_{j=1}^n T_j\right)} \left(\frac{p'}{p}\right)^{2k} \left(\frac{1-p'}{1-p}\right)^{2mn-2k}.$$

We will divide the integral into three parts. Define $J_p = [p - n^{-1+3\varepsilon}, p + n^{-1+3\varepsilon}]$. By Lemma 2.1 and (17), for $p' \in J_p$ and $(\boldsymbol{s}, \boldsymbol{t}) \notin B$, we have

$$(20) \quad \frac{P(p', \boldsymbol{s}, \boldsymbol{t})}{P(p, \boldsymbol{s}, \boldsymbol{t})} = \exp\left(\frac{2(k-pmn)}{pq}(p'-p) - \frac{mn}{pq}(p'-p)^2 + O(n^{-1/2})\right),$$

which gives

$$\int_{J_p} K_p(p') P(p', \boldsymbol{s}, \boldsymbol{t}) \, dp' = 2^{-1/2} P(p, \boldsymbol{s}, \boldsymbol{t}) \exp\left(\frac{(k-pmn)^2}{2pqmn} + O(n^{-1/2})\right).$$

To bound the integral outside J_p, note that $(p'/p)^{2k}\big((1-p')/(1-p)\big)^{2mn-2k}$ is increasing for $p' \le p - n^{-1+3\varepsilon}$ and decreasing for $p \ge p + n^{-1+3\varepsilon}$. Also, since the mean square of a set of numbers is at least as large as the square of their mean, we can infer from (5) that $\mathrm{Prob}_{\mathcal{I}_{p'}}\left(\sum_{i=1}^m S_i = \sum_{j=1}^n T_j\right) \ge (mn+1)^{-1}$ for all p'. Since $mn \, \tilde{o}(1) = \tilde{o}(1)$, we obtain from (19) that

$$\int_{[0,1]\setminus J_p} K_p(p') P(p', \boldsymbol{s}, \boldsymbol{t}) \, dp' = \tilde{o}(1) P(p, \boldsymbol{s}, \boldsymbol{t}).$$

Recalling Lemma 2.2, we conclude that

$$V(p)^{-1} \int_0^1 K_p(p') P(p', \boldsymbol{s}, \boldsymbol{t}) \, dp'$$
$$= 2^{-1/2} P(p, \boldsymbol{s}, \boldsymbol{t}) \exp\left(\frac{(k-pmn)^2}{2pqmn} + O(n^{-1/2})\right),$$

which matches (18) when the value of $P(p, \boldsymbol{s}, \boldsymbol{t})$ given by Lemma 2.1 is substituted. This completes the proof of the first claim of Theorem 1.1(a).

The next two claims follow on summing the first claim over all $(\boldsymbol{s}, \boldsymbol{t})$. For the variance, we can apply the formula for the expectation to argue

$$
\begin{aligned}
\mathrm{Var}_{\mathcal{G}_p}(X) &= \min_{\mu \in \mathbb{R}} \mathbb{E}_{\mathcal{G}_p}(X - \mu)^2 \\
&= \min_{\mu \in \mathbb{R}} \big(\tilde{o}(1) \max_{(\boldsymbol{s}, \boldsymbol{t})} (X - \mu)^2 + (1 + O(n^{-b})) \, \mathbb{E}_{\mathcal{V}_p}(X - \mu)^2 \big) \\
&= \min_{\mu \in \mathbb{R}} \big(\tilde{o}(1) \max_{(\boldsymbol{s}, \boldsymbol{t})} X^2 + (1 + O(n^{-b})) \, \mathbb{E}_{\mathcal{V}_p}(X - \mu)^2 \big) \\
&= \tilde{o}(1) \max_{(\boldsymbol{s}, \boldsymbol{t})} X^2 + (1 + O(n^{-b})) \min_{\mu \in \mathbb{R}} \mathbb{E}_{\mathcal{V}_p}(X - \mu)^2 \\
&= \tilde{o}(1) \max_{(\boldsymbol{s}, \boldsymbol{t})} X^2 + (1 + O(n^{-b})) \, \mathrm{Var}_{\mathcal{V}_p}(X).
\end{aligned}
$$

For the third line we have used the obvious fact that the minimum in the first line occurs somewhere in the interval $[\min X, \max X]$.

The proof of Theorem 1.1(b) is the same. To prove Theorem 1.1(c), note that according to Theorems 3.4, 3.5 and 4.1, there is an event $B \subseteq I_{m,n}$ such that $\mathrm{Prob}_{\mathcal{G}_k}(B) = \tilde{o}(1)$ and, for $(\boldsymbol{s}, \boldsymbol{t}) \notin B$,

$$
\mathrm{Prob}_{\mathcal{G}_k}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t}) = \exp\big(O(n^{-b})\big) \binom{mn}{k}^{-2} \prod_{i=1}^m \binom{n}{s_i} \prod_{j=1}^n \binom{m}{t_j},
$$

which matches $\mathrm{Prob}_{\mathcal{B}_k}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})$ up to the error term. Similarly for Theorem 1.1(d).

Theorem 1.3 follows from a similar argument, on noting that the ε-regularity of $\boldsymbol{t}$ implies

$$
\sum_{j=1}^n (T_j - \lambda m)^2 \le n^{2+2\varepsilon} \le m^{4\varepsilon} \lambda (1 - \lambda) mn.
$$

Finally, we prove Corollary 1.2 for $\mathcal{D} = \mathcal{G}_p$, which is representative of the four cases. In view of Theorem 1.1, it will suffice to prove that

$$
\tag{21} \mathrm{Prob}_{\mathcal{V}_p}(E) \le \tilde{o}(1) + o(1) \sqrt{\mathrm{Prob}_{\mathcal{B}_p}(E)}
$$

if $\mathrm{Prob}_{\mathcal{B}_p}(E) \to 0$. Define

$$
y = \min\Big\{ n^\varepsilon, \sqrt{-\log(\mathrm{Prob}_{\mathcal{B}_p}(E)) - \tfrac{1}{2}\log(-\log(\mathrm{Prob}_{\mathcal{B}_p}(E)))} \Big\}
$$

and

$$\hat{E} = \big\{ (\boldsymbol{s}, \boldsymbol{t}) \in E : |K - pmn| \leq y\sqrt{pqmn} \,\big\}.$$

By a suitable normal approximation of the binomial distribution, such as [27, Thm. 3], $\mathrm{Prob}_{\mathcal{G}_p}(E \setminus \hat{E}) = O(e^{-y^2/2}/y)$, so by Theorem 1.1, $\mathrm{Prob}_{\mathcal{V}_p}(E \setminus \hat{E}) = \tilde{o}(1) + O(e^{-y^2/2}/y)$. Also note that

$$(22) \qquad \int_{|p'-p|>y\sqrt{pq/2mn}} K_p(p')\,dp' = O(e^{-y^2/2}/y).$$

Therefore, since $V(p)^{-1} = 1 + \tilde{o}(1)$ by Lemma 2.2,

$$\mathrm{Prob}_{\mathcal{V}_p}(E) = \tilde{o}(1) + O(e^{-y^2/2}/y) + \int_{|p'-p|\leq y\sqrt{pq/2mn}} K_p(p')\,\mathrm{Prob}_{\mathcal{B}_{p'}}(\hat{E})\,dp.$$

Now note that, by (20), for $|p' - p| \leq y\sqrt{pq/2mn}$ and $|k - pmn| \leq y\sqrt{pqmn}$ we have

$$\frac{\mathrm{Prob}_{\mathcal{B}_{p'}}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})}{\mathrm{Prob}_{\mathcal{B}_p}(\boldsymbol{S} = \boldsymbol{s} \wedge \boldsymbol{T} = \boldsymbol{t})}$$
$$\leq \exp\left(\frac{2(k - pmn)}{pq}(p' - p) - \frac{mn}{pq}(p' - p)^2 + O(n^{-1/2}) \right)$$
$$\leq (1 + O(n^{-1/2}))e^{y^2/2}$$

and so

$$\mathrm{Prob}_{\mathcal{B}_{p'}}(\hat{E}) \leq (1 + O(n^{-1/2}))e^{y^2/2}\,\mathrm{Prob}_{\mathcal{B}_p}(\hat{E}).$$

Since $\int K_p(p')\,dp < 1$, we have proved that

$$\mathrm{Prob}_{\mathcal{V}_p}(E) \leq \tilde{o}(1) + O(e^{-y^2/2}/y) + (1 + o(1))e^{y^2/2}\,\mathrm{Prob}_{\mathcal{B}_p}(E),$$

which gives (21) when the value of y is substituted. To prove the statement for the case $\mathcal{D} = \mathcal{B}_p$, redefine y and $\hat{E}$ by replacing each instance of $\mathcal{B}_p$ with $\mathcal{V}_p$. and then proceed in the same fashion (although in this case because of the direction of the inequality it is enough to note that the tails of the integral in (22) are positive; we do not need to show an upper bound as in the above proof for $\mathcal{D} = \mathcal{G}_p$).

5. Concluding remarks

A theorem similar to Theorem 4.1 holds also in the sparse domain. This was shown by Greenhill, McKay and Wang in the case that $(\max_i s_i)(\max_j t_j) =$

$o\big((\sum_i s_i)^{2/3}\big)$ [11]. That theorem can be used to develop a parallel theory of degree sequences in that domain, though some of the methods used in this paper must be replaced. However the lack of a precise enumeration in the gap between the sparse domain and the dense domain of Theorem 4.1 currently thwarts a theory which spans both the sparse and dense domains.

References

[1] I. Adler and S. M. Ross (2001). The coupon subset collection problem. *J. Appl. Probab.* **38**(3), 737–746. MR1860210

[2] A. Barvinok and J. A. Hartigan (2010). The number of graphs and a random graph with a given degree sequence. *Adv. Math.* **224**(1), 316–339. MR2600999

[3] A. D. Barbour and L. Holst (1989). Some applications of the Stein-Chen method for proving Poisson convergence. *Adv. Appl. Probab.* **21**(1), 74–90. MR0980737

[4] E. A. Bender (1973). Central and local limit theorems applied to asymptotic enumeration. *J. Combin. Theory Ser. A* **15**, 91–111. MR0375433

[5] J. Bukor (1993). Almost every bipartite graph has not two vertices of minimum degree. *Math. Slovaca* **43**, 113–117. MR1274596

[6] E. R. Canfield, C. Greenhill, and B. D. McKay (2008). Asymptotic enumeration of dense 0-1 matrices with specified line sums. *J. Combin. Theory Ser. A* **115**, 32–66. MR2378856

[7] A. Chen, F. Zhang, and H. Hao Li (2011). The maximum and minimum degrees of random bipartite multigraphs. *Acta Math. Sci. Ser. B Engl. Ed.* **31**(3), 1155–1166. MR2830553

[8] P. J. Eicker, M. M. Siddiqui, and P. W. Mielke Jr. (1972). A matrix occupancy problem. *Ann. Math. Stat.* **43**(3), 988–996. MR0301774

[9] A. P. Godbole, B. Lamorte, and J. K. Sklar (1999). If rooks could kill: vertex degrees in random bipartite graphs. In: Y. Alavi, D. R. Lick, and A. Schwenk (eds.) *Combinatorics, Graph theory, and Algorithms, II*, pp. 445–450. New Issues Press, Kalamazoo. MR1985075

[10] C. Greenhill and B. D. McKay (2009). Random dense bipartite graphs and directed graphs with specified degrees. *Random Struct. Algorithms* **35**, 222–249. MR2544007

[11] C. Greenhill, B. D. McKay, and X. Wang (2006). Asymptotic enumeration of sparse 0-1 matrices with irregular row and column sums. *J. Combin. Theory Ser. A* **113**, 291–324. MR2199276

[12] B. Harris (1989). Poisson limits for generalised random allocation problems. *Stat. Probab. Lett.* **8**(2), 123–127. MR1017878

[13] B. Harris, M. Marden, and C. J. Park (1987). The distribution of the number of empty cells in a generalized random allocation scheme. In: M. Karoński and Z. Palka (eds.) *Random Graphs'85*, *Ann. Discret. Math.* **33**, pp. 77–90, Amsterdam. MR0930485

[14] L. Holst (1980). On matrix occupancy, committee, and capture-recapture problems. *Scand. J. Stat.* **7**(3), 139–146. MR0590406

[15] J. Jaworski and M. Karoński (1993). On the connectivity of graphs generated by a sum of random mappings. *J. Graph Theory* **17**(2), 135–150. MR1217389

[16] A. E. Khakimullin (2005). Distribution of extreme terms in an ordered series in a scheme for allocating particles by groups of random size. *Diskret. Mat.*, **17**(3), 28–44. MR2195649

[17] E. R. Khakimullin (1981). Asymptotic behaviour of the maximum occupancy in an equiprobable scheme of allocation of particles by batches (in Russian). *Mat. Zametki* **30**, 277–289. MR0632790

[18] E. R. Khakimullin and N. Y. Enatskaya (1993). Asymptotic behaviour of the waiting time for a scheme of particle allocation. Probab. Methods in *Discret. Math., Proc. Third Intern. Petrozavodsk Conf. VSP/TVP*, Utrecht/Moscow, 155–185. MR1383134

[19] E. R. Khakimullin and N. Y. Enatskaya (1997). Limit theorems for the number of empty cells. *Discrete Math. Appl.* **7**(2), 209–219. MR1468078

[20] E. R. Khakimullin and N. Y. Enatskaya (2002). Asymptotic behaviour of the waiting time for schemes of particle allocation. Probab. Methods in *Discret. Math., Proc. Fifth Intern. Petrozavodsk Conf. VSP/TVP*, Utrecht/Moscow, 197–209.

[21] J. E. Kobza, S. H. Jacobson, and D. E. Vaughan (2007). A survey of the coupon collector's problem with random sample sizes. *Methodol. Comput. Appl. Probab.* **9**(4), 573–584. MR2404744

[22] V. Kolchin, B. Sevastyanov, and V. Chistyakov (1978). *Random Allocations*. John Wiley, New York. MR0471016

[23] W. Kordecki (1996). Poisson convergence of numbers of vertices of a given degree in random graphs. *Discuss. Math. Graph Theory* **16**, 157–172. MR1446354

[24] D. E. Knuth (1997). *The Art of Computer Programming.* Vol. 3, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 3rd edition.

[25] H. M. Mahmoud (2010). Gaussian phases in generalised coupon collection. *Adv. in Appl. Probab.* **42**(4), 994–1012. MR2796674

[26] C. McDiarmid (1989). On the method of bounded differences. In: *Surveys in Combinatorics.* Lond. Math. Soc. Lect. Notes Ser., **141**, 148–188, Cambridge: Cambridge University Press. MR1036755

[27] B. D. McKay (1989). On Littlewood's estimate for the binomial distribution. *Adv. in Appl. Probab.* **21**, 475–478. MR0997736

[28] B. D. McKay, I. M. Wanless, and N. C. Wormald (2002). Asymptotic enumeration of graphs with a given bound on the maximum degree. *Combin. Prob. Comput.* **11**, 373–392. MR1918723

[29] B. D. McKay and N. C. Wormald (1997). The degree sequence of a random graph. I. The models. *Random Struct. Algorithms* **11**, 97–117. MR1610253

[30] B. D. McKay and N. C. Wormald (2015). The degree sequence of a random graph. II. Order statistics and degree counts, preprint.

[31] S. S. Mirakhmedov and S. M. Mirakhmedov (2010). On asymptotic expansion in the random allocation of particles by sets. *J. Theoret. Probab.* **23**, 904–919. MR2679962

[32] S. M. Mitwalli (2002). An occupancy problem with group drawings of different sizes. *Math. Slovaca* **52**(2), 235–24. MR1935121

[33] Z. Palka (1983). The distribution of degrees in random graphs, In: *Graph Theory* (Łagów, Poland, 1981). Lect. Notes Math, **1018** 161–169. Springer, Berlin. MR0730646

[34] Z. Palka (1984). On the degrees of vertices in a bichromatic random graph. *Period. Math. Hungar.* **15**, 121–126. MR0744973

[35] Z. Palka (1986). Rulers and slaves in a random social group. *Graphs Combin.* **2**, 165–172. MR0932123

[36] Z. Palka (1987). Extreme degrees in random graphs. *J. Graph Theory* **11**, 121–134. MR0889344

[37] Z. Palka and M. Sperling (2006). In-degree sequence in a general model of a random digraph. *Discuss. Math. Graph Theory* **26**, 193–207. MR2330321

[38] C. J. Park (1981). On the distribution of the number of unobserved elements when rn-samples of size n are drawn from a finite population. *Comm. Statist. Theory Methods* **10**, 371–383. MR0612403

[39] V. G. Mikhailov (1977). A Poisson limit theorem in a scheme for group allocations of particles. *Theory Probab. Appl.* **22**, 152–156.

[40] V. G. Mikhailov (1977). An estimate for the rate of convergence to the Poisson distribution in group allocation of particles. *Theory Probab. Appl.* **22**, 554–562.

[41] T. M. Sellke (1995). How many iid samples does it take to see all the balls in a box? *Ann. Appl. Probab.* **5**(1), 294–309. MR1325054

[42] R. T. Smythe (2011). Generalized coupon collection: the superlinear case. *J. Appl. Probab.* **48**(1), 189–199. MR2809895

[43] W. Stadje (1990). The collector's problem with group drawings. *Adv. in Appl. Probab.* **22**(4), 866–882. MR1082197

[44] V. A. Vatutin and V. G. Mikhailov (1982). Limit theorems for the number of empty cells in the equiprobable scheme of allocating groups of particles. *Theory Probab. Appl.* **27**, 734–743. MR0681461

Brendan D. McKay
Research School of Computer Science
Australian National University
Canberra ACT 2601
Australia
E-mail address: Brendan.McKay@anu.edu.au

Fiona Skerman
Department of Statistics
University of Oxford
Oxford OX1 3TG
UK
E-mail address: skerman@stats.ox.ac.uk