Asymptotic enumeration of digraphs and bipartite graphs by degree sequence

Anita Liebenau 1 | Nick Wormald 2

1 School of Mathematics and Statistics, UNSW Sydney, Sydney, New South Wales, Australia
2 School of Mathematics, Monash University, Clayton, Victoria, Australia

Correspondence
Anita Liebenau, School of Mathematics and Statistics, UNSW Sydney, Sydney, NSW 2052, Australia.
Email: a.liebenau@unsw.edu.au

Funding information
Australian Research Council, Grant/Award Number: DP180103684

Abstract
We provide asymptotic formulae for the numbers of bipartite graphs with given degree sequence, and of loopless digraphs with given in- and out-degree sequences, for a wide range of parameters including the biregular case. In particular, for bipartite graphs with part sizes that are not highly unequal, our results cover an existing gap in known results between the sparse and dense cases that were proved by Greenhill, McKay, and Wang in 2006 and by Canfield, Greenhill, and McKay in 2008, respectively. For the range of parameters which our results cover, they imply that the degree sequence of a random bipartite graph with \( m \) edges and given part sizes is accurately modeled by a sequence of independent binomial random variables, conditional upon the sum of variables in each part being equal to \( m \). This extends the known behavior in the sparse and dense cases to essentially the full spectrum of possible densities. A similar model also holds for loopless digraphs.

KEYWORDS
asymptotic enumeration, bipartite graphs, degree sequence, digraphs, random graphs

1 | INTRODUCTION

Enumeration of discrete structures with local constraints has attracted the interest of many researchers and has applications in various areas such as coding theory, statistics, and neurostatistical analysis. Exact formulae are often hard to derive or infeasible to compute. Asymptotic formulae are therefore...
sought and often provide sufficient information for the aforementioned applications. In this article we find such formulae for bipartite graphs with given degree sequence, or loopless digraphs with given in- and out-degree sequences. Our results imply that the degree sequence of a random digraph or bipartite graph with \( m \) edges is close to a sequence of independent binomial random variables, conditional upon the sum of degrees in each part being equal to \( m \).

We frame all our arguments in terms of bipartite graphs: as noted below, digraphs are equivalent to “balanced” bipartite graphs. Thus, if loops are not forbidden, the digraph enumeration problem is the same as the bipartite one. The loopless case for digraphs is equivalent to bipartite graphs with a forbidden perfect matching. Our results on counting bipartite graphs with a given degree sequence imply equivalent results on counting 0–1 matrices with given row and column sums. Similarly, counting (loopless) digraphs is equivalent to counting square 0–1 matrices with given row and column sums where the entries on the diagonal are required to be 0.

Our results are obtained via the method of degree switchings and contraction mappings recently introduced by the authors in [14] to count the number of “nearly” regular graphs of a given degree sequence for medium-range densities, and a wider range of degree sequences for low densities. The basic structure of the argument is very similar in the present case, but it needs significant modifications to account for the fact that we are dealing with bipartite graphs and certain edges are not admissible.

### 1.1 Enumeration results

The formulae in [14] are stated in terms of a relationship between the degree sequence of the Erdős-Rényi random graph and a sequence of independent binomial random variables. We shall do the same here for appropriate bipartite random graphs and suitable independent binomials. We first introduce appropriate graph theoretic notation. Let \( \ell, n \) be integers and let \( S = [\ell] \) and \( T = [n+\ell] \setminus [\ell] \). We use \( S \) and \( T \) as the two parts of the vertex set of a bipartite graph \( G \), that is, a graph \( G \) with bipartition \((S,T)\). Such a graph is said to have degree sequence \((s,t)\) if vertex \( a \) has degree \( s_a \) for all \( a \in S \), and \( v \) has degree \( t_v \) for all \( v \in T \). (Our convention is to denote elements of \( S \) by \( \{a,b,\ldots\} \) and elements of \( T \) by \( \{v,w,\ldots\} \).) We let \( D(G) \) denote the degree sequence of \( G \). When \( \ell = n \), we use the fact that a digraph on \( n \) vertices with out-degree sequence \( s \) and in-degree sequence \( t \) corresponds to a bipartite graph with degree sequence \((s,t)\), the equivalence obtained by directing all edges from \( S \) to \( T \). For use in the digraph case, if \( a \in S \) we define \( \tilde{a} = a + \ell \in T \), and for \( v \in T \) we define \( \tilde{v} = v - \ell \in S \). The digraph contains a loop if and only if the bipartite graph has an edge \( a\tilde{a} \) joining \( a \) to \( \tilde{a} \).

The following probability spaces play an important role in this article. Let \( \tilde{G}(\ell,n,m) \) denote the bipartite graph chosen uniformly at random among all bipartite graphs with bipartition \((S,T)\) and with \( m \) edges. In the case when \( \ell = n \), conditioning on the event that none of those \( m \) edges is of the form \( a\tilde{a} \) yields a model of random directed graphs without loops which we call \( \tilde{G}(n,m) \). We define \( D(\tilde{G}(\ell,n,m)) \) and \( D(\tilde{G}(n,m)) \) to be the corresponding probability spaces of degree sequences of \( G(\ell,n,m) \) or of \( \tilde{G}(n,m) \), respectively. Let \( B_p(\ell,n) \) be the probability space of vectors of length \( \ell + n \) where the first \( \ell \) elements are distributed as \( \text{Bin}(n,p) \) and the next \( n \) are distributed as \( \text{Bin}(\ell,p) \), all components of the vectors being distributed independently. Furthermore, let \( B_m(\ell,n) \) be the restriction of \( B_p(\ell,n) \) to the event \( \Sigma_1 = \Sigma_2 = m \), where \( \Sigma_1 \) is the sum of the first \( \ell \) elements of the vector, and \( \Sigma_2 \) the sum of the other \( n \) elements. Similarly, define \( \tilde{B}_p(n) \) to be the probability space of random vectors of length \( 2n \), every component being independently distributed as \( \text{Bin}(n-1,p) \). Finally, let \( \tilde{B}_m(n) \) be the restriction of \( \tilde{B}_p(n) \) to the event \( \Sigma_1 = \Sigma_2 = m \), where \( \Sigma_i \) is defined as above with \( \ell = n \). Note that if \( \sum s_a = \sum t_v = m \), then...
\[ \mathbb{P}_{B_m(\ell, n)}(s, t) = \left( \frac{\ell n}{m} \right)^{-2} \prod_{a \in S} \binom{n}{s_a} \prod_{v \in T} \binom{\ell}{t_v} \] and
\[ \mathbb{P}_{B_m(n)}(s, t) = \left( \frac{n(n-1)}{m} \right)^{-2} \prod_{a \in S} \binom{n-1}{s_a} \prod_{v \in T} \binom{n-1}{t_v}, \]

which we note are both independent of \( p \).

Our main result is Theorem 1.1, which states essentially that for certain sequences \( d \), the probability \( \mathbb{P}_{D(G)}(d) \) is asymptotically equal to \( \mathbb{P}_B(d)\bar{H}(d) \), where \( G = G(\ell, n, m) \) and \( B = B_m(\ell, n) \) in the bipartite case, \( G = \tilde{G}(n, m) \) and \( \bar{B} = \bar{B}_m(n) \) in the digraph case, and where \( \bar{H} \) is a correction factor which we define next. The range of degree sequences to which our result applies covers the typical range of vertex degrees in the random graphs \( G(\ell, n, m) \) and \( \tilde{G}(n, m) \), for a wide range of values of \( m \). This range includes all but very sparse cases (or their complements) and cases where the density of the graph is between a certain constant \( c \) and \( 1-c \). For asymptotics in this article, we take \( n \to \infty \); the restrictions on \( \ell \) will also ensure that \( \ell \to \infty \). (See Section 2 for definitions of asymptotic and other notation.)

With \( S \) and \( T \) as above, let \( d \) be a sequence of length \( N = \ell + n \). We set \( M_1 = M_1(d) = \sum_{i=1}^N d_i \) and use \( s = s(d) \) and \( t = t(d) \) to denote the vectors consisting of the first \( \ell \), and of the last \( n \), entries of \( d \) respectively. Thus, \( d = (s, t) \). We also let \( s = s(d) \) and \( t = t(d) \) denote the average of the components of \( s \), and of \( t \), respectively. Then we set
\[ \sigma^2(s) = \frac{1}{\ell} \sum_{a \in S} (s_a - s)^2, \quad \sigma^2(t) = \frac{1}{n} \sum_{v \in T} (t_v - t)^2, \]
and, in the digraph case,
\[ \sigma(s, t) = \frac{1}{n} \sum_{a \in S} (s_a - s)(t_a - t). \]

We unify our analysis of the two cases, bipartite graphs and digraphs, by introducing the indicator variable \( \delta^{\bar{d}} \) which is 1 in the digraph case (in which case \( \ell = n \) is assumed) and 0 in the bipartite case (in which case terms containing \( \delta^{\bar{d}} \) as a factor may be undefined). This significantly simplifies notation and permits us to emphasize the similarities between the two cases. Define \( \mu = \mu(d) = M_1(d)/(2n(\ell - \delta^{\bar{d}})) \). (This is the edge density of a bipartite graph or a digraph with degree sequence \( d \).) Then we set
\[ \bar{H}(d) = \exp \left( -\frac{1}{2} \left( 1 - \frac{\sigma^2(s)}{s(1-\mu)} \right) \left( 1 - \frac{\sigma^2(t)}{t(1-\mu)} \right) - \frac{\delta^{\bar{d}}\sigma(s, t)}{s(1-\mu)} \right), \]

for a sequence \( d \) of length \( \ell + n \), where \( \mu = \mu(d) \), \( s = s(d) \), \( t = t(d) \). We can now state our main result.

**Theorem 1.1.** For a sufficiently small constant \( \mu_0 \), the following holds. Let \( 0 < \kappa < 1/10 \). Let \( \ell \), \( n \) and \( m \) be integers that satisfy
\[ \bar{\mu} := \frac{m}{n\ell} < \mu_0, \]
\[ \frac{\ell}{n^{3/2-5\kappa}} + \frac{n}{\ell^{3/2-5\kappa}} = o(\bar{\mu}^{1/2-5\kappa}), \]

(3)
Let $\mathfrak{D}$ be the set of sequences $\mathbf{d} = (s, t)$ with $s$ and $t$ of lengths $\ell$ and $n$ respectively, satisfying $M_1(s) = M_1(t) = m$, $|s_a - |s| \leq s^{1/2+\kappa}$ and $|t_v - |t| \leq t^{1/2+\kappa}$ for all $a \in S$ and all $v \in T$, where $s = m/\ell$ and $t = m/n$. Either set $\mathcal{G} = \mathcal{G}(\ell, n, m)$ and $B = B_m(\ell, n)$ (the bipartite case), or set $\mathcal{G} = \mathcal{G}(n, m)$ and $B = \mathcal{B}_m(n)$ and restrict to $\ell = n$ (the digraph case). Then uniformly for all $\mathbf{d} \in \mathfrak{D}$,

$$
P_{D(G)}(\mathbf{d}) = P_B(\mathbf{d})\mathcal{H}(\mathbf{d}) \left( 1 + O \left( \frac{\log \ell}{\ell} + \frac{\log n}{n} + \frac{\ell \overline{\mu}^{5/2-1/2}}{n^{3/2-5\kappa}} + \frac{n \overline{\mu}^{5/2-1/2}}{\ell^{3/2-5\kappa}} \right) \right).$$

(5)

We prove Theorem 1.1 in Section 3. Recall that in this article, asymptotic statements refer to $n \to \infty$. Condition (4), however, together with the trivial upper bound $m \leq n\ell$ implies that $\ell \to \infty$ as well.

In order to explain the range of the parameters in the theorem it is instructive to consider the relatively balanced case when $\ell = \Theta(n)$. Here, the condition (3), which can be regarded as a restriction on the imbalance between $\ell$ and $n$, defaults to $n = o(m)$ which is in particular implied by (4). This latter assumption can be viewed as a quite mild lower bound on the density, given $n$ and $\ell$, and so the theorem applies to relatively balanced $n$ and $\ell$ as long as $n^{1+\epsilon} < \overline{\mu} < \mu_0$ for some $\epsilon > 0$. This easily bridges the gap between existing results for the relatively balanced sparse and dense cases, as we discuss in detail below. Note that at the densest range of our result, where $\overline{\mu}$ (which is our convenient approximation to the density) is constant, our theorem applies as long as $\ell = o(n^{3/2-5\kappa})$ and $n = o(\ell^{3/2-5\kappa})$. Moreover, for (3) to hold, it is required that $\ell = o(\overline{\mu}^{1/2}) = o(n^{1/2})$, and similarly with $n$ and $\ell$ swapped, and hence $\log \ell = \Theta(\log n)$. Finally, note that if the parameter $\kappa$, a measure of the maximum relative degree spread, takes any value greater than 0, then, as mentioned above, this captures the range of degrees in a typical graph in $\mathcal{G}(\ell, n, m)$, a fact that we use in Theorem 1.3 below.

Remark 1.2. In view of (1) and the fact that $|\mathcal{G}(\ell, n, m)| = \binom{\ell n}{m}$, the formula in Theorem 1.1 is equivalent to the assertion that the number of bipartite graphs with degree sequence $(s, t)$ is

$$
\binom{\ell n}{m}^{-1} \prod_{a \in S} \binom{n}{s_a} \prod_{v \in T} \binom{\ell}{t_v} \exp \left( -\frac{1}{2} \left( 1 - \frac{\sigma^2(s)}{\mu(1-\mu)n} \right) \left( 1 - \frac{\sigma^2(t)}{\mu(1-\mu)\ell} \right) + O(\xi) \right),
$$

where $\xi$ is the error term from (5). Similarly, (1) and the fact that $|\mathcal{G}(n, m)| = \binom{n(n-1)}{m}$ gives an asymptotic formula for the number of directed graphs with given degree sequences of in- and out-degrees, and $m$ edges. Note that in the directed graph case, the theorem applies provided that $m < \mu_0 n^2$, $n \log K n = o(m)$ for all fixed $K > 0$, and $|d_i - d| \leq d^{1/2+\kappa}$ for every in- or out-degree $d_i$, where $d = m/n$.

We remark also that a simplified version of our method can also be used to obtain formulae in the sparse case (see the first preprint version of this article [15] and also [14]), but it appears that no sets of parameters not already covered by previous results would be reached without lengthy computations.

There have been many contributions to this topic in the past. Finding (asymptotic) formulae for the number of bipartite graphs with a given degree sequence goes back to Read’s thesis [24] and gained wider interest since the 1970’s, including [3–5, 7, 8, 17, 19, 21, 23, 25]. In particular, the sparse case is best covered by Greenhill, McKay, and Wang [11], who proved an asymptotic formula for the number of bipartite graphs of a given sequence $(s, t)$, provided that $m := M_1(s) = M_1(t)$ and $\Delta(s) \Delta(t) = \Theta \left( m^{2/3} \right)$, where $\Delta(\mathbf{b}) = \max_i b_i$ for any vector $\mathbf{b}$. This is supplemented by a formula for the asymptotic
number of relatively balanced, dense bipartite graphs with specified degree sequences by Canfield, Greenhill, and McKay [6]. Their main result, amongst other conditions, requires $\ell$ and $n$ to differ by at most a $\log n$ factor, and also $\min\{\tilde{\mu}, 1 - \tilde{\mu}\} = \Omega(1/\log n)$ where $\tilde{\mu} = m/n\ell$ as in Theorem 1.1. These values of $n$ and $\ell$ are much more balanced than we described above for Theorem 1.1 in the dense range where it applies. Their result also imposes a bound on the range of degrees that essentially requires the positive quantity $\kappa$ in our Theorem 1.1 to be sufficiently small. Barvinok and Hartigan [2] later derived an asymptotic formula valid for a wider range of degree sequences, but not as wide as for the much cruder estimates given by Barvinok [1]. Isaev and McKay [12] introduced a general technique using complex martingales and applied it to obtain formulae for the same range as in [2], additionally permitting a certain set of edges to be forbidden, and another set to be forcibly included, in the graphs being counted (subject to certain conditions on those edge sets). All the formulae explicitly given in [2] and [12] require the density, and the imbalance between $n$ and $\ell$, to be constrained at least as much as in [6].

Returning to the relatively balanced parameters, where $\ell = \Theta(n)$, if in addition we require the spread of degrees to be relatively small compared to average degrees, the above are the best existing results. They cover $\tilde{\mu} = o(1/\sqrt{n})$ and $\tilde{\mu} = \Omega(1/\log n)$. As mentioned above Theorem 1.1 is applicable for all densities in between.

Actually, in [6] it was found that the formulae for the sparse and the dense case can be unified to produce the formula in Theorem 1.1, which was implicitly conjectured in [6] to hold for the cases in between. This conjecture is essentially verified by Theorem 1.1 for a wide range of parameters $s$ and $t$.

It is interesting to compare the various results in the special case of “biregular” sequences, in which all vertices on one side of the bipartition have degree $s$, say, and all vertices on the other side have degree $t$. This case has received special attention in the literature. So let $s$ denote the constant vector of length $\ell$ in which every entry is $s$, and $t$ denote the constant vector of length $n$ in which every entry is $t$. In 1977, Good and Crook [9] suggested that the number of bipartite graphs with degree sequence $(s, t)$ is roughly
\[
\binom{n}{s}^{\ell} \binom{\ell}{t}^n / \binom{\ell n}{m}
\]
when $m = st = tn$. Some of the references mentioned above verify that this formula is correct up to a constant factor, for particular ranges of $m$, $n$, $s$, and $t$, by showing that the number is
\[
\binom{n}{s}^{\ell} \binom{\ell}{t}^n / \binom{\ell n}{m} \sim e^{-1/2+o(1)}.
\]  
(6)

In particular, McKay and Wang [19] established this formula in the sparse biregular case when $st = o \left( (\ell n)^{1/2} \right)$, and Canfield and McKay [7] covered the relatively balanced and dense case. For this they required $\tilde{\mu}(1 - \tilde{\mu}) = \Omega(1/\log n)$ and also $\ell / \log \ell = O(n) = O(\ell \log \ell)$. The asymptotic assertion in (6) is immediately equivalent to $\mathbb{P}_{\mathcal{G}(\ell, n, m)}(s, t) \sim \mathbb{P}_{B_n(\ell, n)}(s, t)H(s, t)$. Consequently, Theorem 1.1 verifies (6) for a new range of parameters for biregular degree sequences of intermediate density.

In [7], the formula in (6) is also verified for a second range of parameters, which can be written as $2 \leq \ell^{2+\epsilon} = O(n\tilde{\mu}(1 - \tilde{\mu}))$ for some $\epsilon > 0$, with $n \to \infty$. This condition is valid when the bipartite graph is quite unbalanced: it requires that the average degree of the vertices in $S$ is considerably larger than the square of the number of vertices in $S$. Those parameters do not overlap greatly with other known cases. For instance, when $\mu$ is close to 1, there remains a gap between these parameters and those covered by Theorem 1.1, namely for $\ell$ approximately between $\sqrt{n}$ and $n^{2/3}$.

Methods for enumerating bipartite graphs with a given degree sequence often lend themselves to digraphs without loops as well (just as in the current paper). O’Neill [23], for example, provides an asymptotic formula for the number of $d$-regular digraphs on $n$ vertices when $d < \left( (\log n)^{1/4-\epsilon} \right)$ for
some $\varepsilon > 0$. A formula by Bender [4] works for arbitrary sequences when all the degrees are bounded. Bollobás and McKay [5] extended the result in [23] to slightly larger $d$ and irregular sequences. The best result in the sparse range that we are aware of is due to McKay [17] which covers all sequences with maximum entry $\Delta$ satisfying $\Delta^3 = o(m)$. For the dense case, that is, when the number of edges is $\Theta(n^2)$, a result by Greenhill and McKay [10] implies an asymptotic formula. Barvinok [1] provides upper and lower bounds which are coarser but their bounds apply to a wider range of in- and out-degree sequences. Theorem 1.1 again bridges the gap between the earlier sparse case results and the dense range as the average $d$ may be anywhere between $n^\varepsilon$ and $\varepsilon n$.

1.2 Models for the degree sequences of random graphs

In 1997, McKay and Wormald [20] showed that if a certain enumeration formula holds for the number of graphs of a given degree sequence then the degree sequences of the random graph models $G(n, m)$ and $G(n, p)$ can be modeled by certain binomial-based models. The model for $G(n, p)$ showed that the degree sequence was distributed almost the same as a sequence of independent binomial random variables, subject to having even sum, but with a slight twist that introduces dependency. It was also shown there that for properties of the degree sequence satisfying some quite general conditions, this conditioning and dependency make no significant difference, and hence those properties are essentially the same as for a sequence of independent binomials.

At that time, the existing formulae for the sparse and the dense case supplied that relationship of the models. Recently, the enumeration results of [14] for the medium range provide the missing formulae for the gap range of densities, establishing a conjecture from [20]. A natural supposition of the arguments in [20] to show that the existing enumeration results for dense bipartite graphs and directed graphs imply a binomial-based model of the degree sequences of such graphs. This is quite analogous to the model in the graph case, except that it contains an extra complicating conditioning required because the sum of degrees of the vertices in each part must be equal. McKay and Skerman point out that, once the enumeration formulae are proved in the missing ranges, one would expect the model results to follow. Our enumeration results stated above provide what is necessary to immediately establish the relevant conjecture in the case of $G(\ell, n, m)$ and $\tilde{G}(n, m)$, as described below, provided that $\ell$ and $n$ are not too disparate. For their binomial random graph siblings $G(\ell, n, p)$ and $\tilde{G}(n, p)$, in which edges are selected independently with probability $p$, one would expect that arguments similar to those in [18], in conjunction with our results, will now suffice.

Let $A_n$ and $B_n$ be two sequences of probability spaces with the same underlying set for each $n$. Suppose that whenever a sequence of events $H_n$ satisfies $\mathbb{P}(H_n) = n^{-O(1)}$ in either model, it is true that $\mathbb{P}_{A_n}(H_n) \sim \mathbb{P}_{B_n}(H_n)$. Then we call $A_n$ and $B_n$ asymptotically quite equivalent (a.q.e.). We use $o$ to mean a function going to infinity as $n \to \infty$, possibly different in all instances.

Theorem 1.3.

(a) The probability spaces $D(\tilde{G}(n, m))$ and $\tilde{B}_m(n)$ are a.q.e. provided that $\log^4 n = o(\min\{m, n^2 - m\})$.

(b) The probability spaces $D(G(\ell, n, m))$ and $B_m(\ell, n)$ are a.q.e. provided that at least one of the following holds:

(i) (Sparse case) $\log^3 n + \log^3 \ell = o(m)$ and $m = o(\min\{\ell / \log n)^3, (n/ \log \ell)^3, (n\ell)^{3/4}\})$;

(ii) (Intermediate case) for a sufficiently small constant $\mu_0 > 0$ we have $m < \mu_0 n \ell$, and for some $\varepsilon > 0$ we have $\ell^3 = o(n^2 m^{1-\varepsilon})$ and $n^3 = o(\ell^2 m^{1-\varepsilon})$;
(iii) (Dense case) For some constant $\mu_1 > 0$ we have $\min\{m, n \ell - m\} > \mu_1 \ell n$, and for sufficiently small $c = c(\mu_1) > 1$ we have $\max\{\ell / n, n / \ell\} < c \log n$.

We prove this theorem in Section 3. Note that larger values of $m$ than in the dense case of (b) are covered by complementation of the results for smaller values. We remark that McKay and Skerman [18, Theorem I(d) and (c)] immediately imply the assertions of (a) for the range $\min\{m, n^2 - m\} > n^2 / \log n$ (and slightly more), and of (b)(iii) for a slightly wider range of parameters than stated here. For the other parts we combine concentration results with enumeration results. Previously existing enumeration results are used from McKay [17, Theorem 4.6] for (a) in the range $\ell \leq \min\{m, n \ell - m\}$, and of (b)(iii) for a slightly wider range of parameters than stated here. For the bipartite case (b)(i). Theorem 1.1 provides the enumeration results required for the intermediate density ranges of both (a) and (b). We note that under the “moderately balanced” condition $\min\{\ell, n\} \geq \max\{\ell, n\}^{10/11+\varepsilon}$ for fixed $\varepsilon > 0$, all values of $m < \mu_0 n \ell$ are covered by Theorem 1.3, provided of course that $\log n = o(\min\{m, n \ell - m\})$. Note also that by symmetry, the theorem would be unchanged if we had defined $\alpha > 0$ with reference to $\ell^{-O(1)}$ instead of $n^{-O(1)}$.

1.3 | Edge probabilities

As a by-product of our proof of Theorem 1.1 in Section 3, we obtain asymptotic formulae for the edge probabilities in a random bipartite graph with a given degree sequence, and of a random digraph with a given sequence of out- and in-degrees.

**Theorem 1.4.** Let $n$, $\ell$, $m$, and $\mathfrak{D}$ be as in Theorem 1.1 and let $\mathcal{G} = \mathcal{G}(\ell, n, m)$ or $\mathcal{G} = \tilde{\mathcal{G}}(n, m)$. Let $a \in S$ and $v \in T$, with $v \neq \tilde{a}$ in the digraph case. Then uniformly for $d = (s, t) \in \mathfrak{D}$, the probability that $av$ is an edge of $G \in \mathcal{G}$, conditional on the event that $D(G) = d$, is

$$
\frac{s_a t_v}{m - \delta d t} \left(1 - \frac{(s_a - s)(t_v - t)}{m - \delta d t - ts \ell} + \frac{(s_a - s)\sigma^2(t)}{ts (\ell - t)} + \frac{(t_v - t)\sigma^2(s)}{ts (n - s)} + \frac{\delta^2 d (t_v + s_v)}{n - 1} + O\left(\min\{s, t\}^{3k - 2} \frac{m}{n \ell}\right)\right),
$$

where $s = m / \ell$ and $t = m / n$.

We prove this theorem in Section 3. Greenhill and McKay [10, Theorems 2.2 and 3.2] similarly provide a formula (one for the bipartite, one for the digraph case) for the probability of a fixed edge in the dense range when $n = \ell^{1+o(1)}$, $m$ is roughly quadratic in $n$ and the sequence is nearly regular.

2 | PRELIMINARIES

As we indicated in the introduction, the argument in this article derives from that in [14], whose notation and structure we will follow quite closely. Differences occur though to account for the fact that we are dealing with certain forbidden edges. Naturally, we resort to notation used in [14] and add notation that is special to the bipartite case. We then state several intermediate results from [14].

2.1 | Notation

Our graphs are simple, that is, they have no loops or multiple edges. We write $a \sim b$ to mean that $a / b \to 1$, $f = O(g)$ if $|f| \leq C g$ for some constant $C$, and $f = o(g)$ if $f / g \to 0$. We use $\omega$ to mean a function going to infinity, possibly different in all instances, and $f = \Theta(g)$ to mean $f = O(g)$ and
g = O(f). Also \( \binom{V}{2} \) denotes the set of 2-subsets of the set \( V \), and \( V \) is often of the form \([N]\), which denotes \( \{1, \ldots, N\} \). In this article multiplication by juxtaposition has precedence over “/”, so for example \( j/\mu N^2 = j/(\mu N^2) \).

Let \( N \) be an integer and let \( V = [N] \). Assume that \( A = A(N) \subseteq \binom{[N]}{2} \) is specified; we call this the set of \textit{allowable pairs}. Note that as usual we regard the edge joining vertices \( u \) and \( v \) as the unordered pair \( \{u, v\} \), and denote this edge by \( uv \) following standard graph theoretic notation. A sequence \( d = (d_1, \ldots, d_N) \) is called \textit{\( A \)-realisable} if there is a graph \( G \) on vertex set \( V \) such that vertex \( a \in V \) has degree \( d_a \) and all edges of \( G \) are allowable pairs. In this case, we say \( G \) realises \( d \) over \( A \). In standard terminology, if \( d \) is \( \binom{V}{2} \)-realisable, it is \textit{graphical}. Let \( G_A(d) \) be the set of all graphs that realize \( d \) over \( A \). The graph case when \( A = \binom{V}{2} \) is dealt with in [14]. In this article, we are particularly interested in the following two special cases of \( A \).

- **Bipartite graph case**
  Let \( \ell, n \) be integers and set \( N = \ell + n \). Set \( A = A^{bi} = \{uw : u \in [\ell], w \in [N] \setminus [\ell]\} \). Then \( G_A(d) \) is the set of all bipartite graphs \( G \) on vertex set \([N]\) that realize the degree sequence \( d = (s, t) \) with one part being \( S = [\ell] \) and the other part \( T = [N] \setminus [\ell] \).

- **Digraph case**
  Assume that \( N \) is even and let \( n \) be an integer such that \( N = 2n \). Set \( A = A^{di} = \{uw : u \in [n], w \in [n + 1, 2n], u + n \neq w\} \). Then \( G_A(d) \) corresponds to the set of all bipartite graphs \( G \) on vertex set \([2n]\) that realize the degree sequence \( d = (s, t) \) with one part being \( S = [n] \) and the other part \( T = [2n] \setminus [n] \) that do not contain any edge of a predefined matching, or equivalently, \( G_A(d) \) corresponds to the set of all digraphs \( G \) on vertex set \([n]\) that have no loops and that realize the out-degree sequence \( s \) and in-degree sequence \( t \). Recall that \( \tilde{a} = a + n \) for \( a \in S \) and \( \tilde{v} = v - n \) for \( v \in T \) so that edges of the form \( a\tilde{a} \) are forbidden.

Let \( \ell, n \) be integers and suppose that \( d \) is a sequence of length \( \ell + n \). Recall the definitions of \( s = s(d), t = t(d), s, t, M_1(d), \sigma(s, t) \) and \( \sigma^2(d) \) from the introduction. We also use \( \Delta \) or \( \Delta(d) \) to denote \( \max d_i \) for any vector \( d \), in line with the notation for maximum degree of a graph. With \( A \) understood (to be either \( A^{bi} \) or \( A^{di} \) in this article) we write \( \mu = \mu(d) \) for the quantity \( M_1(d)/2|A| \) and note that this agrees with the definition of \( \mu \) given just above (2) in the introduction. Throughout this article we use \( e_i \) to denote the elementary unit vector with 1 in its coordinate indexed by \( i \). We say \( d \) is \textit{matchable} if \( M_1(s) = M_1(t) \). Clearly being matchable is necessary for \( d \) to be \( A \)-realisable in either of the cases \( A = A^{bi} \) or \( A = A^{di} \). Furthermore, we say that \( d \) is \textit{S-heavy} if \( M_1(s) = M_1(t) + 1 \), and we call it \textit{T-heavy} if \( M_1(s) = M_1(t) - 1 \).

Finally, we use \( 1 \pm \xi \) to denote a quantity between \( 1 - \xi \) and \( 1 + \xi \) inclusively.

### 2.2 | Cardinalities, probabilities, and ratios

We first quote a simple result by which we leverage absolute estimates of probabilities from comparisons of related probabilities.

**Lemma 2.1** (Lemma 2.1 in [14]). Let \( S \) and \( S' \) be probability spaces with the same underlying set \( \Omega \). Let \( G \) be a graph with vertex set \( \mathfrak{G} \subseteq \Omega \) such that \( \mathbb{P}_{S}(v), \mathbb{P}_{S'}(v) > 0 \) for all \( v \in \mathfrak{G} \). Suppose that \( \varepsilon_0, \delta > 0 \) such that \( \min\{\mathbb{P}_{S}(\mathfrak{G}), \mathbb{P}_{S'}(\mathfrak{G})\} > 1 - \varepsilon_0 > 1/2 \), and such that for every edge \( uv \) of \( G \),

\[
\frac{\mathbb{P}_{S'}(u)}{\mathbb{P}_{S'}(v)} = e^{O(\delta)} \frac{\mathbb{P}_{S}(u)}{\mathbb{P}_{S}(v)},
\]
where the constant implicit in $O(\cdot)$ is absolute. Let $r$ be an upper bound on the diameter of $G$ and assume $r < \infty$. Then for each $v \in \mathcal{W}$ we have

$$\mathbb{P}_{S'}(v) = e^{O(r^{\delta+\epsilon_3})}\mathbb{P}_{S}(v),$$

with the bound in $O(\cdot)$ uniform overall $v$.

The lemma gives the top level structure of the proof of our main theorem. It allows us to approximate probabilities in the “true” probability space $S'$ (which will be the degree sequence of $\mathcal{G}(\ell', n, m)$ or of $\mathcal{G}(n, m)$) by probabilities in an “ideal” probability space $S$ (which we define in the proof). To apply it, we need to find a set $\mathcal{W}$ of probability roughly 1, and a graph $G$ on the set of possible sequences $d$ for which we can find a fine estimate on the ratios of point probabilities in either space. This leads to computing ratios of closely related instances of the expression on the right hand side of (5). Let $d = (s, t)$ be a sequence where $s$ and $t$ are of length $\ell'$ and $n$, respectively, let $a, b \in S$, and assume that $d$ is $S$-heavy. We will compare the probabilities of the two sequences $(s - e_a, t)$ and $(s - e_b, t)$.

To analyze the ratios of such nearby sequences in the “true” probability space note that, with the above notation, $\mathbb{P}_{D_{\mathcal{G}}}(d)$ in Theorem 1.1 is just $|\mathcal{G}_{A}(d)|/|\mathcal{G}|$ where $\mathcal{G}$ is the random graph space $\mathcal{G}(\ell', n, m)$ or $\mathcal{G}(n, m)$. Let us introduce some more notation. Let $F \subseteq \mathcal{A}$, that is, is a subset of the allowable edges. We write $\mathcal{N}_F(d)$ and $\mathcal{N}_F^e(d)$ for the number of graphs $G \in \mathcal{G}_{A}(d)$ that contain, or do not contain, the edge set $F$, respectively. (When $\mathcal{N}$ and similar notation is used, the set $A$ should be clear by context.) We abbreviate $\mathcal{N}_F(d)$ to $\mathcal{N}_{ab}(d)$ if $F = \{ab\}$ (i.e., contains the single edge $ab$), and put $\mathcal{N}(d) = |\mathcal{G}_{A}(d)|$. Additionally, for a vertex $a \in V$, we set $\mathcal{A}(a) = \{v \in V : av \in A\}$, and, with $d$ understood, we use $A^+(a)$ for the set of $v \in A(a)$ such that $\mathcal{N}_{av}(d) > 0$.

For vertices $a, b \in V$, if $d$ is a sequence such that $d - e_b$ is $A$-realisable, we define

$$R_{a,b}(d) = \frac{\mathcal{N}(d - e_a)}{\mathcal{N}(d - e_b)},$$

and note that this is exactly $\mathbb{P}_{D_{\mathcal{G}}}(d - e_a)/\mathbb{P}_{D_{\mathcal{G}}}(d - e_b)$. Estimating those “true” ratios will be tightly linked to estimating the following. For $F \subseteq \mathcal{A}$, let

$$P_F(d) = \frac{\mathcal{N}_F(d)}{\mathcal{N}(d)},$$

which is the probability that the edges in $F$ are present in a graph $G$ that is drawn uniformly at random from $\mathcal{G}_{A}(d)$. Of particular interest are the probability of a single edge $av$ and a two-edge path on vertices $a, v, b$, for which we simplify the notation to

$$P_{av}(d) = P_{\{av\}}(d), \quad Y_{a,v,b}(d) = P_{\{av,bv\}}(d).$$

We use $Y$ for the two-edge path to distinguish it from the edge probability. In Section 2.4, we will describe relations that let us recursively compute approximations to $R_{a,b}(d)$, $P_{av}(d)$, and $Y_{a,v,b}(d)$, with the ultimate goal of using $R_{a,b}(d)$ to estimate the ratio of probabilities of degree sequences that are “near” each other. For considering the ratios of nearby sequences in the “ideal” probability space, note first that the following are immediate from (1):

$$\frac{\mathbb{P}_{B_n(\ell,n)}(s - e_a, t)}{\mathbb{P}_{B_n(\ell,n)}(s - e_b, t)} = \frac{s_a(n + 1 - s_b)}{s_b(n + 1 - s_a)}$$

and

$$\frac{\mathbb{P}_{\overline{B}_n(n)}(s - e_a, t)}{\mathbb{P}_{\overline{B}_n(n)}(s - e_b, t)} = \frac{s_a(n - s_b)}{s_b(n - s_a)}.$$
Similarly, straight from the definition of $\tilde{H}$ in (2) we have

$$\frac{\tilde{H}(s - e_a, t)}{\tilde{H}(s - e_b, t)} = \exp \left( \frac{s_b - s_a}{s'(1 - \mu')} \left( 1 - \frac{\sigma(t)}{t(1 - \mu')} \right) + \frac{\delta^d(t_a - t_b)}{sn(1 - \mu')} \right),$$

where $\mu' = \mu(s - e_a, t)$, $s = s(s - e_a, t)$, $t = t(s - e_a, t)$, which are, in this case, equal to $\mu(s - e_b, t)$, $s(s - e_b, t)$, and $t(s - e_b, t)$, respectively (recalling that $\mu$ is slightly different in the two cases of $A^{bi}$ and $A^{di}$), and where, we recall, $\delta^d$ is the indicator variable for the digraph case. Therefore, denoting by $H(d')$ the function $\mathbb{P}_B(d')/\tilde{H}(d')$, we get a “combined goal ratio” in the two cases which is

$$\frac{H(s - e_a, t)}{H(s - e_b, t)} = \frac{s_a(n + \delta^{bi} - s_b)}{s_b(n + \delta^{bi} - s_a)} \exp \left( \frac{s_b - s_a}{s'(1 - \mu')} \left( 1 - \frac{\sigma(t)}{t(1 - \mu')} \right) + \frac{\delta^d(t_a - t_b)}{sn(1 - \mu')} \right),$$

(9)

where $\delta^{bi} = 1 - \delta^d$.

The recursive relations mentioned above also require us to have information on the ratios of numbers regarding two degree sequences of differing total degree, for which we use the following, which is [14, Lemma 2.2].

Lemma 2.2. Let $av \in A$ and let $d$ be a sequence of length $N$. Then

$$\mathcal{N}_{av}(d) = \mathcal{N}'(d - e_a - e_v) - \mathcal{N}_{av}(d - e_a - e_v)$$

$$\begin{cases} \mathcal{N}'(d - e_a - e_v)(1 - P_{av}(d - e_a - e_v)) & \text{if } \mathcal{N}'(d - e_a - e_v) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

In Lemma 2.3 in [14] we bound the probability of an edge of a random graph in $G(d)$ in the graph case. A similar switching argument is used to obtain corresponding bounds in the bipartite and digraph cases. Recall that by $\Delta(d)$ we denote $\max_i d_i$, and that $M_1(d) = \sum_i d_i$.

Lemma 2.3. Let $A$ be $A^{bi}$ or $A^{di}$ and let $(s, t)$ be an $A$-realisable sequence. Then for any $av \in A$ we have

$$P_{av}(s, t) \leq \frac{\Delta(s)\Delta(t)}{M_1(s)(1 - 2(\Delta(s) + 1)(\Delta(t) + 1)/M_1(s))}.$$

Proof. Assume without loss of generality that $a \in S$ which forces $v \in T$ in both the digraph and bipartite cases. For each bipartite graph $G$ with degree sequence $(s, t)$ and an edge joining $a$ and $v$, we can perform a switching (of a type often used previously in graphical enumeration) by removing both $av$ and another randomly chosen edge $bw$ (with $b \in S$, and $w \in T$), and inserting the edges $aw$ and $bv$, provided that no multiple edges are formed. Note that the way we choose $b$ and $w$ no loops can occur this way. In the digraph case, we should also make sure that $w \neq \tilde{a}$ and that $b \neq \tilde{v}$, since the pairs $a\tilde{a}$ and $\tilde{v}v$ are not allowable (see the definition of the digraph case in Section 2.1). The number of such switchings that can be applied to $G$ with the vertices of each edge ordered, is at least

$$M_1(s) - (\Delta(s) + 1)\Delta(t) - \Delta(s)(\Delta(t) + 1),$$

since there are $M_1(s)$ ways to choose $b \in S$ and $w \in T$, whereas the number of such choices that are ineligible is at most the number of choices with $b$ being a neighbor of $v$ (which automatically rules out $b = a$) or $b = \tilde{v}$, or similarly for $w$. On the other hand, for each graph $G'$ in which $av$ is not an edge,
the number of ways that it is created by performing such a switching backwards is at most \( \Delta(s) \Delta(t) \). Counting the set of all possible switchings over all such graphs \( G \) and \( G' \) two different ways shows that the ratio of the number of graphs with \( av \) to the number without \( av \) is at most

\[
\beta := \frac{\Delta(s) \Delta(t)}{M_1(s) - 2(\Delta(s) + 1)(\Delta(t) + 1)}.
\]

Hence \( P_{av}(d) \leq \beta/(1 + \beta) \), and the lemma follows in both cases.

2.3 Realisability

As in the graph case in [14], before estimating how many (bipartite) graphs have degree sequence \( d \), for preparation we need to know that there is at least one such graph for various \( d \). Mirsky [22, p. 205] gives a necessary and sufficient condition for the existence of a non-negative integer matrix with row and column sums in specified intervals. For the case that those sums are specified precisely, the statement is the following.

**Theorem 2.4** (Corollary of Mirsky [22]). Let \( 0 \leq r_i, 0 \leq c_j, m_{ij} \geq 0 \) be integers for all \( 1 \leq i \leq \ell \), \( 1 \leq j \leq n \) such that \( \sum_{1 \leq i \leq \ell} r_i = \sum_{1 \leq j \leq n} c_j \). Then there exists an \( \ell \times n \) integer matrix \( B = (b_{ij}) \) with row sums \( r_1, \ldots, r_\ell \) and column sums \( c_1, \ldots, c_n \) such that \( 0 \leq b_{ij} \leq m_{ij} \) for all such \( i \) and \( j \) if and only if, for all \( X \subseteq \{1, \ldots, \ell\} \) and \( Y \subseteq \{1, \ldots, n\} \),

\[
\sum_{i \in X, j \in Y} m_{ij} \geq \sum_{i \in X} r_i - \sum_{j \in Y} c_j.
\]

We use this to show existence of bipartite graphs with given degrees and forbidden edges under rather simple conditions that suffice for the cases of interest. In order to apply this to digraphs, one would set \( \ell = n \) and regard the edges as directed from the first part to the second. For loopless digraphs, we merely forbid all edges of the form \( (i, i + n) \). Recall that, with \( \ell \) and \( n \) understood, we set \( S = [\ell] \) and \( T = [n + \ell] \setminus [\ell] \) for convenience.

**Lemma 2.5.** Given a constant \( C \geq 1 \), the following holds for \( \ell, n \) sufficiently large. Let \( s_a \geq 1 \) and \( t_v \geq 1 \) be integers for all \( a \in S, v \in T \), with \( m := \sum_{a \in S} s_a = \sum_{v \in T} t_v \) satisfying \( m \leq \ell n/9 \). Also let \( F \subseteq \{ av : a \in S, v \in T \} \) be a set of unordered pairs, representing forbidden edges, with no more than \( C \) pairs in \( F \) containing any \( w \in S \cup T \). Let \( s \) and \( t \) denote the vectors \((s_a)_{a \in S}\) and \((t_v)_{v \in T}\). If \( \Delta(s) \leq 2m/\ell \) and \( \Delta(t) \leq 2m/n \), then there exists a bipartite graph with bipartition \((S, T)\) with degrees \( s_a \) for \( a \in S \) and \( t_v \) for \( v \in T \), and containing no edge in the forbidden set \( F \).

**Proof.** We will apply Theorem 2.4 with \( m_{ij} = 0 \) if \( \{i, j + \ell\} \) is a forbidden edge, and \( m_{ij} = 1 \) otherwise, and with \( r_i = s_i \) and \( c_j = t_{j+\ell} \). Note that \( \sum_{a \in X, v \in Y} m_{ij} \geq x - C \min\{x, y\} \) for all subsets \( X \subseteq S, Y \subseteq T \), where \( x = |X| \) and \( y = |Y| \). We will show that for all \( X \subseteq S \) and \( Y \subseteq T \), with \( x = |X| \) and \( y = |Y| \), we have

\[
m - C \min\{x, y\} \geq \sum_{a \in X} s_a + \sum_{v \in Y} t_v - xy.
\]

Equivalently, \( xy - C \min\{x, y\} \geq \sum_{a \in X} s_a - \sum_{v \in T \setminus Y} t_v \). Note that with the previous observation and Theorem 2.4, this implies that there is a matrix \( B \) which is the adjacency matrix of the desired bipartite graph.
Let $s = m/\ell$ and $t = m/n$. Suppose first that $x \geq 2t + C$. Then using $\sum_{a \in \mathcal{X}} s_a \leq m$ and $\sum_{v \in \mathcal{Y}} t_v \leq y\Delta(t) \leq 2yt$, we find that the right hand side of (10) is at most $m - yC$, and (10) follows. A symmetric argument works if $y \geq 2s + C$. So we may assume that neither of these occurs. Then

$$\sum_{a \in \mathcal{X}} s_a + \sum_{v \in \mathcal{Y}} t_v \leq x\Delta(s) + y\Delta(t) \leq (2t + C)2s + (2s + C)2t,$$

and the left hand side is at least $m - C(x + y)/2 \geq m - C(s + t) - C^2$. Thus (10) follows if we show that $m \geq 8st + 5C(s + t) + C^2$, that is, if $1 \geq 8m/\ell n + 5C(1/\ell + 1/n) + C^2/m$ (since $s = m/\ell$ and $t = m/n$). This holds for $\ell$ and $n$ sufficiently large because $n \leq m \leq \ell n/9$. □

### 2.4 Recursive relations

In this subsection we collect results about recursive relations that were obtained, with a quite short proof, in [14]. The results were stated for an arbitrary set $\mathcal{A}$ of allowable pairs in $\binom{V}{2}$. Recall the definitions of the probabilities $P_{av}(d)$ and $Y_{a,v,b}(d)$ in (8), and of the ratio $R_{a,b}(d)$ in (7).

**Proposition 2.6** (Proposition 3.1 in [14]). Let $d$ be a sequence of length $n$ and let $\mathcal{A} \subseteq \binom{[n]}{2}$.

(a) Let $a, v \in V$. If $\mathcal{N}_{av}(d) > 0$ then

$$P_{av}(d) = d_v \left( \sum_{b \in \mathcal{A} \cap \{v\}} R_{b,a}(d - e_v) \frac{1 - P_{bv}(d - e_b - e_v)}{1 - P_{av}(d - e_a - e_v)} \right)^{-1}.$$

(b) Let $a, b \in V$. If $d - e_a$ and $d - e_b$ are $\mathcal{A}$-realisable then

$$R_{a,b}(d) = \frac{d_a}{d_b} \frac{1 - B(a, b, d - e_b)}{1 - B(b, a, d - e_a)}, \quad (11)$$

where

$$B(i, j, d') = \frac{1}{d_i} \left( \sum_{v \in \mathcal{A}(i) \setminus \mathcal{A}(j)} P_{iv}(d') + \sum_{v \in \mathcal{A}(i) \cap \mathcal{A}(j)} Y_{i,v,j}(d') \right), \quad (12)$$

provided that $B(b, a, d - e_a) \neq 1$.

(c) Let $a, v, b$ be distinct elements of $V$. If $d - e_a - e_v$ is $\mathcal{A}$-realisable then

$$Y_{a,v,b}(d) = \frac{P_{av}(d) \left( P_{bv}(d - e_a - e_v) - Y_{a,v,b}(d - e_a - e_v) \right)}{1 - P_{av}(d - e_a - e_v)}.$$

Roughly speaking, these recursive relations can be used to feed in initial approximations for $P$ and $Y$ (and thus for $R$) and obtaining updated approximations with a better relative error, which can be further improved by iterating the procedure. Useful formulae arise after only few iterations (as in the sparse case for ordinary graphs treated in [14]), but to reach the highest density range we need an unbounded number of iterations. To facilitate the analysis, we define operators that mimic the updating step.

Before we define the operators, let us briefly explain the combinatorial idea behind (11) and (12). Recall that $R_{ab}(d)$ is just the ratio $\mathcal{N}(d - e_a)/\mathcal{N}(d - e_b)$. To compare these two values, we pick a
graph $G$ in $G(d - e_b)$ uniformly at random, and then replace a random edge $av$ with a new edge $bv$. This procedure creates a graph in $G(d - e_b)$, unless $ab$ is an edge in $G$ (whence we would create a loop) or $bv$ is already an edge (whence we would create a double edge). Then $B(a, b, d - e_b)$ just turns out to be the probability of one of these “bad” events to happen. Similar combinatorial ideas yield the relations for (a) and (c).

We next present the operators from [14] that are motivated by the recursive relations in Proposition 2.6.

Let $\mathbb{Z}_\geq 0$ denote the set of non-negative integer sequences of length $N$. For a given integer $N$ and a set $\mathcal{A} \subseteq \binom{[N]}{2}$ we define $\tilde{\mathcal{A}}$ to be the set of ordered pairs $(u, v)$ with $\{u, v\} \in \mathcal{A}$. Ordered pairs are needed here because, although the functions of interest are symmetric in the sense that the probability of an edge $uv$ is the same as $vu$, our approximations to the probability do not obey this symmetry. Similarly, let $\tilde{\mathcal{A}}_2$ denote the set of ordered triples $(u, v, w)$ with $u, v$ and $w$ all distinct and $(u, v), \{v, w\} \in \mathcal{A}$. Suppose we are given $p : \tilde{\mathcal{A}} \times \mathbb{Z}_{\geq 0}^N \to \mathbb{R}_{\geq 0}$, $y : \tilde{\mathcal{A}}_2 \times \mathbb{Z}_{\geq 0}^N \to \mathbb{R}_{\geq 0}$ and $r : [N]^2 \times \mathbb{Z}_{\geq 0}^N \to \mathbb{R}_{\geq 0}$. These three functions can be regarded as approximations for $P$, $Y$, and $R$, respectively. It may be instructive to have the examples $p, y \equiv 0$ and $r \equiv 1$ in mind in the following definitions. We write $p_{a,v}(d)$ for $p(a, v, d)$ (where $d \in \mathbb{Z}_\geq 0^N$), $y_{a,v,b}(d)$ for $y(a, v, b, d)$ and $r_{a,b}(d)$ for $r(a, b, d)$. (The notation such as $a, v$ in subscripts here is for an ordered pair, distinguishing it from the unordered pair $av$.) We also define an associated function $\text{bad}(p, y)$ as follows. For $d \in \mathbb{Z}_{\geq 0}^N$ and $a, b \in [N]$ with $a \neq b$, set

$$\text{bad}(p, y)(a, a, d) = 0 \quad \text{and} \quad \text{bad}(p, y)(a, b, d) = \frac{1}{d_a} \left( \sum_{v \in \mathcal{A}(a) \setminus \mathcal{A}(b)} p_{a,v}(d) + \sum_{v \in \mathcal{A}(a) \cap \mathcal{A}(b)} y_{a,v,b}(d) \right).$$

(This is undefined, and not needed, when $d_a = 0$.)

We define three operators $P(p, r), Y(p, y)$, and $R(p, y)$, acting on $p$, $y$, and $r$ as above, as follows. For $d \in \mathbb{Z}_{\geq 0}^N$ and $a, v, b \in [N]$ we set

$$P(p, r)(a, v, d) = d_v \left( \sum_{b \in \mathcal{A}(a)} r_{a,b}(d - e_v) \frac{1 - p_{b,v}(d - e_b - e_v)}{1 - p_{a,v}(d - e_a - e_v)} \right)^{-1} \quad \text{for} \quad (a, v) \in \tilde{\mathcal{A}},$$

$$Y(p, y)(a, v, b, d) = d_v d_a \frac{p_{a,v}(d)(p_{b,v}(d - e_b - e_v) - y_{a,v,b}(d - e_a - e_v))}{1 - p_{a,v}(d - e_a - e_v)} \quad \text{for} \quad (a, v, b) \in \tilde{\mathcal{A}}_2,$n

$$R(p, y)(a, b, d) = d_a d_b \frac{1 - \text{bad}(p, y)(a, b, d - e_b)}{1 - \text{bad}(p, y)(b, a, d - e_a)} \quad \text{for} \quad a, b \in S \ \text{or} \ a, b \in T.$$

We note that, in light of Proposition 2.6, the “real” functions $P$, $Y$, and $R$ are fixed points of these operators in a certain sense that we make precise inside the proof of Theorem 1.1. We invite the reader to see how the operators act on the functions $p, y \equiv 0$, $r \equiv 1$ in the toy example mentioned above. We observed in [14] that the operators are “contractive” in a certain sense in the vicinity of these fixed point functions. We define a suitable domain of contraction as follows.

**Definition 2.7.** Let $\mathfrak{D}_0 \subseteq \mathbb{Z}_{\geq 0}^N$ and let $\mu \in \mathbb{R}$. We use $\Pi_\mu(\mathfrak{D}_0)$ to denote the set of pairs of functions $(p, y)$ with $p : \tilde{\mathcal{A}} \times \mathbb{Z}_{\geq 0}^N \to \mathbb{R}_{\geq 0}$ and $y : \tilde{\mathcal{A}}_2 \times \mathbb{Z}_{\geq 0}^N \to \mathbb{R}_{\geq 0}$ such that for all matchable $d \in \mathfrak{D}_0$, we have

(Pa) $0 \leq p_{a,v}(d) \leq \mu$ for all $(a, v) \in \tilde{\mathcal{A}},$

(Pb) $\sum_{v \in \mathcal{A}(a) \setminus \mathcal{A}(b)} y_{a,v,b}(d) \leq \mu d_a$ for all $a \neq b \in [N],$ and

(Pc) $0 \leq y_{a,v,b}(d) \leq \mu p_{b,v}(d)$ for all $(a, v, b) \in \tilde{\mathcal{A}}_2$. 


For the next lemma, we need to adapt [14, Lemma 5.2] to the present bipartite setting. Recall the definitions of $A^b_i$ and $A^d_i$, and of $S$ and $T$ and $\tilde{a}$ in Section 2.1. In this setting, $A(a) \cap A(b) \neq \emptyset$ if and only if both $a, b \in S$ or both $a, b \in T$. For such $a, b$ we have $|A(a) \setminus A(b)| \leq 1$. Further note that $(a, v) \in \tilde{A}$ if and only if $a \in S$ and $v \in T$ in the bipartite case or $v \in T \setminus \{\tilde{a}\}$ in the digraph case (or $S$ and $T$ swapped), and thus $(a, v, b) \in \tilde{A}_2$ if and only if both $a, b \in S, a \neq b$ and $v \in T$ (bipartite) or $v \in T \setminus \{\tilde{a}, \tilde{b}\}$ (digraph); or $S$ and $T$ swapped.

Also, for $d = (s, t)$ of length $\ell + n$, we let $Q^0_1(d)$, $Q^0_2(d) \subseteq \mathbb{Z}_{\geq 0}^n$ be the set of matchable and $S$-heavy, respectively, vectors of non-negative integers that have $L_1$-distance at most $r$ from $d$. Recall that we use $1 \pm \xi$ to denote a quantity between $1 - \xi$ and $1 + \xi$ inclusively. With these definitions, Lemma 5.2 in [14] specializes to the following.

**Lemma 2.8.** There is a constant $C > 0$ such that the following holds. Let $\ell, n$ be integers, $N = \ell + n$, and let $A$ be either $A^b_i$ or $A^d_i$. Let $d \in \mathbb{Z}_{\geq 0}^n$ such that $d_a > 1$ for all $a \in [N]$. Let $0 < \xi \leq 1$ and $0 < \mu_0 = \mu_0(\xi, n) < C$. Let $(p, y)$ and $(p', y')$ be members of $\Pi_{\mu_0}(Q^0_2(d))$, and let $r, r' : [N]^2 \times \mathbb{Z}_{\geq 0}^n \to \mathbb{R}$. Let $a, b \in S$, $v \in T$.

(a) If $d$ is $S$-heavy, $p_{c,w}(d') = p'_{c,w}(d')(1 \pm \xi)$ for all $(c, w) \in \tilde{A}$ and all $d' \in Q^0_1(d)$, and $y_{c,w,h}(d') = y'_{c,w,h}(d')(1 \pm \xi)$ for all $(c, w, h) \in \tilde{A}_2$ and all $d' \in Q^0_1(d)$, then

$$R(p, y)_{a,b}(d) = R(p', y')_{a,b}(d)(1 + O(\mu_0 \xi)).$$

(b) If $d$ is matchable, $v \neq \tilde{a}$, $p_{c,v}(d') = p'_{c,v}(d')(1 \pm \xi)$ for all $c \in S \setminus \{v\}$ and all $d' \in Q^0_2(d)$, and $r_{c,a}(d') = r'_{c,a}(d')(1 \pm \mu_0 \xi)$ for all $c \in S \setminus \{v\}$ and all $d' \in Q^0_2(d)$, then

$$P(p, r)_{a,b}(d) = P(p', r')_{a,b}(d)(1 + O(\mu_0 \xi)).$$

(c) If $d$ is matchable, $a \neq b$, $v \notin \{\tilde{a}, \tilde{b}\}$, $p_{c,v}(d') = p'_{c,v}(d')(1 \pm \mu_0 \xi)$ for all $c \in S \setminus \{v\}$ and all $d' \in Q^0_2(d)$, and $y_{c,w,h}(d') = y'_{c,w,h}(d')(1 \pm \xi)$ for all $(c, w, h) \in \tilde{A}_2$ and all $d' \in Q^0_2(d)$, then

$$\mathcal{Y}(p, y)_{a,v,b}(d) = \mathcal{Y}(p', y')_{a,v,b}(d)(1 + O(\mu_0 \xi)).$$

The constants implicit in $O(\cdot)$ are absolute.

### 2.5 | Concentration of random variables

When $d = (s, t)$ is either the degree sequence of $G(\ell, n, m)$ or $B_m(\ell, n)$ then we need that $\sigma^2(s)$, $\sigma^2(t)$ and $\sigma(s, t)$ (in the digraph case) are concentrated. The following is due to McDiarmid [16], see, for example, Lemma 4.1 in [14].

**Lemma 2.9** (McDiarmid). Let $c > 0$ and let $f$ be a function defined on the set of subsets of some set $U$ such that $|f(A) - f(B)| \leq c$ whenever $|A| = |B| = m$ and $|A \cap B| = m - 1$. Let $A$ be a randomly chosen $m$-subset of $U$. Then for all $\alpha > 0$ we have

$$\mathbb{P} \left( |f(A) - \mathbb{E} f(A)| \geq \alpha c \sqrt{m} \right) \leq 2 \exp(-2\alpha^2).$$

The following are the concentration results we need. Recall that $N = \ell + n$. 
Lemma 2.10. Let \( \ell, n, m \) be integers, let \( \mathbf{d} = (\mathbf{s}, \mathbf{t}) \) be a sequence in \( D(G(\ell, n, m)) \), \( D(\overline{G}(n, m)) \), or either of the binomial models \( B_m(\ell, n) \) and \( \overline{B}_m(n) \), and let \( s = m/\ell \) and \( t = m/n, \mu = \mu(\mathbf{d}) = m/n(\ell - \delta^d) \). Let \( \alpha \in S, v \in T \). Then the following hold for all \( \alpha > 0 \).

(i) We have
\[
\Pr(|s_a - s| \geq \alpha) \leq 2 \exp\left(-\frac{\alpha^2}{2(s + \alpha/3)}\right), \quad \Pr(|t_v - t| \geq \alpha) \leq 2 \exp\left(-\frac{\alpha^2}{2(t + \alpha/3)}\right).
\]

(ii) If \( t \log^2 N + \log^3 N = o(\alpha^2 m) \) then
\[
\Pr\left(|\sigma^2(\mathbf{t}) - \text{Var} t_v| \geq \alpha t + 1/N\right) = o(N^{-\omega}),
\]
where \( \text{Var} t_v = t(1 - \mu)(1 - 1/n)(1 + O(1/n\ell)) \). Furthermore, for \( \mathbf{d} = (\mathbf{s}, \mathbf{t}) \) in \( D(\overline{G}(n, m)) \) or \( \overline{B}_m(n) \),
\[
\Pr\left(|\sigma(\mathbf{s}, \mathbf{t}) - \text{Covar}(s_a, t_0)| \geq \alpha s + 1/n\right) = o(N^{-\omega}),
\]
where \( \text{Covar}(s_a, t_0) = O(\mu) \).

Proof. The proofs of these concentration results are quite routine and we follow the proof of [14,Lemma 6.2] with suitable modifications.

Note that each \( s_a \) is distributed hypergeometrically with parameters \( n(\ell - \delta^d), m, n - \delta^d \) in all four models, and that \( \delta^d = 1 \) only when \( \ell = n \). It follows that \( \mathbb{E}s_a = s \) in all four cases. Similarly, each \( t_v \) is distributed hypergeometrically with parameters \( n(\ell - \delta^d), m, \ell - \delta^d \), and expectation \( t \). Hence (i) holds by [13,Theorems 2.10 and 2.1].

For part (ii), we consider in detail the concentration of \( \sigma^2(\mathbf{t}) \) in the bipartite graph case. For a graph \( G \in G(\ell, n, m) \) with degrees \( t_v \) for \( v \in T \), define
\[
f = f(G) = \sum_{v \in T} \min\{(t_v - t)^2, x\},
\]
where \( x > 1 \) is specified below. Since \( G \) is determined by a random \( m \)-subset of all possible edges, Lemma 2.9 applies, and we just need to determine a suitable value of \( c \). Note that if a graph \( G' \in G(\ell, n, m) \) shares exactly \( m - 1 \) edges with \( G \) then there are at most two vertices \( v \) and \( v' \) in \( T \) such that their degrees in \( G' \) are those in \( G \) increased or decreased by 1, and all other degrees in \( T \) remain unchanged. Furthermore, the average \( t = M_1(t)/n = m/n \) is the same for both graphs \( G \) and \( G' \). If \( t \) is held fixed but \( t_v \) is increased or decreased by 1 for a single \( v \) whilst holding all other \( t_{v'} \) fixed, the value of \( f \) can only change by at most \((\sqrt{x} - (\sqrt{x} - 1)^2 < 2\sqrt{x} \). Thus, the difference \( |f(G) - f(G')| \) is at most \( 4\sqrt{x} \) and we may set \( c = 4\sqrt{x} \) in Lemma 2.9. Replacing \( \alpha \) by \( \alpha \sqrt{m}/(16x) \) gives
\[
\Pr(|f(G) - \mathbb{E}f(G)| \geq \alpha t n) \leq 2 \exp(-\alpha^2 mn/(8x)) = o(N^{-\omega}),
\]
provided that \( \alpha^2 m/(x \log N) \to \infty \). On the other hand, let \( F \) denote the event \( \max_v |t_v - t| > \sqrt{x} \). By (i) and the union bound applied overall \( n \) values of \( v \), we have \( \Pr(F) = o(N^{-\omega}) \) as long as we choose \( x = \omega(t \log N + \log^2 N) \). By the bound assumed in the hypothesis, since \( m = tn \) there exists \( x \) satisfying both of the two conditions just found, and at this point we set \( x \) as such. Provided that \( F \) does not hold, we have \( f(G) = \sum (t_v - t)^2 = n\sigma^2(\mathbf{t}) \). Hence
\[
\Pr\left(|\sigma^2(\mathbf{t}) - \mathbb{E}f(G)/n| \geq \alpha t\right) = o(N^{-\omega}) + O(\Pr(F)) = o(N^{-\omega}).
\]
Moreover, since \( t_v \leq \ell' \) we trivially have \( f(G) \leq n\ell'^2 \) and so

\[
|E f(G) - nE\sigma^2(t)| = O(n\ell'^2)P(f(G) \neq n\sigma^2(t)) = o(N^{-\omega}) = o(N^{-1}).
\]

Combining these two, we obtain

\[
P(\{|\sigma^2(t) - E\sigma^2(t)| \geq at + 1/N\}) = o(N^{-\omega}).
\]

Noting that \( nE\sigma^2(t) = E \sum_v (t_v - \ell^2) = nE(t_v - \ell^2) = n\text{Var} t_v \) for any \( v \in T \), we obtain the first claim in (ii) for the bipartite case. The estimate for \( \text{Var} t_v \) follows from the standard formula for variance of this hypergeometric random variable.

The analogous statements for \( \sigma^2 \) in the other three probability spaces follow by the same proof after very minor modifications. It only remains to consider \( \sigma(s, t) \) in \( D(\vec{G}(n, m)) \) or \( \vec{B}_m(n) \), for which the argument is slightly different. Define

\[
g(a, b) = \text{sign}(a - b) \min\{|a - b|, \sqrt{x}\},
\]

where \( \text{sign}(y) \) is 1, -1 or 0 if \( y \) is positive, negative or 0, respectively, and where \( x \) is a function that satisfies \( s \log n + \log^2 n = o(x) \) and \( x = o(a^2 \text{sn} / \log n) \) as in the proof above. Let \( f = \sum_{i=1}^{n} g(s_i, t_i) \), and adapt the rest of the proof above in the obvious way. This gives the required concentration bound for \( \sigma(s, t) \) near its expected value, which is \( \frac{1}{n} \sum_{b \in S} \text{Covar}(s_b, t_b) = \text{Covar}(s_a, t_a) \). On the other hand, we can bound \( \text{Covar}(s_a, t_a) \) as follows. In \( D(\vec{G}(n, m)) \), the joint distribution of \( (s_a, t_a) \) is multivariate hypergeometric, with \( m \) edges chosen from \( n(n - 1) \) positions, and \( s_a \) and \( t_a \) are the counts for disjoint subsets of size \( n - 1 \) each. Thus, the well known formula gives

\[
\text{Covar}(s_a, t_a) = \frac{m(n - 1)^2(n(n - 1) - m)}{(n(n - 1))^2(n(n - 1) - 1)} = O(m/n^2),
\]

which establishes the final claim for \( D(\vec{G}(n, m)) \). In \( \vec{B}_m(n) \) the random variables \( s_a \) and \( t_a \) are independent since we condition on \( M_1(s) = m \) and \( M_1(t) = m \) separately. Thus the covariance is 0.  

## 3 PROOF OF THEOREMS 1.1, 1.3, AND 1.4

In this section we prove Theorem 1.1. In light of Proposition 2.6, we first obtain an estimate of the ratio \( R_{a,b}(d) \) between the numbers of graphs of related degree sequences. This step is the crux of the whole argument.

To estimate those ratios we first present functions that approximate the ratios and the probabilities. We write these approximations parameterized to facilitate identifying negligible terms. We express the formulae for the approximations of \( P, Y, \) and \( R \) in both the bipartite graph case and the digraph case simultaneously, again using \( \delta_{di} \) as the indicator variable which is 1 in the digraph case and 0 in the bipartite case. For integers \( \ell' \) and \( n \), and a sequence of real numbers \( \mu, \varepsilon_a, \) and so forth, we define the expressions

\[
\pi = \mu(1 + \varepsilon_a)(1 + \varepsilon_v) \left(1 - \frac{\mu \varepsilon_a \varepsilon_v - \varepsilon_a \sigma^2_{\ell'} / t\ell' - \varepsilon_v \sigma^2_s / sn}{1 - \mu} + \frac{\delta_{di}(\varepsilon_a + \varepsilon_v)\mu}{s}\right),
\]

where

\[
\delta_{di} = \left\{ \begin{array}{ll}
0 & \text{if } d_{di} = 0, \\
1 & \text{if } d_{di} = 1,
\end{array} \right.
\]

and

\[
\sigma^2_s = \left( \frac{1}{n} \sum_{i=1}^{n} s_i \right)^2 - \left( \frac{1}{n} \sum_{i=1}^{n} s_i \right)^2.
\]
In the calculations below, there are small changes in most of the variables that turn out to have a negligible effect. Changes in the various occurrences of $\varepsilon$-type terms, however, need to be tracked precisely. In particular, we use

- $\pi(x, y)$ to stand for $\pi$ with $\varepsilon_a, \varepsilon_y$ replaced by $x, y$,
- $\pi(x, y, z, w)$ to stand for $\pi$ with $\varepsilon_a, \varepsilon_y, \varepsilon_z, \varepsilon_w$ replaced by $x, y, z, w$,
- $\rho(x, y, z, w)$ to stand for $\rho$ with $\varepsilon_a, \varepsilon_y, \varepsilon_z, \varepsilon_w$ replaced by $x, y, z, w$.

Recall that we consider sequences $d = (s, t)$, where $s$ and $t$ have length $\ell$ and $n$, respectively, and where $\ell = n$ in the digraph case. Also $\mu = M_1(d)/2|A|$, where $|A| = (1 - \delta_{\ell_1})n\ell + \delta_{\ell_2}n(n - 1)$, $s = (\sum_{a \in S} s_a)/\ell$ and $t = (\sum_{v \in T} t_v)/n$. Noting that $S \cap T = \emptyset$, we set $\varepsilon_a = (s_v - s)/s$ for $a \in S$, $\varepsilon_v = (t_v - t)/t$ for $v \in T$, $\sigma_v^2 = \sigma^2(s)$, and $\sigma_v^2 = \sigma^2(t)$. Thus, $\varepsilon_a$ for example stands for the relative deviation of $s_a$ from the average degree in $S$, whereas, in the digraph case, $\varepsilon_v$ refers to the relative deviation of $t_a$ from the average degree in $T$. Then with these definitions, and with $*$ standing for either of $bi$ and $di$, we define the following "guesses" for the probability and ratio functions:

$$
P^* = \pi, \quad R^* = \rho,
$$

$$
Y^* = \pi \cdot \pi(\varepsilon_b, \varepsilon_v - 1/t, \varepsilon_{\bar{b}}, \varepsilon_{\bar{v}}) \cdot \left(1 + \frac{\mu(1 + \varepsilon_a) - \mu^2(1 + \varepsilon_a + \varepsilon_b)}{t(1 - \mu)}\right),
$$

so that $P^*_{a,b}$, and so forth are functions of degree sequences. For example, the guessed edge probability function for bipartite graphs with "classical" parameters $d_a$ and so forth is given by

$$
P_{a,v}^*(d) = P_{a,v}^*(d) = \frac{d_a d_v}{m} \left(1 - \frac{n\ell(d_a - s)(d_v - t)}{m(n\ell - m)} - \frac{m(n\ell - m)}{m(n\ell - m)} - \frac{n\ell d_a\sigma_v^2(t)n}{m(n\ell - m)} - \frac{(d_v - t)\sigma_v^2(s)\ell}{m(n\ell - m)}\right).
$$

for $a \in S, v \in T$. Note that this expression agrees with the one given in Theorem 1.4 for the bipartite case. One can compute similar expressions for $R_{a,v}^*$, the ratio functions $R_{a,b}^*$ and $R_{a,b}^*$, and the two-edge-path probabilities $Y_{a,v,b}^*$ and $Y_{a,v,b}^*$. In particular, we remark that $R_{a,b}^*(d)$ written with "classical" parameters is just the expression in (9). We will show that the functions $P^*, Y^*$, and $R^*$ approximate the actual probability and ratio functions sufficiently well. We comment on how to find these expressions in Section 4. The following implies that they are close to fixed points of the operators defined in (14)–(16).

**Lemma 3.1.** Let $n, \ell$ be integers and let $0 \leq \kappa < 1/10$. Let $A$ be as in either the bipartite or the digraph case, let $d = (s, t)$ be a sequence of length $\ell + n$, let $\mu = M_1(d)/2|A|$, and assume that $\mu < 1/4$. Furthermore, let $s$ and $t$ be the average of $s$ and $t$, respectively, and $\varepsilon = \frac{d - \min(s, t)}{s}$, and assume that $\max_{a \in S} |s_a - s|/s \leq \varepsilon$, $\max_{v \in T} |t_v - t|/t \leq \varepsilon$. Let $*$ stand for either of $bi$ or $di$. Then

(a) $R(P^*, Y^*)_{a,b}(d) = R_{a,b}^*(d) \left(1 + O\left(\mu \varepsilon^4\right)\right)$ for all $a, b \in S$,
(b) $P(P^*, R^*)_{a,v}(d) = P_{a,v}^*(d) \left(1 + O(\mu \varepsilon^4)\right)$ for all $a \in S, av \in A$
(c) $Y(P^*, Y^*)_{a,v,b}(d) = Y_{a,v,b}^*(d) \left(1 + O(\mu \varepsilon^4)\right)$ for all distinct $a, b \in S, v \in T, av, vb \in A$. 


Proof. To make comparisons with the proof of [14, Lemma 7.1] easy, we follow that proof very closely, with modifications due to the bipartite setting. Details of routine calculations, for which we often used Maple, are omitted.

We first present some convenient approximations of $P^*$ and $\pi$ for use when their parameters have been slightly altered. Let $d' = (s', t')$ be a sequence where $s'$ and $t'$ are of length $\ell$ and $n$, respectively, such that $d'$ is at $L_1$ distance $O(1)$ from $d$, and with $d'_{i} = d_a_{i} - j_{a_{i}}$ and $d'_{i} = d_v_{i} - j_{v_{i}}$ for some $a \in S, v \in T$. Here and in the following, the bare symbols $\mu, \epsilon_a$ and so on are defined with respect to the original sequence $d$, whilst $\mu', \epsilon'_a$ and so forth are defined with respect to $d'$. For such a sequence $d'$ we have that $\mu(d') = \lambda' = \mu + O(1/n\ell)$ by definition of $\mu(d')$. Therefore, the variables $\epsilon_a$ and $\epsilon_v$ change to $\epsilon_a' = \epsilon_a - j_{a}/s + O(1/\mu n\ell)$ and $\epsilon_v' = \epsilon_v - j_{v}/t + O(1/\mu n\ell)$. (Note that this takes into account that $\epsilon_a$ and $\epsilon_v$ are defined with respect to $s(d')$ and $t(d')$.) Furthermore,

$$\sigma^2(s') - \sigma^2(s) = O((\max_{c \in S}|s_c - s| + 1)/n) = O(\epsilon s/n),$$

and similarly, for $\sigma^2(t')$. Hence, by definition of $P^* = \pi$ and the preceding considerations,

$$P^*_{a,v}(d') = \pi(\epsilon_a - j_a/s, \epsilon_v - j_v/t) + O(1/\mu n\ell).$$

That is, the changes from $\mu, s, t, \sigma^2$ and $\sigma_T^2$ to $\mu', s', t', (\sigma^2)'$ and $(\sigma_T^2)'$ are negligible in the formula for $P^*$.

For (a), we note first that $R(P^*, Y^*)_{a,a}(d) = 1 = R_{a,a}(d)$ by definition of $\rho$ and $R$ in (16). Assume now that $a \neq b$. Using (16) to evaluate $R(P^*, Y^*)_{a,b}(d)$, we estimate the expression $\text{bad}(a, b, d - e_b) = \text{bad}(P^*, Y^*)(a, b, d - e_b)$ for which, in turn, we need to estimate $\sum Y^*_{a,v,b}(d - e_b)$, where the sum is over all $v \in T$ such that both $av$ and $bv$ are allowable (see (13)). By definition and (17),

$$Y^*_{a,v,b}(d - e_b) = \pi \cdot \pi(\epsilon_b - 1/s, \epsilon_v - 1/t) \cdot \left(1 + \frac{\mu(1 + \epsilon_a) - \mu^2(1 + \epsilon_a + \epsilon_b)}{t(1 - \mu)} + O(1/\mu n\ell)\right),$$

where we use $\epsilon_a$, $\epsilon_b$, and $\mu$ in the third factor (rather than the altered versions $\epsilon'_a$ etc.) using the same reasoning as for obtaining (17). Consider expanding this expression for $Y^*_{a,v,b}(d - e_b)$ ignoring terms of order $\epsilon^4$, and hence also ignoring those of order $\epsilon^2$, $\epsilon^3$, $\epsilon^4$, $1/\ell^2 = \max(1/t, 1/s)$. A convenient way to do this is to make substitutions $\epsilon_y = y_1 \epsilon_y$, $1/t = y_2^2/t$, $\mu = y_2 \mu, 1/n = y_1^2 y_2/n$, and so on where $y_1$ represents a parameter of size $O(\epsilon)$ and $y_2$ one of size $O(\mu)$ (for instance, $\sigma_T^2/t\ell^2$ is $O(\epsilon^2 \mu)$), and then expand about $y_1 = 0$ and drop terms of order $y_1$. Since $Y^*_{a,v,b}(d - e_b)$ gains a factor $y_2^2$ via the factors of $\mu$ in $\pi$, each term containing $y_1^2$ is of order $y_1^2 y_2^2$ and is hence $O(\mu^2 \epsilon^2)$. Next, removing the “sizing” variables $y_1$ by setting them equal to 1, and then expanding the result about $(\epsilon_\gamma, \epsilon_v) = (0, 0)$ and retaining all terms of total degree at most 3 in $\epsilon_\gamma$ and $\epsilon_v$, we get

$$Y^*_{a,v,b}(d - e_b) = c_0 + c_{10} \epsilon_v + c_{01} \epsilon_\gamma + c_{20} \epsilon_\gamma^2 + O(\mu^2 \epsilon^4),$$

where the functions $c_0, c_{10}, c_{01},$ and $c_{20}$ are independent of $\epsilon_v$ and $\epsilon_\gamma$, with $c_0, c_{10}$ and $c_{01}$ linear in $1/s$ and $1/t$, and $c_{20}$ constant in those variables. (By calculation, the third order terms all turn out to be absorbed by the error term. Furthermore, the relative error $1/\mu \ell n$ in the previous expression for $Y^*$ yields an absolute error $O(\mu/\ell n) = O(\mu^2 \epsilon^4)$ since $Y^*$ is $O(\mu^2)$.) We note that $c_{01}$ has a factor $\sigma^2$ since $\epsilon_\gamma$ in $\pi$ has such a factor.
Then considering the definition of \( \text{bad}(a, b, d - e_b) \) in (13) we find that the second summation can be written as

\[
\Sigma_{\text{bad}} := \sum_{v \in \mathcal{A}(a) \cap \mathcal{A}(b)} Y^s_{a,v,b}(d - e_b)
\]

\[
= \sum_{v \in \mathcal{A}(a) \cap \mathcal{A}(b)} (c_0 + c_{10} \epsilon_v + c_{01} \epsilon_{\bar{v}} + c_{20} \epsilon_v^2 + O(\mu^2 \epsilon^4)) + \delta^d \left( -2c_0 - c_{10}(\epsilon_\bar{v} + \epsilon_{\bar{b}}) - c_{01}(\epsilon_a + \epsilon_{\bar{b}}) - c_{20}(\epsilon_a^2 + \epsilon_{\bar{b}}^2) \right) + O(n\mu^2 \epsilon^4),
\]

since \( a \) and \( b \) are distinct elements of \( S \) (in which case \( \mathcal{A}(a) \cap \mathcal{A}(b) \) is \( T \) in the bipartite case and is \( T \setminus \{\bar{a}, \bar{b}\} \) in the digraph case), and where we also use that \( \sum_{v \in T} \epsilon_v = 0 \) and that, in the digraph case, \( \sum_{v \in T} \epsilon_{\bar{v}} = \sum_{a \in S} \epsilon_a = 0 \). Noting that \( \mathcal{A}(a) \setminus \mathcal{A}(b) \) is \( \emptyset \) in the bipartite case and consists of just \( \bar{b} \) in \( T \) in the digraph case, we can write \( \text{bad}(a, b, d - e_b) \) in the numerator of (16), by using (13), as

\[
\text{bad}(a, b, d - e_b) = \frac{1}{d_a} (\Sigma_{\text{bad}} + \delta^d \pi(\epsilon_\bar{a}, \epsilon_{\bar{b}}, \epsilon_{\bar{a}}, \epsilon_{\bar{b}} - 1/t) + O(1/\ell n)),
\]

where \( d_a = s_a = (1 + \epsilon_a)s \), and the \( O(1/n\ell) \) term captures the fact that we use \( \mu = \mu(d), \sigma^2(s) \), and \( \sigma^2(t) \), respectively, instead of \( \mu(d'), \sigma^2(s'), \) and \( \sigma^2(t') \), in the formula for \( \pi \) applied to an altered sequence \( d' \). The error term \( O(1/\ell n) \) here, together with the one from \( \Sigma_{\text{bad}} \) above, produce an additive error in \( \text{bad}(a, b, d - e_b) \) of \( O(\mu \epsilon^4) \) since \( n/d_a = n/s_a \sim 1/\mu \) and \( 1/\ell n < \mu \epsilon^4 \). Substituting the above expression, stripped of its error terms, into

\[
\frac{R(P^*, Y^s)_{a,b}(d)}{\rho} - 1 = \frac{1}{\rho} \cdot \frac{(1 + \epsilon_a)(1 - \text{bad}(a, b, d - e_b))}{(1 + \epsilon_{\bar{b}})(1 - \text{bad}(b, a, d - e_b))} - 1,
\]

and simplifying gives a rational function \( \tilde{F} \) which satisfies \( R(P^*, Y^s)_{a,b}(d)/\rho = 1 + \tilde{F} + O(\mu \epsilon^4) \). After inserting the size variables \( y_1 \) and \( y_2 \) into \( \tilde{F} \) as specified above (and here it may be convenient to use \( n = \delta^{d_i} + s/\mu \)), and simplifying, we find that \( \tilde{F} \) has \( y_2 \) as a factor (of multiplicity 1), and its denominator is non-zero at \( y_1 = 0 \). Then expanding the expression in powers of \( y_1 \) shows that \( \tilde{F} = O(y_1^2) \). Along with the extra factor \( y_2 \), this implies \( \tilde{F} = O(\mu \epsilon^4) \). Part (a) follows.

To prove part (b) let \( av \in \mathcal{A} \) with \( a \in S \) and let \( d' \) be the sequence \( d - e_v \). Note that, analogous to (17), the differences in the values of \( \mu, s, t \) in \( \rho \) between \( d \) and \( d' \) are negligible, as are the differences in \( \epsilon_a \) and \( \epsilon_{\bar{b}} \) for \( b \in \mathcal{A}(v) = S \setminus \{\bar{v}\} \) (note that with these assumptions \( v \) is never equal to \( a, \bar{a}, b, \) or \( \bar{b} \); so the values of \( \epsilon_a, \epsilon_{\bar{b}}, \epsilon_{\bar{a}}, \epsilon_{\bar{b}} \) only change since \( s \) changes). Hence, we also have

\[
R^*_{a,b}(d') = \rho \cdot (1 + O(1/\mu n \ell)),
\]

for \( a, b \in S, v \in \mathcal{A}(a) \cap \mathcal{A}(b) \). Therefore, by definition (14),

\[
P(P^*, R^s)_{a,v}(d)
\]

\[
= d_v \left( \sum_{b \in \mathcal{A}(v)} R^s_{b,a}(d - e_v) \frac{1 - P^s_{b,v}(d - e_b - e_v)}{1 - P^s_{a,v}(d - e_a - e_v)} \right)^{-1}
\]

\[
= d_v \left( \sum_{b \in \mathcal{A}(v)} \rho(\epsilon_b, \epsilon_a, \epsilon_{\bar{b}}, \epsilon_{\bar{a}}) \cdot \frac{1 - \pi(\epsilon_b - 1/s, \epsilon_v - 1/t, \epsilon_{\bar{b}}, \epsilon_{\bar{a}})}{1 - \pi(\epsilon_a - 1/s, \epsilon_v - 1/t, \epsilon_{\bar{a}}, \epsilon_{\bar{b}})} \right)^{-1} \left(1 + O\left(1/\mu n \ell\right)\right). \quad (18)
\]
By expanding in powers of $\varepsilon_b$ and $\varepsilon_b^*$ we obtain

$$
\rho(\varepsilon_b, \varepsilon_a, \varepsilon_b^*, \varepsilon_a^*) \cdot \frac{1 - \pi(\varepsilon_b - 1/s, \varepsilon_a - 1/t, \varepsilon_b^*, \varepsilon_a^*)}{1 - \pi(\varepsilon_a - 1/s, \varepsilon_a - 1/t, \varepsilon_b^*, \varepsilon_a^*)} = K + O(\varepsilon^4),
$$

where $K$ is a polynomial in $\varepsilon_b$ and $\varepsilon_b^*$ of total degree at most 3 in $\{\varepsilon_b, \varepsilon_b^*\}$. Calculations using the size variables $y_1$ and $y_2$ as for (a) show that $K = k_{00} + k_{10} \varepsilon_b + k_{01} \varepsilon_b^* + k_{20} \varepsilon_b^2 + O(\varepsilon^4)$ for some $k_{ij}$ (and in particular, the coefficients $k_{11}$ and $k_{02}$, and those for terms of third order, are all absorbed by the error terms). We also note from the definition of $\pi$ and $\rho$ that $k_{01}$ has $\delta_{d^i}$ as a factor. Also, for $v \in T$ we have $A(v) = S$ in the bipartite case and $A(v) = S \setminus \{\vec{v}\}$ in the digraph cases. Then the main summation over $b$ in (18) can be evaluated as

$$
\ell \cdot k_{00} + \ell \sigma^2 k_{20}/s^2 + \delta_{d^i} (-k_{00} - k_{10} \varepsilon_b^* - k_{01} \varepsilon_a - k_{20} \varepsilon_b^2),
$$

with relative error $O(\varepsilon^4)$, noting that $K$ has constant order, where we use that $\sum_{b \in S} \varepsilon_b = 0$ and that $\sum_{b \in S} \varepsilon_b^* = \ell \sigma^2/s^2$. Using the size variables $y_1$ and $y_2$ as described above, we then find that $P(P^*, R^*)_{a,v}(d) = \pi(1 + O(\mu \varepsilon^4))$, with the extra factor $\mu$ arising in the error term in the same way as for $R$ in part (a). Part (b) follows.

Part (c) is more straightforward than the first two parts and is easily verified by similar considerations, so we omit details. \hfill \blacksquare

**Proof of Theorem 1.1.** Our aim is to apply Lemma 2.1 to suitably defined probability spaces $S$ and $S'$.

Let $\mu_0, \kappa, \ell$, $n, m$, $D$ be as in the theorem statement. Note that all sequences $d$ in $D$ have the same values $\mu = \mu(d) = m/n(\ell - \delta_d)$, $s = s(d) = m/\ell$ and $t = t(d) = m/n$. All other sequences $d'$ in this proof will be close enough to $D$ (in Hamming distance) so that $\mu' := \mu(d') \sim \mu$. Let $\varepsilon = \max\{s^{-1/2}, r^{-1/2}\}$ (in accordance with Lemma 3.1, since $\kappa - 1/2 < 0$) and note that $\varepsilon \geq \max\{1/\sqrt{s}, 1/\sqrt{t}\}$ since $\kappa > 0$. By the assumptions on $D$, the variable $\varepsilon$ is an upper bound on the relative deviation of degrees from the average degree on a given side for sequences $d \in D$.

We first consider the ratios $R_{a,b}$ for $a, b \in S$. Recall that $P, Y$ and $R$ denote the actual edge probability, two-edge-path probability and ratio functions (cf. (7) and (8)) and that $P^*, Y^*$, and $R^*$ are functions of degree sequences with closed form given at the beginning of this section. Recall that $P, Y$, and $R$ are always defined with respect to some underlying set $A$ which is either $A^{bi}$ or $A^{di}$ here. The next claim states that the functions $P^*$ and $R^*$ approximate $P$ and $R$ sufficiently well on $D$. An analogous statement for $Y^*$ and $Y$ appears in the proof of the claim but is not needed elsewhere. It is easy to see inductively that we only require $R_{a,b}(d)$ when $a$ and $b$ are in the same set of the bipartition. Let $Q_1^S$ denote the set of sequences $d = (s, t) \in \mathbb{Z}_{\geq 0}^{\leq n}$ such that $d - e_v \in D$ for some $a \in S$; and let $Q_1^T$ denote the set of sequences $d = (s, t) \in \mathbb{Z}_{\geq 0}^{\leq n}$ such that $d - e_v \in D$ for some $v \in T$.

**Claim 3.2.** Let $\ast$ be either bi or di. For $d \in D$, $av \in A$,

$$
P_{av}(d) = P^*_{av}(d)(1 + O(\mu \varepsilon^4)), \quad (19)
$$

and uniformly for all $d \in Q_1^S$ and all $a, b \in S$

$$
R_{a,b}(d) = R^*_{a,b}(d)(1 + O(\mu \varepsilon^4)). \quad (20)
$$
By symmetry, (20) also holds for all \( d \in Q_1^T \) and all distinct \( a, b \in T \). The proof is very similar to the proof of Claim 6.4 in [14] with some adaptations to the bipartite setting. We include a full proof for the sake of completeness.

**Proof of the claim.** To show that \( P \) and \( P^* \) (and \( R \) and \( R^* \)) are \((\mu e^4)\)-close in the sense of (19) and (20), we define the compositional operator

\[
C(p, y) = (\hat{p}, \mathcal{Y}(\hat{p}, y)), \quad \text{where } \hat{p} = P(p, R(p, y)).
\]

We first observe that Proposition 2.6 implies that \( C \) fixes \((P, Y)\), where in this context we regard \( P \) to be the function \( p \) with \( p_{a,v} = p_{av} \) for all \( av \in A \), and similarly \( Y \) to be \( y \) with \( y_{a,v,b} = Y_{a,v,b} \). We will deduce a certain contraction property of \( C \) by applying Lemma 2.8(a)–(c) one after the other, and then show that for any integer \( k > 0 \), \( C^n(P^*, Y^*) \) and \( C^n(P, Y) \) are \( O(\mu^k) \)-close. We will also show that \( P^* \) and \( C^n(P^*) \) are \( O(\mu e^4) \)-close. These observations will then be shown to imply Claim 3.2.

Fix \( k_0 = 4 \log n \) and \( r = 4k_0 + 4 = O(\log n) \). Let \( \Omega(0) \) be the set of sequences \( d' \in \mathbb{Z}_{\geq 0}^n \) that are at \( L_1 \) distance at most \( r \) from a sequence in \( \Box \). Let \( \mu_1 = 5\mu \), and define \( \Omega(0) \) to be the set of sequences \( d' \in \Omega(0) \) of \( L_1 \) distance at least \( s + 1 \) from all sequences outside \( \Omega(0) \).

Towards applying Lemma 2.8 we first establish that \((P, Y)\) and \((P^*, Y^*)\) are elements of \( \Pi_{\mu_1}(\Omega(2)) \) (see Definition 2.7). Note that for \( d' \in \Omega(0) \), the values of \( s(d') \), \( t(d') \), and \( \mu(d') \) are asymptotically equal to \( s \), \( t \), and \( \mu \), respectively, since \( M_1(d') = M_1(d_0) + O(\log n) \) for some sequence \( d_0 \in \Box \). Thus, \( \mu \) and \( \mu(d') \) are interchangeable in the error terms below. Furthermore, we note that the bounds on \( s_a \) and \( t_v \) in the theorem statement imply that for all matchable \( d \in \Omega(0) \), \( s_a \sim s \) and \( t_v \sim t \) uniformly for all \( a \in S \) and \( v \in T \). By Lemma 2.5 we obtain \( N(d) > 0 \) for all matchable \( d \in \Omega(0) \) in both cases \( A^i \) and \( A^i \). In doing so, we may take \( C = 1 \) and \( F \) to be either empty in the bipartite case, or a matching in the digraph case; the condition \( m \leq \ell n/9 \) follows from choosing \( \mu_0 \) small enough, and the conditions on \( \Delta(s) \) and \( \Delta(t) \) can be seen to follow from \( |s_a - s| \leq s^{1/2+k} \) for all \( a \in S \), \( |t_v - t| \leq t^{1/2+k} \) for all \( v \in T \) and \( k < 1/2 \), say. After this, for \( n \) and \( \ell \) sufficiently large, Lemma 2.3, together with the facts that \( s_a \sim s \) and \( t_v \sim t \) uniformly for all \( a \in S \), \( v \in T \), implies that for all \( d \in A \)

\[
P_{av}(d) \leq \frac{\mu}{1 - 2\mu} (1 + o(1)) < \frac{5\mu}{4} \quad \text{for all matchable } d \in \Omega(0), \tag{21}
\]

where for the last inequality we use that \( \mu < \mu_0 < 1/11 \), say. Since \( \Omega(2) \subseteq \Omega(0) \) and \( 5\mu/4 < \mu_1 \), this establishes requirement (IIa) for \( P \) in the definition of \( \Pi_{\mu_1}(\Omega(2)) \). Now restrict slightly to \( d \in \Omega(2) \).

By definition, \( Y_{a,v,b}(d) \) is the probability that both edges \( av \) and \( bv \) are present. Hence Proposition 2.6(c) implies (with the above bounds on \( P_{av}(d) \) applying for all \( d \in \Omega(0) \)) that \( 0 \leq Y_{a,v,b}(d) = Y_{b,v,a}(d) \leq 3P_{bv}(d)/2 \) (easily) assuming, as we may, that \( \mu \) is sufficiently small. Thus \((P, Y)\) satisfies condition (IIc) for membership of \( \Pi_{\mu_1}(\Omega(2)) \), and also

\[
\sum_{a \in A(a) \cap A(b)} Y_{a,v,b}(d) \leq \sum_{a \in A(a) \cap A(b)} \frac{3P_{bv}(d)\mu}{2} \leq 3\mu s_a(1 + o(1)),
\]

for all distinct \( a, b \in S \), using \( P_{bv}(d) \leq 2\mu \sim 2s_a/n \) which follows from (21) and recalling that \( s_a \sim \mu n \) uniformly for all \( a \in S \) for all \( d \in \Omega(0) \). As \( 4\mu < \mu_1 \), this shows \( Y \) satisfies condition (IIb) for membership of \( \Pi_{\mu_1}(\Omega(2)) \) when \( n \) is sufficiently large and \( a, b \in S \) are distinct. The equivalent statement for both \( a, b \in T \) follows analogously. Note that this covers all cases of (IIb) as otherwise \( A(a) \cap A(b) = \emptyset \) and the statement is trivial. To see that \((P^*, Y^*)\) is also in \( \Pi_{\mu_1}(\Omega(2)) \), we first observe
that $P_{a,v}^*(d) = \pi(\epsilon_a, \epsilon_v) \sim \mu$ for all $av \in A$ and all $d \in \Omega^{(0)}$. This is because $\epsilon \to 0$ since $\kappa < 1/2$ and both $s, t \to \infty$, and because $\sigma_\epsilon^2/t\epsilon + \sigma_\kappa^2/s\kappa = O(\epsilon^2\mu)$. Properties (Pi)–(Pc) follow directly from this fact and the definition of $Y^*$ since $\mu = \mu_1/5$.

Now for large $n$ and $\ell$ and for $(a, v) \in \tilde{A}_2$ we have $P_{a,v}(d) = P_{a,v}^*(d)(1 \pm 1)$ for all matchable $d \in \Omega^{(0)}$ since $P_{a,v}(d) \sim \mu$ and by (21). Also, $0 \leq Y_{a,v,b}(d) \leq 3\mu P_{b,v}(d)/2$ implies $Y_{a,v,b}(d) = Y_{a,v,b}^*(d)(1 \pm 1)$ for matchable $d \in \Omega^{(2)}$. We may now apply Lemma 2.8(a) with $\xi = 1$ for any $S$-heavy $d \in \Omega^{(3)}$ to deduce that

$$R(P, Y)_{a,b}(d) = R(P^*, Y^*)_{a,b}(d)(1 + O(\mu_1)),$$

for all $a, b \in S$ (noting that for $a = b$ the claim is trivial, and for distinct $a, b$ we use the previous observations). Writing $r$ for $R(P, Y)$ and $r'$ for $R(P^*, Y^*)$ we obtain from this and Lemma 2.8(b) that

$$P(P, r)_{a,v}(d) = P(P^*, r')_{a,v}(d)(1 + O(\mu_1)),$$

for all matchable $d \in \Omega^{(4)}$ and all $av \in A$. Next applying Lemma 2.8(c) in the same way to matchable $d \in \Omega^{(6)}$ gives

$$\mathcal{Y}(\hat{p}, Y)_{a,v,b}(d) = \mathcal{Y}(\hat{p}', Y^*)_{a,v,b}(d)(1 + O(\mu_1)),$$

for all matchable $d \in \Omega^{(6)}$, and all $(a, v, b) \in \tilde{A}_2$ with $a \in S$, where $\hat{p} = P(P, r)$ and $\hat{p}' = P(P^*, r')$. Recalling the definition of $C(P, y)$ from the beginning of this proof, we note that the three conclusions above imply that, for $C(P, Y) = (P_1, Y_1)$ and $C(P^*, Y^*) = (P^*_1, Y^*_1)$, we have $P^*_k(d) = P_1(d)(1 + O(\mu_1))$ and $Y^*_k(d) = Y_1(d)(1 + O(\mu_1))$ for all $d \in \Omega^{(6)}$. Similarly, making $k - 1$ iterated applications of the three parts of Lemma 2.8 with ever-decreasing $\xi$ produces

$$P_k^*(d) = P_k(d)(1 + O(\mu_1)^k) \quad \text{and} \quad Y_k^*(d) = Y_k(d)(1 + O(\mu_1)^k),$$

for all $d \in \Omega^{(2k+2)}$, where $P_k$, $P_k^*$ and so forth are defined analogously for $C_k$. In the same way, applying Lemma 3.1(a)–(c) in turn, recalling Lemma 2.8 to handle the small error terms, shows that $P_k^*(d) = P^*(d)(1 + O(\mu\xi^4))$ and $Y_k^*(d) = Y^*(d)(1 + O(\mu\xi^4))$ for all $d \in \Omega^{(6)}$. Using the three parts of Lemma 2.8 repeatedly, and bounding the total distance moved during the iterations as for a contraction mapping (as the sum of a geometric series), this gives

$$P_k^*(d) = P^*(d)(1 + O(\mu\xi^4)) \quad \text{and} \quad Y_k^*(d) = Y^*(d)(1 + O(\mu\xi^4)),$$

for all $d \in \Omega^{(2k+2)}$. Using the last two conclusions with $k := k_0 = 4\log n$ and the fact that $P_k = P$ (as $C$ fixes $(P, Y)$) gives (19) for all matchable $d \in \Omega^{(r-2)}$ since we may assume that the $O(\mu_1)$ term is at most $1/\epsilon$ say. (Recall that $\mu_1 = 5\mu < 5\mu_0$ which we may choose to be sufficiently small.) Note that $\Phi \subseteq \Omega^{(r)} \subseteq \Omega^{(r-2)}$ by definition. For (20), we now use that (19) holds for all matchable $d \in \Omega^{(r-2)}$ to deduce from Lemma 2.8(a) that $R(P, Y)_{a,b}(d) = R(P^*, Y^*)_{a,b}(d)(1 + O(\mu\xi^4))$ for all $S$-heavy $d \in \Omega^{(r-1)}$. This, together with (a) above and the fact that $R(P, Y) = R$, implies (20) for all $d \in \Omega^{r-1} \subseteq \Omega^{(r-1)}$.

Claim 3.2 shows that $R^*$ is a precise enough approximation of the “true” ratio $R$. This approximation will be used in conjunction with Lemma 2.1 where this ratio makes its appearance in the form of the expression $P_{S'}(d)/P_{S'}(d')$. Let us now make suitable definitions of probability spaces $S$ and $S'$ in preparation for using Lemma 2.1. Let $\Omega$ be the underlying set of $B_m(\ell', n)$ in the bipartite case and
of $\tilde{B}_m(n)$ in the digraph case, that is, the set of all degree sequences $d = (s, t) \in \mathbb{Z}_{\geq 0}^{2n}$ such that $M_1(s) = M_1(t) = m$ (and $\ell' = n$ in the digraph case). Let $S' = D(G(\ell', n, m))$ in the bipartite case, and $S' = D(\tilde{G}(n, m))$ in the digraph case. Set

$$\mathfrak{B} = \{ d = (s, t) \in \mathfrak{D} : \sigma^2(s) \leq 2s, \sigma^2(t) \leq 2t, |d_{ij}(s, t)| < \xi s \},$$

where $\xi = \max\{\log^2 \ell' / \sqrt{\ell'} , \log^2 n / \sqrt{n}\}$. Define the graph $G$ (to be used in Lemma 2.1) on vertex set $\mathfrak{D}$ by joining two degree sequences by an edge if they are of the form $d - e_a$ and $d - e_b$ for some $d \in Q^\ell_1$ and $a, b \in S$, or for some $d \in Q^\ell_2$ and $a, b \in T$. By virtue of the definition of $G$, any two such degree sequences are adjacent.

We note at this point that the diameter of $G$ is

$$r = O(\ell' s^{1/2+\kappa} + m^{1/2+\kappa} ) = O(\varepsilon mn \ell'),$$

since the constant sequence $(d, \ldots, d)$ is an element of $\mathfrak{D}$ and by the degree constraints for $d \in \mathfrak{D}$. The same bound holds for the diameter of $G[\mathfrak{B}]$. By definition of $R$ in (7) and its approximation in (20) we have, for adjacent vertices/sequences in this graph, that

$$\frac{\mathbb{P}_{S'}(d - e_a)}{\mathbb{P}_{S'}(d - e_b)} = R_{ab}(d) = R^{S'}_{ab}(d) \left( 1 + O(\mu \varepsilon^4) \right).$$

For the setup in Lemma 2.1, it remains to define the probability space $S$, which can be thought of as the ideal probability space by which we approximate the desired space $S'$. Recall the definition of $\tilde{H}(d)$ in (2), which is slightly different in the bipartite and the digraph case (due to $\mu$ being defined slightly differently and the extra term in the exponent in the digraph case). Also recall (from just before (9)) that we use $H(d)$ to denote the product $\mathbb{P}_{\tilde{B}}(d) \tilde{H}(d)$ on the right hand side of (5), where $B = B_m(\ell', n)$ in the bipartite case and $B = \tilde{B}_m(n)$ in the digraph case. Then set

$$\mathbb{P}_S(d) := H(d) / \sum_{d' \in \mathfrak{B}} H(d') = \frac{H(d)}{\mathbb{E}_B(\tilde{H})},$$

for $d \in \mathfrak{B}$ and $\mathbb{P}_S(d) = 0$ otherwise.

In order to use Lemma 2.1, we need to show that $\mathbb{P}(\mathfrak{B})$ is roughly equal to $1$ in the two probability spaces $S$ and $S'$ in both cases, bipartite and digraph. First note that $\mathbb{P}_S(\mathfrak{B}) = 1$ by definition. Towards addressing the claim for $S'$, we simultaneously evaluate $\mathbb{P}_B(\mathfrak{B})$ and $\mathbb{E}_B(\tilde{H})$ for later use. We note that $M_1(s) = M_1(t) = m$ for all $d = (s, t) \in \Omega$, by definition. Let $n = \min\{n, \ell'\}$. For the following, let $d$ be chosen in either of $D(G) = S'$ and $B$, in either of the bipartite and the digraph case. Then $|s_a - s| \leq s^{1/2+\kappa}$ and $|t_v - t| \leq t^{1/2+\kappa}$ for all $a \in S$ and $v \in T$ with probability at least $1 - O(\tilde{n}^{-\omega})$ by Lemma 2.10(i) and since $s > (\log \ell')^K$ and $t > (\log n)^K$ for all $K > 0$. It follows that $\mathbb{P}_{D(G)}(\mathfrak{D}) = \mathbb{P}_{S'}(\mathfrak{D}) = 1 - O(\tilde{n}^{-\omega})$ and $\mathbb{P}_B(\mathfrak{D}) = 1 - O(\tilde{n}^{-\omega})$ in both $S'$ and $B$. Thus $\mathbb{P}(\mathfrak{B}) \geq 1 - \varepsilon_0$ in both $S$ and $S'$ in both cases, bipartite and digraph, for, say, $\varepsilon_0 = 1/\bar{n}$.

Before applying Lemma 2.1, let us use these concentration results straight away to estimate $\mathbb{E}_B(\tilde{H})$. If $\sigma^2(t) = t(1 - \mu)(1 + O(\xi))$ then the term $\sigma^2(t) / t(1 - \mu)$ in the exponent of $\tilde{H}(d)$ is $1 + O(\xi)$. 
Similarly for the term $\sigma^2(s)/s(1-\mu)$. Furthermore, in the digraph case, if $\sigma(s,t) = O(\xi s)$ then the term $\delta^d \sigma(s,t)/s(1-\mu)$ in $\tilde{H}(d)$ is $O(\xi)$. It follows, using the strong concentration results in the previous paragraph, that

$$\tilde{H}(d) = 1 + O(\xi)$$

with probability $1 - O(n^{-\omega})$ for $d \in B$. \hspace{1cm} (25)

In addition, by definition of $\mathfrak{B}$ it follows that $\tilde{H}(d) = \Theta(1)$ for all $d \in \mathfrak{B}$, using the fact that $\mu < 1/2$, say. This and (25) then imply that $E_B(1_{\mathfrak{B}} \tilde{H}) = 1 + O(\xi + n^{-\omega}) = 1 + O(\xi)$. The error term $n^{-\omega}$ can be dropped; it is absorbed by $\xi$ because (3) implies $n = o(\ell^{3/2}) = o(n^{3/2})$.

We now move to the last step of the proof, which is applying Lemma 2.1. For $d = (s,t) \in Q^S_1$ and $a,b \in S$,

$$\frac{H(d-e_a)}{H(d-e_b)} = \frac{H(s-e_a,t)}{H(s-e_b,t)} = \frac{s_a(n + \delta^{bi} - s_b)}{s_b(n + \delta^{bi} - s_a)} \exp\left(\frac{s_b - s_a}{s(1-\mu)\ell} \left(1 - \frac{\sigma^2(t)}{t(1-\mu)}\right) + \frac{\delta^{di}(t_a - t_b)}{t(1-\mu)n}\right),$$

by (9). Compare the expression on the right hand side with $R^\ast = \rho$ at the beginning of this section (after a straight-forward reparameterization) to see that

$$\frac{H(d-e_a)}{H(d-e_b)} = R^\ast_{ab}(d) \left(1 + O\left(1/n^{2}\right)\right),$$

where the error is due to the fact that we use $e^x = 1 + x + O(x^2)$ and that $\mu(d) = \mu + O(1/n\ell)$ for $d \in Q^S_1$. This together with (24) gives

$$\frac{\mathbb{P}_{S'}(d-e_a)}{\mathbb{P}_{S'}(d-e_b)} = e^{O(\mu\ell^4)} \frac{H(d-e_a)}{H(d-e_b)},$$

for $d \in Q^S_1$ and $a,b \in S$. The same applies for $d \in Q^T_1$ and $a,b \in T$ by symmetry. Note that when both $d-e_a$ and $d-e_b$ are elements of $\mathfrak{B}$, the right hand side is in fact equal to $e^{O(\mu\ell^4)}\mathbb{P}_{S}(d-e_a)/\mathbb{P}_{S}(d-e_b)$, by definition of $\mathbb{P}_{S}$ above. Therefore, by Lemma 2.1

$$\mathbb{P}_{S'}(d) = \mathbb{P}_{S}(d) e^{O(r\mu^4 + \epsilon_0)}$$

$$= \mathbb{P}_{B}(d) \tilde{H}(d) \left(1 + O\left(\xi + r\mu\ell^4 + \epsilon_0\right)\right),$$

for all $d \in \mathfrak{B}$, where we use that $E_B(1_{\mathfrak{B}} \tilde{H}) = 1 + O(\xi)$. Now let $d \in \mathfrak{D} \setminus \mathfrak{B}$. Then there is some sequence $d' \in \mathfrak{B}$ such that the distance of $d$ and $d'$ in $G$ is at most $2r$. Any two adjacent sequences $\tilde{d} - e_a$ and $\tilde{d} - e_b$ along that path satisfy (28), so that by telescoping we see that

$$\mathbb{P}_{S'}(d) = \mathbb{P}_{B}(d) \tilde{H}(d) \left(1 + O\left(\xi + r\mu\ell^4 + \epsilon_0\right)\right),$$

for such $d$ as well. This proves the estimate (5) claimed in the theorem, since $\xi = \max\{\log^2\ell/\sqrt{\ell}, \log^2n/\sqrt{n}\}$, $\epsilon_0 = 1/n = o(\ell)$, and $r\mu\ell^4 = O(\xi^5\mu^2n\ell)$ from (23), where we note that $\epsilon = \max\{s^{x-1/2}, t^{x-1/2}\}$, $s = O(m/\ell) = O(n\mu)$, $t = O(m/n) = O(\ell\mu)$, and $\mu \sim \tilde{\mu}$. \hspace{1cm} \hfill \blacksquare

\textbf{Proof of Theorem 1.3.} Given $\ell$, $n$, and $m$, let $G$ denote either $\tilde{G}(n,m)$ or $G(\ell', n, m)$, and let $B$ denote either $\tilde{B}_m(n)$ or $B_m(\ell', n)$, as the case may be. We first claim that if the triple $(\ell', n, m)$ satisfies the
hypotheses of Theorem 1.1 for some \( \kappa > 0 \), then the assertion of Theorem 1.3 holds, that is, \( D(G) \) and \( B \) are a.q.e.

Let \( \tilde{n} = \min\{\ell', n\} \). In the proof of Theorem 1.1, we defined a set \( \mathfrak{B} \subseteq \mathfrak{D} \) in (22) and showed that \( \mathbb{P}_{S'}(\mathfrak{B}) = 1 - o(\tilde{n}^{-\omega}) \) just before (25), where \( S' = D(G) \), and also that \( \mathbb{P}_B(\mathfrak{B}) = 1 - o(\tilde{n}^{-\omega}) \). Theorem 1.1 gives (5) for all \( \mathfrak{d} \in \mathfrak{B} \). Since \( \xi \to 0 \) we have \( \tilde{H}(\mathfrak{d}) \sim 1 \) with probability \( 1 - o(\tilde{n}^{-\omega}) \) for \( \mathfrak{d} \in B \), by (25). The latter was proved using just the concentration results on \( \sigma^2(s), \sigma^2(t) \), and \( \sigma(s, t) \), for which Lemma 2.10 was used. This lemma equally applies to sequences \( \mathfrak{d} = (s, t) \in D(G) \), and so essentially the same proof shows that \( \tilde{H}(\mathfrak{d}) \sim 1 \) with probability \( 1 - o(\tilde{n}^{-\omega}) \) for \( \mathfrak{d} \in D(G) \) as well. Focusing on \( \mathfrak{B} \), which has probability \( 1 - o(\tilde{n}^{-\omega}) \) in each of the four probability spaces under consideration, we deduce that \( \mathbb{P}_{D(G)}(\mathfrak{d}) \sim \mathbb{P}_B(\mathfrak{d}) \) uniformly for \( \mathfrak{d} \) on an event whose probability is \( 1 - o(\tilde{n}^{-\omega}) \) in both \( D(G) \) and \( B \). Noting as above that \( n = o(\tilde{n}^{3/2}) \), this implies the required a.q.e. property, for any triple \((\ell', n, m)\) satisfying the hypotheses of Theorem 1.1.

In the bipartite case (b), we claim that the conditions on \( \ell', m, n \) in (ii) imply that the hypotheses for Theorem 1.1 are satisfied for some \( 0 < \kappa < 1/10 \). We may assume without loss of generality that \( \ell' \leq n \). This and the assumption \( n^3 = o(\ell^2 m^{1-\varepsilon}) \) imply \( n = o(m^{1-\varepsilon}) \) and then also \( m^t = o(\ell') \), so \( m \) and \( \ell' \) tend to \( \infty \) with \( n \). Then \( n = o(m^{1-\varepsilon}) \) also gives, for fixed \( K > 0 \),

\[
\log^K \ell' = o(\ell^2) = o(m^t) = O(\tilde{\mu} \ell n/m^{1-\varepsilon}) = o(\tilde{\mu} \ell'),
\]

which gives (4) (noting that \( \ell' \leq n \)). On the other hand, \( n^3 = o(\ell^2 m^{1-\varepsilon}) \) implies

\[
n^{3-10 \kappa} = o(n^3) = o(\ell^2 m^{1-10 \kappa})
\]

with \( \kappa = \varepsilon /10 \), and rearranging and taking the square root of both sides gives \( n/\ell^3/2-5k = o(\tilde{\mu}^{1/2-5k}) \). Once again using \( \ell' \leq n \), this implies (3), and thus the hypotheses of Theorem 1.1 are satisfied. Thus, part (b)(ii) of Theorem 1.3 holds by the claim in the first paragraph of this proof.

Setting \( \ell' = n \) in the preceding analysis shows that the conditions of Theorem 1.1 are satisfied for the digraph case when \( n^{1+\varepsilon} = o(m) \) and \( m \leq \mu_0 n^2 \) for some constants \( \varepsilon, \mu_0 > 0 \). Thus, part (a) of Theorem 1.3 follows when \( m = m(n) \) satisfies these conditions.

We justifi the remaining range of \( m \) in (a), and parts (i) and (iii) of (b), using previously known enumeration results for the sparse cases. Again we discuss the bipartite cases first, beginning with the sparse case (b)(i). From Greenhill, McKay, and Wang [11, Corollary 5.2] it follows immediately that \( \mathbb{P}_{D(G)}(d) \sim \mathbb{P}_{B_n}(d) \) as long as \( 1 \leq \Delta(s) \Delta(t) = o(m^{2/3}) \) and \( \sigma^2(s)/(s(1 - \mu)) \sim 1 \) and \( \sigma^2(t)/(t(1 - \mu)) \sim 1 \) (or, in the notation of [11], \( \mu_2 \sim \mu_3 \sim 1 \)). The latter two conditions are equivalent to \( \text{Var} s_a = s \) and \( \text{Var} t_v = t \) for all \( a \in S \) and \( v \in T \), since \( \mu = o((n\ell')^{-1/4}) = o(1) \). So it suffi ces to show that the random sequences satisfy these three conditions with probability \( 1 - O(N^{-\omega}) \) in both the random bipartite graph degree sequence model \( D(G) \) and the binomial model \( B \), where we recall \( N = \ell' + n \). For this, we use Lemma 2.10, so the following probability estimates apply to both \( D(G) \) and \( B \) simultaneously.

To consider the first condition, note that Lemma 2.10(i) with \( \alpha = \beta(\sqrt{s \log N} + \log N) \) for any \( \beta \to \infty \) implies that \( s_a = s + \alpha \) with probability \( 1 - O(N^{-\omega}) \). Noting that \( s + \alpha < 2\beta(s + \log N) \) for \( \beta > 1 \), combining with the analogous inequality for \( t_v \), and using the union bound for \( a \in S \) and \( v \in T \), we get

\[
\Delta(s) \Delta(t) < 4\beta(s + \log N)(t + \log N)
\]

with probability \( 1 - O(N^{-\omega}) \). Now \( st = m^2/n\ell' = o(m^{2/3}) \) by the assumption \( m = o((n\ell')^{3/4}) \) of part (b)(i) of the theorem. Second, \( s \log N = (m \log N)/\ell' = o(m^{2/3}) \) since

\[
m^{1/3} = o(\min\{\ell' / \log n, n / \log \ell'\})
\]
by the theorem’s assumption. Third, \( \log^2 N = o(m^{2/3}) \) by the first assumption in part (b) of the theorem. Thus, \( \Delta(s)\Delta(t) = o(m^{2/3}) \) with probability 1 – \( O(N^{-\omega}) \).

The final two conditions, concerning \( \text{Var}_s \) and \( \text{Var}_t \), are implied by Lemma 2.10(ii) for the following reasons. Since \( \log^3 N = o(m) \) we may choose \( a = o(1) \) such that \( \log^3 N = o(a^2 m) \). Also, \( t\log^2 N = (m/n)\log^2 N \), which is \( o(m) \) because in this sparse case we have \( \log^3 \epsilon = o(m) = o((n/\log \ell)^3) \), whence \( \log^2 \epsilon = o(n) \). Thus we may impose the additional condition on the choice of \( a = o(1) \) such that \( \log^2 N = (m/n)\log^2 (\epsilon + n) = o(a^2 m) \). Then Lemma 2.10(ii) gives \( \sigma^2(t) \sim \text{Var}_t \) with probability \( 1 - o(N^{-\omega}) \), where \( \text{Var}_t \sim t \) since \( \mu = o(1) \). The similar statement \( \sigma^2(s) \sim s \) follows by a symmetric argument. This completes the proof of (b)(i).

Case (b)(iii) is a special case of [18, Theorem 1(c)], which also covers slightly lower values of \( m \), to approximately \( n\epsilon / \log n \), at the expense of decreasing the permissible imbalance between \( \ell \) and \( n \).

For the digraph case (a), as mentioned above, our enumeration result of Theorem 1.1 yields the asymptotic equivalence of the models when \( n^{1+\epsilon} = o(m) \) and \( m \leq \mu_0 m^2 \) for any \( \epsilon > 0 \) and for sufficiently small constant \( \mu_0 > 0 \). The denser range is directly covered by [18, Theorem 1(d)], which applies when \( \min\{m, n(n - 1) - m\} > n^2 / c \log n \) for any \( c > 0 \). Only the very sparse case remains. As \( n^{1+\epsilon} = o(m) \) is already covered, we may now assume that \( m = o(n^{4/3}) \). Since \( \log^3 n = o(m) \) is assumed in Theorem 1.3, we may again argue using Lemma 2.10, as in the case of (b) above, to show that, with probability \( 1 - O(n^{-\omega}) \), both \( \Delta(s) \) and \( \Delta(t) \) are bounded from above by \( O(m/n + \log n) \). In particular, \( \Delta(d)^4 = o(m) \) with probability \( 1 - O(n^{-\omega}) \), (using the upper bound on \( m \), and also the assumption of (a) that \( \log^4 n = o(m) \)), for random \( d \) in both the digraph and binomial models. Under this condition, we can apply the enumeration formula of McKay [17, Theorem 4.6], which is for bipartite graphs with given degree sequence and a forbidden set \( X \) of edges. Letting \( X \) be a perfect matching between the two parts of the vertex partition, this gives an asymptotic formula for the number of digraphs with in- and out-degree sequence \( (s, t) \). Using concentration of \( \sigma^2(s) \) and \( \sigma^2(t) \), established again as for the bipartite case, we find that this formula implies \( \text{Pr}_D(d) \sim \text{Pr}_B(d) \) under the conditions assumed. We deduce that \( D(G) \) and \( B \) are asymptotically quite equivalent in this case as well. ■

**Proof of Theorem 1.4.** The stated formula is a by-product of the proof of Theorem 1.1, so here we just point to the relevant spots within that proof. Recall the definition of \( P^* \), which is our approximation to the edge probability, at the beginning of Section 3, and note that a parameterization yields the formula given in the statement of Theorem 1.4. Claim 3.2 then yields the desired approximation since \( \mu \epsilon^4 = O(\min\{s, t\}^{4\epsilon - 2} m/n\ell) \). ■

4 | DISCUSSION

In this article we prove that the number of bipartite graphs with parts of sizes \( \ell \) and \( n \) and with a given degree sequence \( d \) is asymptotically equal to \( \mathbb{P}_B(d)\hat{H}(d) \left( \frac{n\ell}{m} \right) \), where \( m = \sum_i d_i / 2 \), \( B \) is a certain model of independent binomials, and \( \hat{H}(d) \) is defined in (2), under certain conditions on the parameters. These conditions are met if the edge density \( m/n\ell \) is bounded from above by a sufficiently small constant and from below by (roughly) \( n^{\ell - 1} + \ell^{\ell - 1} \) for some \( \epsilon > 0 \), and if the degrees \( d_i \) do not differ from the average degree of their respective parts by much more than a random bipartite graph with \( m \) edges.

For the range of parameters \( (\ell, n, m) \) where our enumeration result is applicable, it implies that the degree sequence of the random bipartite graph \( G(\ell, n, m) \) can be accurately modeled by a sequence of \( \ell + n \) independent binomials (conditioned on their first \( \ell \) and last \( n \) entries summing to \( m \)); see Theorem 1.3.
Discrepancy of part sizes
Our enumeration result permits a large imbalance between $n$ and $\ell$. When $\mu = m/n\ell$ is a small enough constant, for example, then $\ell$ may be roughly between $n^{2/3}$ and $n^{3/2}$. A formula due to Canfield and McKay [7] applies to biregular sequences for constant density when $\ell$ is bounded from above by roughly $n^{1/2}$. We do not know of any results covering the range roughly between $n^{1/2}$ and $n^{2/3}$. Similarly, for smaller $\mu$ a gap is left between Theorem 1.1 and the range covered in [7].

Digraphs and other discrete structures
Our approach to digraphs is to consider balanced bipartite graphs with a forbidden perfect matching. The same process could be used to produce formulae for numbers of (bipartite) graphs with other forbidden subgraphs, though this looks rather more complicated than the procedure in the present paper, and to deal with arbitrary subgraphs in large generality would require a different approach to some parts of the analysis.

Towards finding the approximating formulae
Let us take a moment to discuss how to find the approximating formulae $P^* = \pi$, $R^* = \rho$ and $Y^*$. Although we refer to those functions as “guessed formulae” in Section 3, there is of course a structured way to obtain them. Recall that Proposition 2.6 provides recursive relations between the “real” edge probabilities, ratios and 2-edge-path probabilities in terms of those functions of nearby sequences $d'$. We may feed in a certain approximation, say $P_{av} \equiv 0$ and $Y_{a,v,b} \equiv 0$ into the right hand side of the formulae to obtain first $R_{a,b}(d) = d_a/d_b$ and then updated approximations $P_{av}(d)$ and $Y_{a,v,b}(d)$. Applying expansions similar to those in the proof of Lemma 3.1, we may continue this process to obtain successive approximations to all three functions, of ever increasing accuracy. Noting any geometric series that start appearing in the terms of these approximations, and taking their limits, eventually leads to the “guessed” formulae which are a sufficiently good approximation for our purposes.

The main step remaining is to compare the resulting approximation for $R$ with the ratios of the formula for $N(d)$ we wish to prove. In some cases a conjectured formula for $N(d)$ exists, for example, for ordinary graphs treated in [14] and for bipartite graphs treated in this article. To apply the method when a conjectured formula does not exist would require an extra step, which in theory can be accomplished by performing a summation, since we have determined the ratios of numbers of graphs with different degree sequences, and we know the total number of graphs.

Finally, we remark that, with suitable adjustments, our method also applies to the sparse case when only upper bounds on the degrees are assumed and no bound on the degree spread. (Details are given in the first preprint version of this article [15].)

ACKNOWLEDGMENTS
We would like to thank the anonymous referees for their thorough reading and their helpful comments that also improved the article. In particular, this included pointing out that the results in [17] are applicable to digraphs, which permitted us to remove our own separate argument for sparse digraphs. Open access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

FUNDING INFORMATION
Research supported by ARC Discovery Project DP180103684
REFERENCES

1. A. Barvinok, *On the number of matrices and a random matrix with prescribed row and column sums and 0–1 entries*, Adv. Math. 224 (2010), 316–339.
2. A. Barvinok and J. A. Hartigan, *The number of graphs and a random graph with a given degree sequence*, Random Struct. Algorithms 42 (2013), no. 3, 301–348.
3. A. Békéssy, P. Békéssy, and J. Komlós, *Asymptotic enumeration of regular matrices*, Stud. Sci. Math. Hungar. 7 (1972), 343–353.
4. E. A. Bender, *The asymptotic number of non-negative integer matrices with given row and column sums*, Discr. Math. 10 (1974), 217–223.
5. B. Bollobás and B. D. McKay, *The number of matchings in random regular graphs and bipartite graphs*, J. Comb. Theory Ser. B 41 (1986), 80–91.
6. E. R. Canfield, C. Greenhill, and B. D. McKay, *Asymptotic enumeration of dense 0–1 matrices with specified line sums*, J. Comb. Theory Ser. A 115 (2008), 32–66.
7. E. R. Canfield and B. D. McKay, *Asymptotic enumeration of dense 0-1 matrices with equal row sums and equal column sums*, Electron. J. Comb. 12 (2005), R29.
8. C. J. Everett and P. R. Stein, *The asymptotic number of integer stochastic matrices*, Discr. Math. 1 (1971), 55–72.
9. I. J. Good and J. F. Crook, *The enumeration of arrays and a generalization related to contingency tables*, Discr. Math. 19 (1977), 23–45.
10. C. Greenhill and B. D. McKay, *Random dense bipartite graphs and directed graphs with specified degrees*, Random Struct. Algorithms 35 (2009), 222–249.
11. C. Greenhill, B. D. McKay, and X. Wang, *Asymptotic enumeration of sparse 0–1 matrices with irregular row and column sums*, J. Comb. Theory Ser. A 113 (2006), 291–324.
12. M. Isaev and B. D. McKay, *Complex martingales and asymptotic enumeration*, Random Struct. Algorithms 52 (2018), 617–661.
13. S. Janson, T. Łuczak, and A. Ruciński, Random graphs, Vol 45, John Wiley & Sons, Boca Raton, 2011.
14. A. Liebenau and N. Wormald, *Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph*, arXiv:1702.08373, 2017.
15. A. Liebenau and N. Wormald, *Asymptotic enumeration of digraphs and bipartite by degree sequence*, arXiv:2006.15797, 2020.
16. C. McDiarmid, *On the method of bounded differences*, Surv. Comb. 141 (1989), 148–188.
17. B. D. McKay, *Asymptotics for 0-1 matrices with prescribed line sums*, Enumer. Design (1984), 225–238.
18. B. D. McKay and F. Skerman, *Degree sequences of random digraphs and bipartite graphs*, J. Comb. 7 (2016), 21–49.
19. B. D. McKay and X. Wang, *Asymptotic enumeration of 0–1 matrices with equal row sums and equal column sums*, Linear Algebra Appl. 373 (2003), 273–287.
20. B. D. McKay and N. C. Wormald, *The degree sequence of a random graph. I. The models*, Random Struct. Algorithms 11 (1997), 97–117.
21. M. P. Mineev and A. I. Pavlov, *On the number of (0, 1) matrices with prescribed numbers of rows and columns*, Doc. Akad. Nauk SSSR 230 (1976), 1276–1282.
22. L. Mirsky, Transversal theory, Vol 197, Academic Press, New York, 1971.
23. P. E. O’Neil, *Asymptotics and random matrices with row-sum and column sum-restrictions*, Bull. Am. Math. Soc. 75 (1969), 1276–1282.
24. R. C. Read, *Some enumeration problems in graph theory*, PhD thesis, University of London, 1958.
25. N. C. Wormald, *Some problems in the enumeration of labelled graphs*, Bull. Austr. Math. Soc. 21 (1980), 159–160.

How to cite this article: A. Liebenau, and N. Wormald, *Asymptotic enumeration of digraphs and bipartite graphs by degree sequence*, Random Struct. Alg. 62 (2023), 259–286. https://doi.org/10.1002/rsa.21105