SIAM J. Discrete Math.
Vol. 27, No. 4, pp. 2064–2089

ASYMPTOTIC ENUMERATION OF SPARSE MULTIGRAPHS WITH GIVEN DEGREES*

CATHERINE GREENHILL† and BRENDAN D. MCKAY‡

Abstract. Let J and J^* be subsets of $\mathbb{N}$ such that $0, 1 \in J$ and $0 \in J^*$. For infinitely many n, let $\boldsymbol{k} = (k_1, \ldots, k_n)$ be a vector of nonnegative integers whose sum M is even. We find an asymptotic expression for the number of multigraphs on the vertex set $\{1, \ldots, n\}$ with degree sequence given by $\boldsymbol{k}$ such that every loop has multiplicity in J^* and every nonloop edge has multiplicity in J. Equivalently, these are symmetric integer matrices with values J^* allowed on the diagonal and J off the diagonal. Our expression holds when the maximum degree $k_{\max}$ satisfies $k_{\max} = o(M^{1/3})$. We prove this result using the switching method, building on an asymptotic enumeration of simple graphs with given degrees [B. D. McKay and N. C. Wormald, *Combinatorica*, 11 (1991), pp. 369–382]. Our application of the switching method introduces a novel way of combining several different switching operations into a single computation.

Key words. multigraph, integer matrix, symmetric matrix, asymptotic enumeration, switching

AMS subject classifications. 05A16, 05C30, 15B36, 62H17

DOI. 10.1137/130913419

1. Introduction. Multigraphs arise in many applications including modeling transportation networks [1], the structure of RNA [6, 10], and nonparametric statistics [7]. We use the terminology "multigraphs" inclusively, with both multiple edges and (possibly multiple) loops allowed. We seek an asymptotic enumeration formula for multigraphs with a given degree sequence, satisfying certain conditions.

Let $k_{i,n}$ be a nonnegative integer for all pairs (i, n) of integers which satisfy $1 \le i \le n$. Then for each $n \ge 1$, let $\boldsymbol{k} = \boldsymbol{k}(n) = (k_{1,n}, \ldots, k_{n,n})$. We usually write k_i instead of $k_{i,n}$. Define $M = \sum_{i=1}^{n} k_i$. We assume that M is even for an infinite number of values of n and tacitly restrict ourselves to such n.

For subsets J, J^* of $\mathbb{N}$, define $\mathcal{G}(\boldsymbol{k}, J, J^*)$ to be the set of all multigraphs on the vertex set $\{1, \ldots, n\}$ with degree sequence given by $\boldsymbol{k}$ such that the multiplicity of every loop belongs to J^* and the multiplicity of every nonloop edge belongs to J. Loops contribute 2 to the degree of their vertex. Also define $\mathcal{G}(\boldsymbol{k})$ to be the set of multigraphs on $\{1, \ldots, n\}$ with degree sequence $\boldsymbol{k}$ (and no restrictions on the multiplicities).

Equivalently we may think of $\mathcal{G}(\boldsymbol{k}, J, J^*)$ as the set of all $n \times n$ symmetric matrices $A = (a_{ij})$ with all diagonal entries in J^*, all off-diagonal entries in J and with

$$2a_{ii} + \sum_{j \ne i} a_{ij} = k_i \tag{1.1}$$

for $i = 1, 2, \ldots, n$. Note that diagonal entries are weighted by 2 in the row sum.

Let $k_{\max} = \max_{i=1}^{n} k_i$. We find an asymptotic expression for

$$G(\boldsymbol{k}, J, J^*) = |\mathcal{G}(\boldsymbol{k}, J, J^*)|$$

*Received by the editors March 18, 2013; accepted for publication (in revised form) September 27, 2013; published electronically December 11, 2013. This research was supported by the Australian Research Council.
http://www.siam.org/journals/sidma/27-4/91341.html

†School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia (csg@unsw.edu.au).
‡Research School of Computer Science, Australian National University, Canberra, ACT 0200, Australia (bdm@cs.anu.edu.au).

that holds when $k_{\max} = o(M^{1/3})$. For $r = 1, 2, \ldots$ let

$$M_r = \sum_{i=1}^{n} [k_i]_r,$$

where $[a]_r = a(a-1)(a-2)\cdots(a-r+1)$ denotes the falling factorial. Then $M_1 = M$ and $M_r \le k_{\max} M_{r-1}$ for all $r \ge 2$.

For $i \ge 0$, define

$$x_i = \begin{cases} 1 & \text{if } i \in J, \\ 0 & \text{otherwise;} \end{cases} \qquad y_i = \begin{cases} 1 & \text{if } i \in J^*, \\ 0 & \text{otherwise.} \end{cases}$$

Our main result is the following.

THEOREM 1.1. *Let J and J^* be subsets of $\mathbb{N}$ such that $0, 1 \in J$ and $0 \in J^*$. Suppose that $n \to \infty$, $M \to \infty$, and $k_{\max}^3 = o(M)$. Then*

$$G(\boldsymbol{k}, J, J^*) = \frac{M!}{(M/2)!\, 2^{M/2} k_1! \cdots k_n!} \exp\biggl(\bigl(y_1 - \tfrac{1}{2}\bigr)\frac{M_2}{M} + \bigl(x_2 - \tfrac{1}{2}\bigr)\frac{M_2^2}{2M^2} + \frac{M_2^4}{4M^5} - \frac{M_2^2 M_3}{2M^4} + \bigl(x_3 - x_2 + \tfrac{1}{3}\bigr)\frac{M_3^2}{2M^3} + O(k_{\max}^3/M)\biggr).$$

For convenience, we restate the formula in the regular case.

COROLLARY 1.2. *Suppose that $\boldsymbol{k} = (k, \ldots, k)$ with kn even, such that $k = o(n^{1/2})$ as $n \to \infty$. Then*

$$G(\boldsymbol{k}, J, J^*) = \frac{(kn)!}{(kn/2)!\, 2^{kn/2} (k!)^n} \exp\bigl(-Q(k,n) + O(k^2/n)\bigr),$$

where

$$Q(k,n) = \tfrac{1}{4}\,(k-1)\bigl((-1)^{x_2}(k-1) + 2(-1)^{y_1}\bigr) + \frac{k^3}{12n}(6x_2 - 6x_3 + 1).$$

Setting $y_1 = x_2 = x_3 = 0$ we recapture the asymptotic expression for sparse simple graphs with given degrees presented in [17, Theorem 5.2]. Similarly, setting $y_1 = 1$, $x_2 = x_3 = 0$ we obtain the asymptotic enumeration by degree sequence of sparse graphs with loops allowed but no multiple edges: this is the second expression in [5, Theorem 1.5].

We remark that the conditions $0 \in J^*$ and $0, 1 \in J$ in Theorem 1.1 can be replaced by weaker conditions, namely, that J^* is nonempty and that J has at least two elements with the smallest two consecutive. Let s denote the least element of J^* and let $t, t+1$ be the smallest two elements of J. This case is reduced to ours by subtracting s from each diagonal element, t from each off-diagonal element, and $2s + (n-1)t$ from each k_i, provided the new degree sequence satisfies the conditions of Theorem 1.1.

Theorem 1.1 is proved using the switching method, building on an asymptotic enumeration of simple graphs with given degrees [17]. Our application of the switching method introduces a novel way of combining several different switching operations into a single computation.

The remainder of the paper is structured as follows. The history of this asymptotic enumeration problem is briefly reviewed in section 1.1, then the new switching theorem is given in section 2. In section 3 we describe 15 types of switchings on multigraphs, which are used to show that certain multiplicities of edges or loops are rare. These switchings are analyzed together in subsection 3.14 using the new switching theorem. In section 4 we complete the enumeration with the help of some calculations performed in [17].

Finally, in section 5 we show that a naïve argument leads to a formula for $G(\boldsymbol{k}, J, J^*)$ which differs asymptotically by a constant factor from the result of Theorem 1.1 in the regular case. The constant factor takes two different values, depending on whether $2 \in J$.

1.1. History. The earliest work on this problem was that of Read [19, p. 156], who found exact and asymptotic formulae when $k_1 = \cdots = k_n = 3$ for all four combinations of $J^* \in \{\{0\}, \mathbb{N}\}$ and $J \in \{\{0,1\}, \mathbb{N}\}$. The best result for $J^* = \{0\}$, $J = \{0,1\}$ in the sparse range is that of McKay and Wormald [17], who treated $k_{\max} = o(M^{1/3})$; see that paper for a survey of the many earlier results on the 0-1 case. In addition to $J^* = \{0\}, J = \{0,1\}$, Bender and Canfield [3] found the asymptotics for $J^* = \{0\}, J = \mathbb{N}$ when $k_{\max} = O(1)$. (Although that paper allows nonzero diagonal entries, they contribute singly to the row sums, not doubly as we have it.)

In [15], McKay and Wormald considered $J^* = \{0\}, J = \{0,1\}$ in the dense domain defined by $\min\{k, n-k-1\} \ge cn/\log n$ for $c > \frac{2}{3}$ and $|k_i - k| = O(n^{1/2+\varepsilon})$ for all i, where k is the average degree. McKay [12] found a better error term under the same conditions, while Barvinok and Hartigan allowed a wider range of degrees [2]. Greenhill and McKay [5] added the option of loops in both the sparse and dense ranges by considering $J^* = \{0,1\}$, $J = \{0,1\}$. Finally, McKay and McLeod [13] analyzed the case of $J^* = \{0\}, J = \mathbb{N}$ when $k_1 = \cdots = k_n > cn/\log n$ for $c > \frac{1}{6}$.

For sparse rectangular matrices which are not necessarily symmetric, see Greenhill and McKay [4].

2. The switching theorem. In order to bound the number of multigraphs having some unusual properties, we will apply the method of switchings. It will be necessary to apply several switching types, which we could analyze one at a time as in [4]. Instead we now prove a generalized switching theorem that allows us to analyze all of them at once.

In order to facilitate use of the method in future work, we will present the theorem in greater generality than is required in this paper.

THEOREM 2.1. *Let $\Gamma = \Gamma(V,E)$ be a directed multigraph, and let $\alpha : E \to \mathbb{R}_+$ be a positive weighting of the edges of Γ. Fix a nonempty finite set C, whose elements we will call colors, and let $c : E \to C$ be an edge coloring of Γ (which need not be proper). For all $v \in V$ and $c \in C$ denote by $\Gamma_c^-(v)$ the set of edges of color c entering v and by $\Gamma_c^+(v)$ the set of edges of color c leaving v.*

We introduce the set of variables $\{N(v) : v \in V\} \cup \{s(e) : e \in E\}$ and consider the following system of linear inequalities on these variables:

$$N(v) \ge 0 \qquad (v \in V), \tag{2.1}$$

$$s(e) \ge 0 \qquad (e \in E), \tag{2.2}$$

$$\sum_{e \in \Gamma_c^-(v)} s(e) \le N(v) \qquad (v \in V,\ c \in C), \tag{2.3}$$

$$\sum_{e \in \Gamma_c^+(v)} \alpha(e) s(e) \ge N(v) \qquad (v \in V,\ c \in C,\ \Gamma_c^+(v) \ne \emptyset). \tag{2.4}$$

Let $C(v) = \{c \in C : \Gamma_c^-(v) \neq \emptyset\}$ be the set of colors entering v. For each $v \in V$ which is not a source, let $\lambda_c(v), c \in C(v)$ be positive numbers such that $\sum_{c\in C(v)} \lambda_c(v) \leq 1$. For each $vw \in E$, define $\hat\alpha(vw) = \alpha(vw)/\lambda_{c(vw)}(w)$. Extend this function to any directed path P by defining $\hat\alpha(P) = \prod_{e\in P} \hat\alpha(e)$.

Now suppose that $Y, Z \subseteq V$ satisfy the following conditions:

1. *$Z \neq \emptyset$ and $Y \cap Z = \emptyset$.*
2. *If $v \in V$ is a sink of Γ, or if $\hat\alpha(vw) \geq 1$ for some $vw \in E$, then $v \in Z$.*

For any $W, W' \subseteq V$, define $\mathcal{P}(W, W')$ to be the set of nontrivial directed paths in Γ which start in W, end in W', and have no internal vertices in $Y \cup Z$. Then every solution of (2.1)–(2.4) *satisfies*

$$\sum_{v\in Y} N(v) \leq \frac{\max_{P\in\mathcal{P}(Y,Z)} \hat\alpha(P)}{1 - \max_{P\in\mathcal{P}(Y,Y)} \hat\alpha(P)} \sum_{v\in Z} N(v), \tag{2.5}$$

where the maximum over an empty set is taken to be 0.

Proof. The case of one color appears in [9, Theorem 3], apart from an inconsequential difference in the conditions on Z. We proceed by reducing the general problem to one in which there is only one color, then modifying Γ slightly so that [9, Theorem 3] applies.

Define $\Gamma^-(v)$ and $\Gamma^+(v)$ to be the set of all edges (regardless of color) entering v or leaving v, respectively. For each edge $vw \in E$, define $\hat s(vw) = s(vw)\lambda_{c(vw)}(w)$. If we weight inequality (2.3) by $\lambda_c(v)$ and sum over $c \in C(v)$, we obtain

$$\sum_{e\in\Gamma^-(v)} \hat s(e) \leq N(v) \qquad (v \in V). \tag{2.6}$$

Similarly, if $\Gamma^+(v) \neq \emptyset$, then we can sum (2.4) over those $c \in C$ with $\Gamma_c^+(v) \neq \emptyset$ to obtain

$$\sum_{e\in\Gamma^+(v)} \hat\alpha(e)\hat s(e) \geq N(v) \qquad (v \in V,\ \Gamma^+(v) \neq \emptyset). \tag{2.7}$$

Together with the nonnegativity of $N(v)$ and $\hat s(e)$, we have equations of the form of (2.1)–(2.4) with only one color.

Next remove from Γ all edges vw where $v \in Z$. Doing so can only weaken the conditions, by decreasing the left-hand side of some inequalities in (2.6) or removing some inequalities in (2.7). Note that none of the quantities in (2.5) are changed by removal of these edges. After this change to Γ, Z satisfies the requirements of [9, Theorem 3] with the variable X defined there set equal to Z. Applying that theorem to the system defined by (2.6), (2.7) and the nonnegativity of $N(v)$ and $\hat s(e)$ gives (2.5). □

We now describe how Theorem 2.1 can be used for counting.

Suppose we have a finite set of "objects" partitioned into disjoint classes $\mathcal{S}(v)$, where $v \in V$ for some index set V. Define $N(v) = |\mathcal{S}(v)|$ for each $v \in V$. Also suppose that for each $c \in C$ we have a relation Ψ_c between objects: to be precise, Ψ_c is a multiset of ordered pairs (Q, R) of objects. (We call Ψ_c a *switching* and usually define it by some operation that modifies Q to make R.)

Now define an edge-colored directed multigraph $\Gamma = (V, E)$ with vertex set V, where Γ has a directed edge vw of color c if and only if $(Q, R) \in \Psi_c$ for some $Q \in \mathcal{S}(v)$ and $R \in \mathcal{S}(w)$. (There is at most one edge of each color between any pair of distinct vertices of Γ.) For each $vw \in E$ let $s'(vw) = |\{(Q,R) \in \Psi_c : Q \in \mathcal{S}(v), R \in \mathcal{S}(w)\}|$, counting multiplicities, where c is the color of vw.

Fix $v \in V$ and $c \in C$ such that $\Gamma_c^+(v) \neq \emptyset$. Suppose that for any $Q \in \mathcal{S}(v)$ there are at least $a_c(v) > 0$ objects R with $(Q,R) \in \Psi_c$, counting multiplicities. Then

$$\sum_{e \in \Gamma_c^+(v)} s'(e) \geq a_c(v) N(v).$$

Similarly, for fixed $w \in V$ and $c \in C$, suppose that for every $R \in \mathcal{S}(w)$ there are at most $b_c(w) > 0$ objects Q with $(Q,R) \in \Psi_c$, counting multiplicities. Then

$$\sum_{e \in \Gamma_c^-(w)} s'(e) \leq b_c(w) N(w).$$

Defining $s(vw) = s'(vw)/b_{c(vw)}(w)$ and $\alpha(vw) = b_{c(vw)}(w)/a_{c(vw)}(v)$ we obtain (2.1)–(2.4). Theorem 2.1 can thus be used to bound the relative values of $\sum_{v \in Z}|\mathcal{S}(v)|$ and $\sum_{v \in Y}|\mathcal{S}(v)|$ if Y, Z satisfy the requirements of the lemma. Since $\sum_{v \in Z}|\mathcal{S}(v)| \leq \sum_{v \in V}|\mathcal{S}(v)|$, this also bounds $\sum_{v \in Y}|\mathcal{S}(v)|$ relative to $\sum_{v \in V}|\mathcal{S}(v)|$; i.e., it bounds the fraction of all objects that lie in $\bigcup_{v \in Y} \mathcal{S}(v)$.

3. Switchings on multigraphs.

Define

$$\begin{aligned} N_1 &= \max\bigl\{\lceil \log M \rceil, \lceil 480 M_2/M \rceil\bigr\}, \\ N_2 &= \max\bigl\{\lceil \log M \rceil, \lceil 240 M_2^2/M^2 \rceil\bigr\}, \\ N_3 &= \max\bigl\{\lceil \log M \rceil, \lceil 240 M_3^2/M^3 \rceil\bigr\}. \end{aligned} \tag{3.1}$$

(We have not attempted to optimize constants.) Given a multigraph Q, let $\ell_D(Q)$ denote the number of loops with multiplicity D and let $e_D(Q)$ denote the number of nonloop edges with multiplicity D, for $D \geq 1$.

Let

$$\mathcal{G}_0 = \mathcal{G}(\boldsymbol{k}, J \cup \{4,5,6,\ldots\}, J^* \cup \{2,3,4,\ldots\})$$

be the set of all multigraphs with degree sequence $\boldsymbol{k}$ and allowing all multiplicities except for those in $\{1\} - J^*$ on loops and those in $\{2,3\} - J$ on nonloops. Note that $\mathcal{G}(\boldsymbol{k}, J, J^*) \subseteq \mathcal{G}_0$.

We also define the subsets Y, Z of $\mathcal{G}_0$ by

$$\begin{aligned} Y = \mathcal{G}_0 - \{Q \in \mathcal{G}_0 \mid &\ \ell_D(Q) = 0 \text{ for } D \geq 2,\ e_D(Q) = 0 \text{ for } D \geq 4, \\ &\ \ell_1(Q) \leq N_1,\ e_2(Q) \leq N_2,\ e_3(Q) \leq N_3\}, \end{aligned} \tag{3.2}$$

$$\begin{aligned} Z = \{Q \in \mathcal{G}_0 \mid &\ \ell_D(Q) = 0 \text{ for } D \geq 2,\ e_D(Q) = 0 \text{ for } D \geq 4, \\ &\ \ell_1(Q) \leq \lceil N_1/2 \rceil,\ e_2(Q) \leq \lceil N_2/2 \rceil,\ e_3(Q) \leq \lceil N_3/2 \rceil\}. \end{aligned} \tag{3.3}$$

We define 15 colored switchings which act to reduce the number of loops or edges with high multiplicities, moving in steps from Y toward Z. These switchings are defined below, together with a description of when each should be used. Indeed, any given switching will only be used on multigraphs in $\mathcal{G}_0$ for which none of the switchings with a lower-labeled color are applicable. An important property of all the switchings is that they do not create simple loops (with multiplicity 1) and they do not create nonloop edges of multiplicity 2 or 3, as these may not be allowed for multigraphs in $\mathcal{G}_0$.

For each switching color c and multigraphs Q, R we will define the following parameters:

- $a_c(Q)$ is a lower bound on the number of ways in which a switching of color c can be applied to Q.
- $b_c(R)$ is an upper bound on the number of ways in which a switching of color c can produce R.
- If R can be obtained from Q by performing a switching of color c, then we let $\alpha(Q, R)$ be an upper bound on $b_c(R)/a_c(Q)$. (Note that the color c is determined by the edge QR.)

Finally in section 3.14 we will analyze all the switchings at once by applying Theorem 2.1.

3.1. Switchings of color 1. A switching of color 1 is used to reduce the number of loops of multiplicity equal to 2 or multiplicity at least 4. It is applied to multigraphs $Q \in \mathcal{G}_0$ with $L(Q) > \lceil 3M^{1/2} \rceil$, where

$$L(Q) = \ell_2(Q) + \sum_{D \geq 4} \ell_D(Q) \tag{3.4}$$

is the number of loops in Q with multiplicity 2 or multiplicity at least 4. The switching is described by the sequence (v_1, v_2) of distinct vertices such that there is a loop of multiplicity D_1 at v_1 and a loop of multiplicity D_2 at v_2 with $D_1, D_2 \in \{2\} \cup \{4, 5, \ldots\}$.

Let m be the multiplicity of the edge from v_1 to v_2 in Q (which may equal zero). The switching reduces the multiplicity of both loops by 2 and increases the multiplicity of the edge $\{v_1, v_2\}$ by 4. This operation is depicted in Figure 3.1.

Suppose that R can be produced from Q using a switching of color 1. The number of ways to perform a switching of color 1 in Q is exactly

$$[L(Q)]_2 \geq 9M$$

as there is no restriction on the value of m. The number of ways that R can be produced using a switching of color 1 is bounded above by $M/4$, which is a bound on the number of ways to choose an oriented edge with multiplicity at least 4. Hence we can set $a_1(Q) = 9M$ and $b_1(R) = M/4$, giving

$$\alpha(Q, R) = \frac{1}{36}.$$

3.2. Switchings of color 2. A switching of color 2 is used to reduce the number of loops of multiplicity 3 to at most $\lceil 3M^{1/2} \rceil$. It is performed for multigraphs Q for which switchings of color 1 do not apply and such that $\ell_3(Q) > \lceil 3M^{1/2} \rceil$. The switching is described by the sequence (v_1, v_2) of distinct vertices such that there is a loop of multiplicity 3 at v_1 and at v_2.

Let m be the multiplicity of the edge from v_1 to v_2 in Q (which may equal zero). The switching removes the loop at v_1 and at v_2, and increases the multiplicity of the edge $\{v_1, v_2\}$ by 6, as illustrated in Figure 3.2.

Suppose that the multigraph R can be produced from Q using a switching of color 2. The number of ways to perform a switching of color 2 in Q is exactly $[\ell_3(Q)]_2 \geq 9M$, and the number of ways that R can be produced using a switching of color 2 is at most $M/6$ (since an oriented edge of multiplicity at least 6 determines the reverse

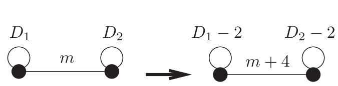

FIG. 3.1. *A switching of color* 1.

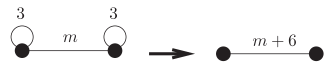

FIG. 3.2. *A switching of color* 2.

operation, from R to Q). Hence we can take $a_2(Q) = 9M$ and $b_2(R) = M/6$, leading to

$$\alpha(Q, R) = \frac{1}{54}.$$

3.3. Switchings of color 3. A switching of color 3 is used to reduce the number of loops of multiplicity 1 to at most $\lceil M^{1/2} \rceil$. It is applied to multigraphs Q for which switchings of color 1 and 2 do not apply and such that $\ell_1(Q) > \lceil M^{1/2} \rceil$. The switching is defined by the sequence (v_1, v_2, v_3) of distinct vertices such that each of v_1, v_2, v_3 has a simple loop in Q (of multiplicity 1), and none of the edges $\{v_1, v_2\}$, $\{v_1, v_3\}$, $\{v_2, v_3\}$ are present in Q. The switching removes these three loops and joins the three vertices pairwise by simple edges (of multiplicity 1), as illustrated in Figure 3.3.

Suppose that the multigraph R can be produced from Q using a switching of color 3. The number of ways to perform a switching of color 3 in Q is at least

$$[\ell_1(Q)]_3 - O\big(k_{\max}\ell_1(Q)^2\big) = \ell_1(Q)^3(1 - o(1)) \geq \tfrac{1}{2}\, M^{3/2},$$

while the number of ways that R can be produced using a switching of color 3 is at most $k_{\max}M$. Therefore we can take $a_3(Q) = \frac{1}{2}M^{3/2}$ and $b_3(R) = k_{\max}M$, leading to

$$\alpha(Q, R) = \frac{2k_{\max}}{M^{1/2}} = o(1).$$

3.4. Switchings of color 4. A switching of color 4 is used to reduce the number of nonloop edges of multiplicity greater than $\max\{4, \lceil k_{\max}^{1/2} \rceil\}$. It is applied to multigraphs Q for which switchings of colors $1, 2, 3$ do not apply and such that $E^+(Q) > \lceil 4k_{\max}^{1/2}\, M^{1/2} \rceil$, where

$$E^+(Q) = \sum_{D=\max\{4,\, \lceil k_{\max}^{1/2} \rceil\}+1}^{k_{\max}} e_D(Q). \tag{3.5}$$

A switching of color 4 in Q is described by a sequence (v_1, w_1, v_2, w_2) of distinct vertices such that

- the multiplicity of edge (v_1, w_1) in Q is D_1 and the multiplicity of (v_2, w_2) in Q is D_2, where $D_1, D_2 > \max\{4, \lceil k_{\max}^{1/2} \rceil\}$,
- the edges $\{v_1, v_2\}$ and $\{w_1, w_2\}$ have multiplicity zero in Q (they are nonedges).

The switching reduces the multiplicity of edges (v_1, w_1) and (v_2, w_2) by one and gives multiplicity 1 to edges $\{v_1, v_2\}$ and $\{w_1, w_2\}$, as shown in Figure 3.4.

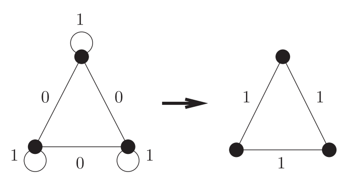

FIG. 3.3. *A switching of color* 3.

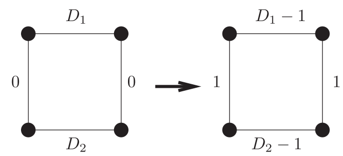

FIG. 3.4. *A switching of color* 4 *or* 5.

Now suppose that the multigraph R can be produced from Q using a switching of color 4. The number of switchings of color 4 that can be performed in Q is bounded below by

$$4[E^+(Q)]_2 - O(k_{\max}^2\, E^+(Q)) = 4E^+(Q)^2(1-o(1)) \geq 64\, k_{\max}\, M(1-o(1)) > 60k_{\max}\, M,$$

and the number of switchings of color 4 that can produce R is at most

$$\frac{2k_{\max}^2 M}{\max\{4,\, \lceil k_{\max}^{1/2}\rceil\}^2} \leq 2k_{\max}\, M.$$

Hence we can take $a_4(Q) = 60\, k_{\max}\, M$ and $b_4(R) = 2k_{\max} M$, leading to

$$\alpha(Q,R) = \frac{1}{30}.$$

3.5. Switchings of color 5. A switching of color 5 is used to reduce the number of nonloop edges of multiplicity at least 5 and at most $\lceil k_{\max}^{1/2}\rceil$. For a multigraph $Q \in \mathcal{G}_0$, let

$$E^-(Q) = \sum_{D=5}^{\lceil k_{\max}^{1/2}\rceil} e_D(Q).$$

Switchings of color 5 are applied to multigraphs Q for which switchings of colors $1,\ldots,4$ do not apply and such that $E^-(Q) > \lceil 3k_{\max} M^{1/2}\rceil$.

A switching of color 5 is described by a sequence (v_1, w_1, v_2, w_2) of distinct vertices such that

- the multiplicity of $\{v_1, w_1\}$ in Q is D_1 and the multiplicity of $\{v_2, w_2\}$ in Q is D_2, where $D_1, D_2 \in \{5, 6, \ldots, \lceil k_{\max}^{1/2}\rceil\}$, and
- the multiplicity of $\{v_1, v_2\}$ and $\{w_1, w_2\}$ in Q is zero (these are nonedges).

This switching is also illustrated by Figure 3.4, but with different conditions on D_1, D_2 as above.

Suppose that R can be produced from Q using a switching of color 5. The number of ways that a switching of color 5 can be performed in Q is bounded below by

$$4[E^-(Q)]_2 - O(k_{\max}^2 E^-(Q)) = 4E^-(Q)^2(1-o(1)) \geq 30k_{\max}^2 M,$$

and the number of ways to produce R using a switching of color 5 is at most

$$k_{\max}^2 M.$$

Hence we can let $a_5(Q) = 30k_{\max}^2 M$ and $b_5(R) = k_{\max}^2 M$, leading to

$$\alpha(Q,R) = \frac{1}{30}.$$

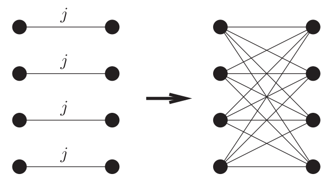

FIG. 3.5. *A switching of color* 8.

3.6. Switchings of color 6, 7, 8. A switching of color $4+j$ is used to reduce the number of edges of multiplicity j to at most $\lceil M^{5/6}\rceil$ for $j=2,3,4$. Switchings of color $4+j$ are applied to multigraphs Q for which switchings of colors $1,\dots,3+j$ do not apply and such that

$$e_j(Q) > \lceil M^{5/6}\rceil.$$

Given such a multigraph Q, a switching of color $4+j$ is defined by a sequence

$$(v_1, w_1, v_2, w_2, \dots, v_j, w_j)$$

of distinct vertices such that

- (v_r, w_r) is an edge of multiplicity j in Q for $r=1,\dots,j$,
- every edge $\{v_r, w_s\}$ with $1 \le r \ne s \le j$ has multiplicity 0 in Q (it is a nonedge).

The switching deletes these j edges of multiplicity j and inserts a complete bipartite graph $K_{j,j}$ on $\{v_1,\dots,v_j\}\cup\{w_1,\dots,w_j\}$ with all j^2 new edges simple (that is, multiplicity 1). This operation is illustrated in Figure 3.5 for the case $j=4$.

Suppose that the multigraph R can be obtained from Q using a switching of color $4+j$ for $j\in\{2,3,4\}$. The number of ways in which a switching of color $4+j$ can be performed in Q is bounded below by

$$2^j\,[e_j(Q)]_j - O(k_{\max}^2\, e_j(Q)^{j-1}) \ge \tfrac{1}{2}M^{5j/6},$$

while the number of switchings of color $4+j$ which produce R is at most $k_{\max}^{2j-2}M$. Hence we can define $a_{4+j}(Q) = \frac{1}{2}M^{5j/6}$ and $b_{4+j}(R) = k_{\max}^{2j-2}\,M$ for $j\in\{2,3,4\}$. Since $j\ge 2$, this leads to

$$\alpha(Q,R) = \frac{2k_{\max}^{2j-2}M}{M^{5j/6}} = O\biggl(\frac{k_{\max}^2}{M^{2/3}}\biggr) = o(1).$$

Before describing more colored switchings, we prove a useful fact.

LEMMA 3.1. *Suppose that* $Q\in\mathcal{G}_0$ *is such that no switching of color* 1 *to* 8 *applies to* Q. *Then* $e_1(Q) = (\frac{1}{2} - o(1))M$.

Proof. Since no switching of color 1, 2, or 3 applies we have

$$\sum_{D\ge 1} D\,\ell_D(Q) = O(k_{\max}\,M^{1/2}) = O(M^{5/6}),$$

and since no switching of colors 4–8 applies we know that

$$\begin{aligned}\sum_{D\ge 2} D\,e_D(Q) &\le 9\lceil M^{5/6}\rceil + k_{\max}\,E^+(Q) + \lceil k_{\max}^{1/2}\rceil\, E^-(Q)\\ &= O(M^{5/6} + k_{\max}^{3/2}\,M^{1/2})\\ &= o(M).\end{aligned}$$

The result follows. ∎

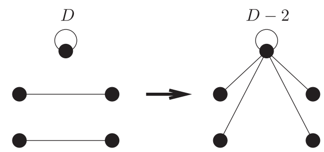

FIG. 3.6. *A switching of color* 9.

3.7. Switchings of color 9. Switchings of color 9 reduce the number of loops of multiplicity 2 or multiplicity at least 4, until this number is zero. They are applied to multigraphs Q for which the switchings of colors $1, \dots, 8$ do not apply and such that $L(Q) \geq 1$.

Let Q be such a multigraph. A switching of color 9 in Q is defined by a sequence $(v_0, v_1, w_1, v_2, w_2)$ of distinct vertices such that

- there is a loop at vertex v_0 in Q with multiplicity D, where $D = 2$ or $D \geq 4$,
- $\{v_1, w_1\}$ and $\{v_2, w_2\}$ are simple edges (of multiplicity 1),
- $\{v_0, v_1\}$, $\{v_0, v_2\}$, $\{v_0, w_1\}$, $\{v_0, w_2\}$ are all nonedges in Q (with multiplicity zero).

The switching reduces the multiplicity of the loop at v_0 to $D - 2$, removes the two simple edges $\{v_1, w_1\}$ and $\{v_2, w_2\}$ and inserts the four simple edges $\{v_0, v_1\}$, $\{v_0, v_2\}$, $\{v_0, w_1\}$, $\{v_0, w_2\}$, as shown in Figure 3.6.

Suppose that the multigraph R can be produced from Q by a switching of color 9. The number of ways to perform a switching of color 9 in Q is at least

$$L(Q)\big([2e_1(Q)]_2 - O(k_{\max}^2 e_1(Q))\big) \geq \tfrac{1}{2}\, L(Q)M^2$$

using Lemma 3.1, while the number of ways to produce R using a switching of color 9 is at most M_4. (We ignore the presence of the loop at v_0 in this upper bound, since no such loop exists when $D = 2$.)

Hence we can let $a_9(Q) = \frac{1}{2}L(Q)M^2$ and $b_9(R) = M_4$, leading to

$$\alpha(Q, R) = \frac{2M_4}{L(Q)\, M^2} = O\bigg(\frac{k_{\max}^3}{M}\bigg)$$

since $L(Q) \geq 1$.

3.8. Switchings of color 10. Switchings of color 10 reduce the number of loops of multiplicity 3, until this number is zero. They are applied to multigraphs Q such that switchings of colors $1, \dots, 9$ do not apply and such that $\ell_3(Q) \geq 1$.

Let Q be such a multigraph. A switching of color 10 is defined by a sequence of distinct vertices $(v_0, v_1, w_1, v_2, w_2, v_3, w_3)$ such that

- there is a loop of multiplicity 3 at v_0 in Q,
- there is a simple edge $\{v_j, w_j\}$ in Q for $j = 1, 2, 3$,
- the edges $\{v_0, v_j\}$ and $\{v_0, w_j\}$ all have multiplicity 0 in Q for $j = 1, 2, 3$.

The switching removes the loop of multiplicity 3 and the simple edges $\{v_j, w_j\}$ for $j = 1, 2, 3$ and adds the six simple edges $\{\{v_0, v_j\}, \{v_0, w_j\} \mid j = 1, 2, 3\}$, as shown in Figure 3.7.

Now let R be a multigraph which can be formed from Q by a switching of color 10. The number of ways that a switching of color 10 can be performed in Q is at least

$$\ell_3(Q)\,\big([2e_1(Q)]_3 - O(k_{\max}^2 e_1(Q)^2)\big) \geq \tfrac{1}{2}\, \ell_3(Q)M^3,$$

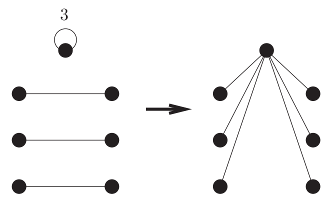

FIG. 3.7. *A switching of color* 10.

using Lemma 3.1, and the number of ways that R can be produced using a switching of color 10 is bounded above by M_6. Hence we can take $a_{10}(Q) = \frac{1}{2}\,\ell_3(Q)\,M^3$ and $b_{10}(R) = M_6$. This leads to

$$\alpha(Q,R) = \frac{2\,M_6}{\ell_3(Q)\,M^3} = O\biggl(\frac{k_{\max}^5}{M^2}\biggr) = o\biggl(\frac{k_{\max}^3}{M}\biggr).$$

We do not tackle single loops yet. Instead, the next two switchings reduce nonloop edges of high multiplicity down to zero.

3.9. Switchings of color 11. Switchings of color 11 reduce the number of nonloop edges with multiplicity 4 or multiplicity at least 7 until this number is zero. Switchings of color 11 are applied to multigraphs Q for which switchings of colors $1,\ldots,10$ do not apply and $E(Q) \ge 1$, where

$$E(Q) = e_4(Q) + \sum_{D\ge 7} e_D(Q).$$

Let Q be such a multigraph. A switching of color 11 in Q is defined by a sequence

$$(v_0, w_0, v_1, w_1, v_2, w_2, v_3, w_3)$$

of distinct vertices such that

- the edge $\{v_0, w_0\}$ has multiplicity D in Q, where $D = 4$ or $D \ge 7$,
- edges $\{v_j, w_j\}$ are simple edges in Q (with multiplicity 1) for $j = 1, 2, 3$,
- the edges $\{v_0, v_j\}$ and $\{w_0, w_j\}$ are nonedges in Q for $j = 1, 2, 3$ (with multiplicity zero).

The switching reduces the multiplicity of $\{v_0, w_0\}$ to $D-3$, removes the edges $\{v_j, w_j\}$ for $j = 1, 2, 3$, and adds the simple edges $\{v_0, v_j\}$ and $\{w_0, w_j\}$ for $j = 1, 2, 3$. This operation is illustrated in Figure 3.8.

Suppose that the multigraph R can be obtained from Q by performing a switching of color 11. The number of ways to perform a switching of color 11 in Q is at least

$$2E(Q)\,\bigl([2e_1(Q)]_3 - O(k_{\max}^2 e_1(Q)^2)\bigr) \ge E(Q)\,M^3$$

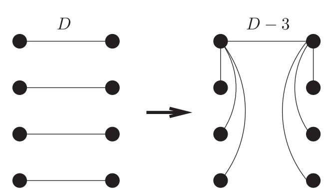

FIG. 3.8. *A switching of color* 11.

using Lemma 3.1, and the number of ways to produce R using a switching of color 11 is at most $k_{\max}^3 M_4$. Hence we can let $a_{11}(Q) = E(Q)\,M^3$ and $b_{11}(R) = k_{\max}^3 M_4$. This leads to

$$\alpha(Q,R) = \frac{k_{\max}^3 M_4}{E(Q) M^3} = O\biggl(\frac{k_{\max}^6}{M^2}\biggr) = o\biggl(\frac{k_{\max}^3}{M}\biggr).$$

3.10. Switchings of color 12. Switchings of color 12 reduce the number of nonloop edges with multiplicity 5 or 6, until this number is zero. They are applied to multigraphs Q such that switchings of colors $1, \dots, 11$ do not apply and $e_5(Q) + e_6(Q) \geq 1$.

Let Q be such a multigraph. Then a switching of color 12 is defined by a sequence of distinct vertices $(v_0, w_0, v_1, w_1, v_2, w_2, v_3, w_3, v_4, w_4, v_5, w_5)$ such that

- the edge $\{v_0, w_0\}$ has multiplicity D in Q, where $D \in \{5, 6\}$,
- each edge $\{v_j, w_j\}$ is a simple edge in Q with multiplicity 1, for $j = 1, 2, 3, 4, 5$,
- the edges $\{v_0, v_j\}$ and $\{w_0, w_j\}$ all have multiplicity 0 in Q for $j = 1, 2, 3, 4, 5$ (that is, they are nonedges).

The switching reduces the multiplicity of the edge $\{v_0, w_0\}$ to $D - 5$, removes the edges $\{v_j, w_j\}$, $j = 1, 2, 3, 4, 5$, and inserts the simple edges $\{v_0, v_j\}$, $\{v_0, w_j\}$ for $j = 1, 2, 3, 4, 5$, as shown in Figure 3.9.

Suppose that the multigraph R can be formed from Q by a switching of color 12. The number of ways that a switching of color 12 can be performed in Q is at least

$$2(e_5(Q) + e_6(Q))\,\bigl([2e_1(Q)]_5 - O(k_{\max}^2 e_1(Q)^4)\bigr) \geq (e_5(Q) + e_6(Q))\,M^5$$

using Lemma 3.1, and the number of ways that R can be produced by a switching of color 12 is at most M_5^2. Hence we can set $a_{12}(Q) = (e_5(Q) + e_6(Q))M^5$ and $b_{12}(R) = M_5^2$. This gives

$$\alpha(Q,R) = \frac{M_5^2}{(e_5(Q) + e_6(Q))M^5} = O\biggl(\frac{k_{\max}^8}{M^3}\biggr) = o\biggl(\frac{k_{\max}^3}{M}\biggr).$$

3.11. Switchings of color 13. Switchings of color 13 are used to reduce the number of simple loops until this number is at most $\lceil N_1/2 \rceil$. They are applied to multigraphs Q for which switchings of colors $1, \dots, 12$ do not apply and $\ell_1(Q) > \lceil N_1/2 \rceil$.

Let Q be such a multigraph. Then a switching of color 13 is defined by a sequence of distinct vertices (v_0, v_1, v_2) such that there is a simple loop on v_0 in Q, the edge $\{v_1, v_2\}$ is a simple edge in Q, and the edges $\{v_0, v_1\}$, $\{v_0, v_2\}$ are both absent in Q

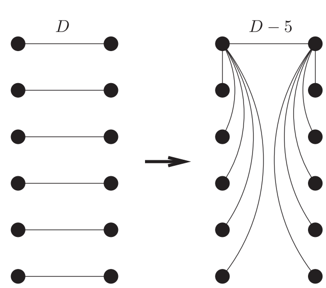

FIG. 3.9. *A switching of color* 12.

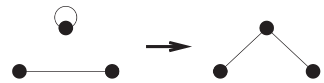

FIG. 3.10. *A switching of color* 13.

(that is, they have multiplicity zero). The switching removes the simple loop at v_0 and the simple edge $\{v_1, v_2\}$ and inserts the two simple edges $\{v_0, v_1\}$, $\{v_0, v_2\}$, as shown in Figure 3.10.

Suppose that the multigraph R can be obtained from Q by performing a switching of color 13. The number of ways that a switching of color 13 can be performed in Q is at least

$$\ell_1(Q)\big(2e_1(Q) - O(k_{\max}^2)\big) \geq \tfrac{1}{2}\ell_1(Q)M$$

using Lemma 3.1, and the number of ways that a switching of color 13 can produce R is at most M_2. Therefore we can define $a_{13}(Q) = \frac{1}{2}\ell_1(Q)M$ and $b_{13}(R) = M_2$, leading to

$$\alpha(Q, R) = \frac{2\,M_2}{\ell_1(Q)M} \leq \frac{4\,M_2}{N_1 M} \leq \frac{1}{120}$$

using the definition of N_1.

3.12. Switchings of color 14. Switchings of color 14 reduce the number of nonloop edges with multiplicity 2 until this number is at most $\lceil N_2/2\rceil$. This switching is applied to multigraphs Q such that switchings of colors $1, \ldots, 13$ do not apply and $e_2(Q) > \lceil N_2/2\rceil$.

Let Q be such a multigraph. A switching of color 14 in Q is described by a sequence $(v_0, w_0, v_1, w_1, v_2, w_2)$ of distinct vertices such that

- $\{v_0, w_0\}$ is an edge of multiplicity 2 in Q,
- $\{v_1, w_1\}$ and $\{v_2, w_2\}$ are simple edges in Q (with multiplicity 1),
- none of the edges $\{v_0, v_j\}$ or $\{w_0, w_j\}$ are present in Q, for $j = 1, 2$ (these are all nonedges with multiplicity zero).

The switching removes these three edges and replaces them with two copies of $K_{1,2}$, one on $\{v_0, v_1, v_2\}$ centered at v_0 and the other on $\{w_0, w_1, w_2\}$ centered at w_0. This operation is shown in Figure 3.11.

Suppose that the multigraph R can be produced from Q using a switching of color 14. The number of ways to perform a switching of color 14 in Q is at least

$$2e_2(Q)\big([2e_1(Q)]_2 - O(k_{\max}^2 e_1(Q))\big) \geq e_2(Q)M^2$$

using Lemma 3.1, and the number of ways that R can be produced using a switching of color 14 is at most M_2^2. Therefore we can take $a_{14}(Q) = e_2(Q)M^2$ and $b_{14}(R) = M_2^2$, leading to

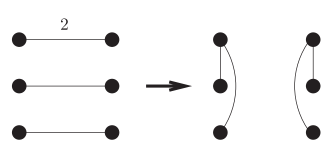

FIG. 3.11. *A switching of color* 14.

$$\alpha(Q,R) = \frac{M_2^2}{e_2(Q)M^2} < \frac{2M_2^2}{N_2\,M^2} \le \frac{1}{120}$$

by the definition of N_2.

3.13. Switchings of color 15. Switchings of color 15 are used to reduce the number of nonloop edges with multiplicity 3 until this number is at most $\lceil N_3/2\rceil$. They are applied to multigraphs Q such that switchings of colors $1,\dots,14$ do not apply and $e_3(Q) > \lceil N_3/2\rceil$.

Let Q be such a multigraph. Then a switching of color 15 is defined by a sequence of distinct vertices $(v_0, w_0, v_1, w_1, v_2, w_2, v_3, w_3)$ such that

- $\{v_0, w_0\}$ has multiplicity 3 in Q,
- each of $\{v_j, w_j\}$ is a simple edge in Q for $j = 1, 2, 3$,
- the edges $\{v_0, v_j\}$ and $\{w_0, w_j\}$ are absent in Q for $j = 1, 2, 3$.

The switching removes the edges $\{v_j, w_j\}$ for $j = 0, 1, 2, 3$ (setting the multiplicity of each to zero) and inserts the simple edges $\{v_0, v_j\}$, $\{w_0, w_j\}$ for $j = 1, 2, 3$, as illustrated in Figure 3.8 for the case that $D = 3$.

Suppose that the multigraph R can be produced by performing a switching of color 15 from Q. The number of ways to perform a switching of color 15 in Q is at least

$$2e_3(Q)\left([2e_1(Q)]_3 - O(k_{\max}^2 e_1(Q)^2)\right) \ge e_3(Q)M^3$$

using Lemma 3.1, while the number of ways that R could be produced using a switching of color 15 is at most M_3^2. Therefore we can take $a_{15}(Q) = e_3(Q)M^3$ and $b_{15}(R) = M_3^2$, leading to

$$\alpha(Q,R) = \frac{M_3^2}{e_3(Q)M^3} \le \frac{2\,M_3^2}{N_3\,M^3} \le \frac{1}{120}$$

by the definition of N_3.

3.14. Analysis of the switchings. We now explain how to apply Theorem 2.1 to analyze these switchings on $\mathcal{G}_0$. (See the statement of Theorem 2.1 for the necessary notation.)

We now define a directed graph Γ with vertex set $V(\Gamma) = \{v_Q : Q \in \mathcal{G}_0\}$, where v_Q is associated with the set $\mathcal{S}(v_Q) = \{Q\}$ containing one object. These sets are certainly disjoint. By a slight abuse of notation we will identify v_Q and Q from now on and write Q for both. The edge set of Γ is defined as follows: there is an edge QR in Γ with color $c \in \{1,\dots,15\}$ if and only if R can be obtained from Q using a switching of color c. Since a switching is used only where no switching with a lower-labeled color applies, there is at most one edge from Q to R in Γ. Hence the endvertices of an edge uniquely determine the color of the edge. We take $\lambda_c(v) = \frac{1}{15}$ for all c, v, and so for each edge QR in Γ we have

$$\hat{\alpha}(QR) = \frac{\alpha(QR)}{\lambda_{c(QR)}(R)} = 15\,\alpha(QR).$$

Let Y, Z be as defined in (3.2), (3.3).

LEMMA 3.2. *With notation as established above, we have*

$$G(\boldsymbol{k}, J, J^*) = \left(1 + O(k_{\max}^3/M)\right)|\mathcal{G}_0 - Y|.$$

Proof. It follows from the analysis of the previous sections that $\alpha(QR) \le \frac{1}{30}$, and therefore $\hat\alpha(QR) \le \frac{1}{2}$, for all edges QR in $\varGamma$. Moreover, we have provided at least one switching which can be performed from Q for each graph Q in $\mathcal{G}_0 - Z$. Hence Y and Z satisfy the requirements of Theorem 2.1 and we can conclude that

$$|Y| \le \frac{\hat\alpha(Y,Z)}{1-\hat\alpha(Y,Y)}\,|Z| \le \frac{\hat\alpha(Y,Z)}{1-\hat\alpha(Y,Y)}\,|\mathcal{G}_0|,$$

where

$$\hat\alpha(W,W') = \max_{P\in\mathcal{P}(W,W')} \hat\alpha(P)$$

for all $W, W' \subseteq \mathcal{G}_0$.

Since $\mathcal{P}(Y,Y)$ by definition has only nontrivial paths, and $\hat\alpha(QR) \le \frac{1}{2}$ for all edges in $\varGamma$, we know that $\hat\alpha(Y,Y) \le \frac{1}{2}$, and therefore

$$|Y| \le 2\,\hat\alpha(Y,Z)\,|\mathcal{G}_0|.$$

Now let P be a path in $\varGamma$ from some $Q \in Y$ to some element of Z, such that all internal vertices of P belong to $\mathcal{G}_0 - (Y \cup Z)$. Let QR be the first edge in P. Then QR cannot have color in $\{1,\dots,8\}$, since these switchings only produce graphs in Y. If QR is colored with a color in $\{9,10,11,12\}$, then $\hat\alpha(QR) = O(k_{\max}^3/M)$, so $\hat\alpha(P) = O(k_{\max}^3/M)$.

The remaining possibility is that QR has color $12+r$ for some $r \in \{1,2,3\}$, in which case P must contain at least $\lfloor N_r/2\rfloor$ edges of color $12+r$. Our analysis showed that $\alpha(e) \le \frac{1}{120}$ and hence that $\hat\alpha(e) \le \frac{1}{8}$ for all such edges e. By definition $\lfloor N_r/2\rfloor \ge \frac{1}{2}(\log M - 1)$ for $r \in \{1,2,3\}$. Therefore,

$$\hat\alpha(P) \le 8^{-(\log M-1)/2} = O(M^{-1}),$$

which implies that $\hat\alpha(Y,Z) = O(k_{\max}^3/M)$. We conclude that

$$|\mathcal{G}_0 - Y| = \bigl(1 + O(k_{\max}^3/M)\bigr)|\mathcal{G}_0|.$$

But $\mathcal{G}_0 - Y \subseteq \mathcal{G}(\boldsymbol{k}, J, J^*) \subseteq \mathcal{G}_0$, and hence

$$G(\boldsymbol{k}, J, J^*) = \bigl(1 + O(k_{\max}^3/M)\bigr)|\mathcal{G}_0 - Y|,$$

completing the proof. ∎

It remains to obtain an asymptotic expression for $\mathcal{G}_0 - Y$, which we do in the next section.

4. From pairings to multigraphs. In this section we work in the *pairing model* (also called the *configuration model*), which we now describe. This model is standard for working with random graphs of fixed degrees; see, for example, [11]. Consider a set of M *points* partitioned into *cells* $c_1,\dots,c_n$, where cell c_i contains k_i points for $i = 1,2,\dots,n$. Take a partition P (called a *pairing*) of the M points into $M/2$ *pairs* with each pair having the form $\{y,z\}$, where $y \in c_i$ and $z \in c_j$ for some i,j. The set of all such pairings, of which there are $M!/(M/2)!\,2^{M/2}$, will be denoted by $\mathcal{C}(\boldsymbol{k})$.

A *loop* is a pair whose two points lie in the same cell, while a *link* is a pair involving two distinct cells. Two pairs are *parallel* if they involve the same cells. A *parallel class* is a maximal set of mutually parallel pairs. The *multiplicity* of a parallel class (and of the pairs in the class) is the cardinality of the class. As important special cases, a *simple pair* is a parallel class of multiplicity 1, a *double pair* is a parallel class of multiplicity 2, and a *triple pair* is a parallel class of multiplicity 3.

Each pairing gives rise to a multigraph in $\mathcal{G}(\boldsymbol{k})$ by replacing each cell by a vertex and letting the multiplicity of the edge $\{v, w\}$ equal the multiplicity of the parallel class between the corresponding cells.

Let $\mathcal{C}_{\ell,d,t}$ be the set of all pairings in $\mathcal{C}(\boldsymbol{k})$ with exactly ℓ simple loops, exactly d double pairs, and exactly t triple pairs, but with no loops of multiplicity greater than one and no links of multiplicity greater than three. If $G \in \mathcal{G}_0 - Y$ can be formed from a pairing $P \in \mathcal{C}_{\ell,d,t}$, then exactly

$$2^{-(\ell+d)}\, 6^{-t} \prod_{i=1}^{n} k_i!$$

pairings in $\mathcal{C}(\boldsymbol{k})$ give rise to G. Now defining

$$w(\ell, d, t) = 2^{\ell+d}\, 6^t\, |\mathcal{C}_{\ell,d,t}|,$$

we can write

$$|\mathcal{G}_0 - Y| = \left(\prod_{i=1}^{n} k_i!\right)^{-1} \sum_{\ell=0}^{N_1} \sum_{d=0}^{N_2} \sum_{t=0}^{N_3} w(\ell, d, t). \tag{4.1}$$

Hence it suffices to obtain an asymptotic expression for the above sum.

We will need the following two summation lemmas adapted from [8].

LEMMA 4.1 (see [8, Corollary 4.3]). *Let $0 \le A_1 \le A_2$ and $B_1 \le B_2$ be real numbers. Suppose that there exist integers N, K with $N \ge 2$ and $0 \le K \le N$, and a real number $c > 2e$ such that $Ac < N - K + 1$ and $|BN| < 1$ for all $A \in [A_1, A_2]$ and $B \in [B_1, B_2]$. Further suppose that there are real numbers δ_i, for $1 \le i \le N$, and $\gamma_i \ge 0$, for $0 \le i \le K$, such that $\sum_{j=1}^{i} |\delta_j| \le \sum_{j=0}^{K} \gamma_j [i]_j < \frac{1}{5}$ for $1 \le i \le N$.*

Given $A(1), \ldots, A(N) \in [A_1, A_2]$ and $B(1), \ldots, B(N) \in [B_1, B_2]$, define $n_0, n_1, \ldots, n_N$ by $n_0 = 1$ and

$$n_i = \frac{1}{i} A(i)\bigl(1 - (i-1)B(i)\bigr)\bigl(1 + \delta_i\bigr)\, n_{i-1}$$

for $1 \le i \le N$. Then

$$\Sigma_1 \le \sum_{i=0}^{N} n_i \le \Sigma_2,$$

where

$$\Sigma_1 = \exp\left(A_1 - \tfrac{1}{2}A_1^2 B_2 - 4\sum_{j=0}^{K} \gamma_j (3A_1)^j\right) - \tfrac{1}{4}(2e/c)^N,$$

$$\Sigma_2 = \exp\left(A_2 - \tfrac{1}{2}A_2^2 B_1 + \tfrac{1}{2}A_2^3 B_1^2 + 4\sum_{j=0}^{K} \gamma_j (3A_2)^j\right) + \tfrac{1}{4}(2e/c)^N. \qquad \square$$

LEMMA 4.2 (see [8, Corollary 4.5]). *Let $N \ge 2$ be an integer and, for $1 \le i \le N$, let real numbers $A(i)$, $C(i)$ be given such that $A(i) \ge 0$ and $A(i) - (i-1)C(i) \ge 0$. Define*

$A_1 = \min_{i=1}^N A(i)$, $A_2 = \max_{i=1}^N A(i)$, $C_1 = \min_{i=1}^N C(i)$, and $C_2 = \max_{i=1}^N C(i)$. Suppose that there exists a real number $\hat{c}$ with $0 < \hat{c} < \frac{1}{3}$ such that $\max\{A/N, |C|\} \le \hat{c}$ for all $A \in [A_1, A_2]$, $C \in [C_1, C_2]$. Define $n_0, \ldots, n_N$ by $n_0 = 1$ and

$$n_i = \frac{1}{i}\bigl(A(i) - (i-1)C(i)\bigr)\, n_{i-1}$$

for $1 \le i \le N$. Then

$$\Sigma_1 \le \sum_{i=0}^{N} n_i \le \Sigma_2,$$

where

$$\begin{aligned}\Sigma_1 &= \exp\bigl(A_1 - \tfrac{1}{2}A_1C_2\bigr) - (2e\hat{c})^N,\\ \Sigma_2 &= \exp\bigl(A_2 - \tfrac{1}{2}A_2C_1 + \tfrac{1}{2}A_2C_1^2\bigr) + (2e\hat{c})^N. \qquad \square\end{aligned}$$

In the proofs of Lemmas 4.3–4.5, we will use several results which were proved in [17] (specifically, Lemmas 4.1–4.4 of that paper and some details of their proofs). That paper uses values of N_1, N_2, N_3 that differ from ours, but only by bounded factors, and examination of the proofs in [17] shows that the results we wish to apply remain valid when our values of N_1, N_2, N_3 are used.

First we perform a summation over the number of edges of multiplicity 3.

LEMMA 4.3. *Uniformly for $0 \le d \le N_2$ and $0 \le \ell \le N_1$, we have*

$$\sum_{t=0}^{N_3} w(\ell, d, t) = w(\ell, d, 0)\, \exp\biggl(\frac{M_3^2}{2M^3} + O\bigl(k_{\max}^3/M\bigr)\biggr).$$

Proof. We will apply the proof of [17, Lemma 4.1].

Let t' be the first value of $t \le N_3$ for which $\mathcal{C}_{\ell,d,t} = \emptyset$, or $t' = N_3 + 1$ if there is no such value. In [17, Lemma 4.1], a switching is described that converts any pairing in $\mathcal{C}_{\ell,d,t}$ to at least one in $\mathcal{C}_{\ell,d,t-1}$ for $1 \le t \le N_3$, so we know that $w(\ell, d, t) = 0$ for $t' \le t \le N_3$. In particular, the present lemma is true when $w(\ell, d, 0) = 0$, so we assume that $t' \ge 1$.

Noting that

$$\frac{w(\ell, d, t)}{w(\ell, d, t-1)} = \frac{6\,|\mathcal{C}_{\ell,d,t}|}{|\mathcal{C}_{\ell,d,t-1}|}$$

when the denominators are nonzero, the calculation in [17, Lemma 4.1] shows that there is some uniformly bounded function $\alpha_t = \alpha_t(\ell, d)$ such that

$$\frac{w(\ell, d, t)}{w(\ell, d, 0)} = \frac{1}{t}\, \frac{w(\ell, d, t-1)}{w(\ell, d, 0)}\, \bigl(A(t) - (t-1)C(t)\bigr) \tag{4.2}$$

for $1 \le t \le N_3$, where

$$A(t) = \frac{M_3^2 - \alpha_t k_{\max}^2 (k_{\max}^2 + \ell + d) M_3}{2M^3}, \quad C(t) = \frac{\alpha_t k_{\max}^2 M_3}{2M^3}$$

for $1 \le t < t'$ and $A(t) = C(t) = 0$ for $t \ge t'$.

Now we can apply Lemma 4.2. It is clear that $A(t) - (t-1)C(t) \geq 0$ by (4.2). If $\alpha_t \geq 0$, then $A(t) \geq A(t) - (t-1)C(t) \geq 0$, while if $\alpha_t < 0$, then the definition of $A(t)$ makes it evidently nonnegative. Now define A_1, A_2, C_1, C_2 by taking the minimum and maximum of $A(t)$ and $C(t)$ over $1 \leq t \leq N_3$. Let $A \in [A_1, A_2]$ and $C \in [C_1, C_2]$, and set $\hat{c} = \frac{1}{80}$. Since $A = M_3^2/2M^3 + o(1)$ and $C = o(1)$, we have that $\max\{A/N_3, |C|\} < \hat{c}$ for M sufficiently large, by the definition of N_3.

Therefore Lemma 4.2 applies and gives an upper bound

$$\sum_{t=0}^{N_3} \frac{w(\ell,d,t)}{w(\ell,d,0)} \leq \exp\biggl(\frac{M_3^2}{2M^3} + O\bigl(k_{\max}^4(k_{\max}^2 + \ell + d)/M^2\bigr)\biggr) + O\bigl((e/40)^{N_3}\bigr).$$

Since $\ell + d \leq N_1 + N_2 = O(k_{\max}^2 + \log M)$ and $(e/40)^{N_3} \leq (e/40)^{\log M} \leq M^{-2}$,

$$\sum_{t=0}^{N_3} \frac{w(\ell,d,t)}{w(\ell,d,0)} \leq \exp\biggl(\frac{M_3^2}{2M^3} + O\bigl(k_{\max}^3/M\bigr)\biggr).$$

In the case that $t' = N_3 + 1$, the lower bound given by Lemma 4.2 is the same within the stated error term, so we are done.

This leaves the case $1 \leq t' \leq N_3$. By the counts of the second switching in the proof of [17, Lemma 4.1], $|\mathcal{C}_{\ell,d,t}| = 0$ is possible only if $M_3 = O(k_{\max}^2(k_{\max}^2 + \ell + d + t))$. If this happens for $t \leq N_3$ we have

$$M_3 = O(k_{\max}^2(k_{\max}^2 + N_1 + N_2 + N_3)) = O(k_{\max}^2(k_{\max}^2 + \log M)).$$

However, this implies that $M_3^2/M^3 = O(k_{\max}^3/M)$ so the upper bound matches the trivial lower bound 1 within the error term. This completes the proof. ∎

Next we perform a summation over the number of simple loops.

LEMMA 4.4. *Uniformly for $0 \leq d \leq N_2$, we have*

$$\sum_{\ell=0}^{N_1} w(\ell,d,0) = w(0,d,0)\,\exp\biggl(\frac{M_2}{M} + O\biggl(\frac{k_{\max}d}{M} + \frac{k_{\max}^3}{M}\biggr)\biggr).$$

Proof. Let ℓ' be the first value of $\ell \leq N_1$ for which $\mathcal{C}_{\ell,d,0} = \emptyset$, or $\ell' = N_1 + 1$ if there is no such value. In the proof of [17, Lemma 4.2], a switching is described that converts any pairing in $\mathcal{C}_{\ell,d,0}$ to at least one in $\mathcal{C}_{\ell-1,d,0}$ for $1 \leq \ell \leq N_1$, so we know that $w(\ell,d,0) = 0$ for $\ell' \leq \ell \leq N_1$. In particular, the present lemma is true when $w(0,d,0) = 0$, so we assume that $\ell' \geq 1$.

Noting that

$$\frac{w(\ell,d,0)}{w(\ell-1,d,0)} = \frac{2\,|\mathcal{C}_{\ell,d,0}|}{|\mathcal{C}_{\ell-1,d,0}|}$$

when the denominators are nonzero, the calculation in [17, Lemma 4.2] shows that there is some uniformly bounded function $\beta_\ell = \beta_\ell(d)$ such that

$$\frac{w(\ell,d,0)}{w(0,d,0)} = \frac{1}{\ell}\,\frac{w(\ell-1,d,0)}{w(0,d,0)}\,\bigl(A(\ell) - (\ell-1)C(\ell)\bigr)$$

for $1 \leq t \leq N_1$, where

$$A(\ell) = \frac{M_2 - \beta_\ell(k_{\max}^3 + k_{\max}d)}{M}, \quad C(\ell) = \frac{\beta_\ell k_{\max}^2}{M}$$

for $1 \leq \ell < \ell'$ and $A(\ell) = C(\ell) = 0$ for $\ell \geq \ell'$.

We can now complete the proof using $\hat{c} = \frac{1}{80}$ and following the argument used in the previous lemma. The treatment of the lower bound when $\ell' \le N_1$ needs some additional care. From the analysis of the second switching used in [17, Lemma 4.2], $\ell' \le N_1$ can happen only if $M_2 = O(k_{\max}d + k_{\max}^2\ell')$. For $\ell' = 1$, the trivial lower bound of 1 matches the upper bound within the stated error terms. If $2 \le \ell' \le N_1$ then the first two terms of the summation give the lower bound

$$w(0,d,0)(1 + A(1)) = w(0,d,0)\exp\biggl(\frac{M_2}{M} + O\biggl(\frac{k_{\max}d}{M} + \frac{k_{\max}^3}{M}\biggr)\biggr),$$

which matches the upper bound within the stated error terms. In either case, the proof is complete. ∎

Next we perform a summation over the number of edges of multiplicity 2. The exponential factor in the summand corresponds to the error term from Lemma 4.4 which depends on d.

LEMMA 4.5. *For any constant ρ have*

$$\sum_{d=0}^{N_2} w(0,d,0)\exp\biggl(\frac{\rho k_{\max}d}{M}\biggr) = w(0,0,0)\,\exp\biggl(\frac{M_2^2}{2M^2} - \frac{M_3^2}{2M^3} + O\bigl(k_{\max}^3/M\bigr)\biggr).$$

Proof. Let d' be the first value of $d \le N_2$ for which $\mathcal{C}_{0,d,0} = \emptyset$, or $d' = N_2 + 1$ if no such value of d exists. As in the previous two lemmas, the switchings described in [17] (see also [16, Lemma 4]) show that $w(0,d,0) = 0$ for $d' \le d \le N_2$. They also show that $d' \le N_2$ is possible only if $M_2 = O(k_{\max}^3 + k_{\max}d)$. In particular, the lemma is true if $w(0,0,0) = 0$, so we assume that $d' \ge 1$.

We divide the proof into two cases, following the division used in [17, Lemmas 4.3–4.4]. For $0 \le d \le N_2$, define

$$m_d = w(0,d,0)\,\exp\biggl(\frac{\rho k_{\max}d}{M}\biggr).$$

First suppose that $M_2 \le M$. From [17, Lemma 4.3] we know that for $1 \le d < d'$,

$$\frac{m_d}{m_{d-1}} = \frac{2\,\exp(\rho k_{\max}/M)\,|\mathcal{C}_{0,d,0}|}{|\mathcal{C}_{0,d-1,0}|} = \frac{M_2^2}{2dM^2} + O\biggl(\frac{k_{\max}(k_{\max}^2 + d)M_2}{dM^2}\biggr).$$

The current lemma follows from this, arguing as in the proofs of the previous two lemmas. (Note that the term $M_3^2/2M^3$ in the answer is absorbed into the error term by the assumption that $M_2 \le M$.) In the case that $d' \le N_2$, the condition $M_2 = O(k_{\max}^3 + k_{\max}d)$ implies that 1 is a sufficient lower bound.

Now suppose that $M_2 > M$. In this case $M_2 = O(k_{\max}^3 + k_{\max}d)$ is not possible for $d \le N_2$, so $d' = N_2 + 1$ and the series contains only positive terms. By [17, Lemma 4.4] we have for $1 \le d \le N_2$,

$$m_d = m_{d-1}\,\frac{A(d)}{d}\bigl(1 - (d-1)B(d)\bigr)\,(1 + \delta_d),$$

where

$$\begin{aligned}
A(d) &= \frac{M_2^2}{2M^2} + \frac{2M_2^2M_3}{M^4} - \frac{M_3^2}{2M^3} - \frac{M_2^4}{M^5} + O(k_{\max}^3/M),\\
B(d) &= B = -\frac{8}{M} + \frac{16M_3}{M_2^2},\\
\delta_d &= O\bigl((d-1)/M_2\bigr) \quad \text{uniformly for } 1 \le d \le N_2.
\end{aligned}$$

(Note that $A(d)\exp(\rho k_{\max}/M)$ equals $A(d)$ within the precision afforded by the error term, since $k_{\max}A(d)/M = O(k_{\max}M_2^2/M^3) = O(k_{\max}^3/M)$.)

Now we will apply Lemma 4.1 to $\sum_d m_d$. Let $A_1 = \min_d A(d)$ and $A_2 = \max_d A(d)$ with the minimum and maximum taken over $1 \le d \le N_2$, and let $B_1 = B_2 = B$. Define $c = 80$ and $K = 2$. Then the conditions of Lemma 4.1 apply with $\gamma_0 = \gamma_1 = 0$ and some value of γ_2 which satisfies $\gamma_2 = O(1/M_2)$. Application of that lemma now gives the desired result. ∎

Recall the definition of y_1, x_2, $x_3 \in \{0,1\}$ given just before the statement of Theorem 1.1. By combining the last three summations we obtain the following.

LEMMA 4.6. *As $M \to \infty$ and $k_{\max}^3 = o(M)$, we have*

$$\sum_{d=0}^{x_2N_2}\sum_{\ell=0}^{y_1N_1}\sum_{t=0}^{x_3N_3} w(\ell,d,t) = w(0,0,0)\,\exp\biggl(y_1\frac{M_2}{M} + x_2\frac{M_2^2}{2M^2} + (x_3-x_2)\frac{M_3^2}{2M^3} + O\bigl(k_{\max}^3/M\bigr)\biggr).$$

Proof. From Lemmas 4.3 and 4.4, we have

$$\sum_{\ell=0}^{y_1N_1}\sum_{t=0}^{x_3N_3} w(\ell,d,t) = w(0,d,0)\,\exp\biggl(y_1\frac{M_2}{M} + x_3\frac{M_3^2}{2M^3} + O\biggl(\frac{k_{\max}d}{M} + \frac{k_{\max}^3}{M}\biggr)\biggr) \tag{4.3}$$

uniformly over d.

If $x_2 = 0$, we are finished. If $x_2 = 1$, choose two constants ρ_1, ρ_2 such that the actual value of the error term $O(k_{\max}d/M)$ in (4.3) lies in $[\rho_1 k_{\max}d/M, \rho_2 k_{\max}d/M]$ for all $d \in \{1,\ldots,N_2\}$. Applying Lemma 4.5 with $\rho \in \{\rho_1,\rho_2\}$ and noting that the result does not depend on ρ within the precision given by the error term, we are done. ∎

Finally we can prove our main result.

Proof of Theorem 1.1. From Lemmas 3.2 and 4.6 together with (4.1), we obtain

$$\begin{aligned} G(\boldsymbol{k},J,J^*) &= \bigl(1+O(k_{\max}^3/M)\bigr)\,|\mathcal{G}_0 - Y| \\ &= \frac{1+O\bigl(k_{\max}^3/M\bigr)}{k_1!\cdots k_n!}\sum_{d=0}^{x_2N_2}\sum_{\ell=0}^{y_1N_2}\sum_{t=0}^{x_3N_3} w(\ell,d,t) \\ &= \frac{w(0,0,0)}{k_1!\cdots k_n!}\,\exp\biggl(y_1\frac{M_2}{M} + x_2\frac{M_2^2}{2M^2} + (x_3-x_2)\frac{M_3^2}{2M^3} + O\bigl(k_{\max}^3/M\bigr)\biggr). \end{aligned}$$

Now $w(0,0,0) = |\mathcal{C}_{0,0,0}|$, and it follows from [17, Lemma 5.1] that

$$|\mathcal{C}_{0,0,0}| = \frac{M!}{(M/2)!\,2^{M/2}}\,\exp\biggl(-\frac{M_2}{2M} - \frac{M_2^2}{4M^2} - \frac{M_2^2M_3}{2M^4} + \frac{M_2^4}{4M^5} + \frac{M_3^2}{6M^3} + O(k_{\max}^3/M)\biggr).$$

Combining these two expressions completes the proof. ∎

5. Comparison to a naïve model. Let $p \in (0,1)$ be a probability which we will define later. Abusing notation slightly, we define the functions

$$J(z) = \sum_{j\in J} p^j z^j, \quad J^*(z) = \sum_{j\in J^*} p^j z^j$$

and the probability generating functions

$$f(z) = \frac{J(z)}{J(1)}, \quad g(z) = \frac{J^*(z)}{J^*(1)}.$$

Consider a random symmetric nonnegative integer matrix $A = (a_{ij})$ which is created as follows:

- All entries on and above the diagonal are independent.
- For $i = 1, \ldots, n$, let a_{ii} be a randomly chosen element of J^*, with $\Pr(a_{ii} = b) = p^b/J^*(1)$ for all $b \in J^*$.
- For $1 \le i < j \le n$ let a_{ij} be a randomly chosen element of J, with $\Pr(a_{ij} = b) = p^b/J(1)$ for all $b \in J$.
- Finally, let $a_{ji} = a_{ij}$ for $1 \le i < j \le n$.

We write $\Pr(\cdot)$ to mean probabilities generated by the above procedure. If A_0 is a fixed matrix which corresponds to an element of $\mathcal{G}(\boldsymbol{k}, J, J^*)$, then the probability that A_0 is produced by the above procedure is

$$\Pr(A_0) = \frac{p^{M/2}}{J(1)^{\binom{n}{2}} J^*(1)^n}. \tag{5.1}$$

This follows as the probability of A_0 depends only on the sum of the diagonal and above-diagonal entries (that is, on the number of edges of the corresponding multigraph).

For $i = 1, \ldots, n$, let $\mathcal{E}_i$ denote the event that the sum of row i equals k_i, recalling that diagonal entries are weighted by a factor of 2 as in (1.1). Now

$$G(\boldsymbol{k}, J, J^*) = \frac{\Pr(\mathcal{E}_1 \wedge \mathcal{E}_2 \wedge \cdots \wedge \mathcal{E}_n)}{\Pr(A_0)}$$

since $\Pr(\cdot)$ is uniform on (the set of matrices corresponding to) $\mathcal{G}(\boldsymbol{k}, J, J^*)$. By (incorrectly) assuming that the events $\mathcal{E}_1, \ldots, \mathcal{E}_n$ are independent, we obtain the following naïve estimate of $G(\boldsymbol{k}, J, J^*)$, using (5.1):

$$\begin{aligned} G_p(\boldsymbol{k}, J, J^*) &= \frac{\prod_{i=1}^n \Pr(\mathcal{E}_i)}{\Pr(A_0)} \\ &= \frac{J(1)^{\binom{n}{2}} J^*(1)^n}{p^{M/2}} \prod_{i=1}^n [z^{k_i}]\, f(z)^{n-1} g(z^2) \\ &= p^{-M/2}\, J(1)^{-\binom{n}{2}} \prod_{i=1}^n [z^{k_i}]\, J(z)^{n-1} J^*(z^2). \end{aligned} \tag{5.2}$$

Let $k = M/n$ denote the target average row sum. The value p_0 of p which makes the expected row sum equal to k satisfies

$$p_0 = \frac{k}{n} + \frac{(1 - 2x_2)k^2}{n^2} + O(k/n^2). \tag{5.3}$$

We write $G_{\text{naive}}(\boldsymbol{k}, J, J^*)$ for $G_{p_0}(\boldsymbol{k}, J, J^*)$ from now on. (As we shall see, the exact value of the $O(k/n^2)$ term does not affect Theorem 5.1 below, within the stated error bound, even though some of the intermediate formulae we give in the proof are affected.)

We now compare the expression given in Theorem 1.1 with the naïve estimate $G_{\text{naive}}(\boldsymbol{k}, J, J^*)$. For future applications, it will be convenient to express the result in terms of the scaled central moments μ_2, μ_3 defined by

$$\mu_r = \frac{1}{M}\sum_{i=1}^{n}(k_i - k)^r$$

for $r = 2, 3$.

THEOREM 5.1.

$$\begin{aligned}(5.4)\quad G(\boldsymbol{k}, J, J^*) = \sqrt{2}\, G_{\text{naive}}(\boldsymbol{k}, J, J^*) \exp\biggl(&\frac{1}{4}(1-\mu_2)\bigl(1 + 2x_2 + \mu_2(1-2x_2)\bigr)\\ &+ \frac{6\mu_2\mu_3(x_3 - x_2) - \mu_2^3}{2n} + \frac{3\mu_2^2(\mu_2^2 - 2\mu_3) + 2\mu_3^2(3x_3 - 3x_2 + 1)}{12M}\\ &+ \frac{\mu_2^2 M}{2n^2}(9x_3 - 9x_2 - 1) + O(k_{\max}^3/M)\biggr).\end{aligned}$$

Proof. Suppose that

$$p = \frac{k}{n} + \frac{(1-2x_2)k^2}{n^2} + \frac{ck}{n^2},$$

where $c = c(n) = O(1)$. By direct computation we find that

$$(5.5)\qquad p^{M/2} = \left(\frac{k}{n}\right)^{M/2} \exp\left(\frac{1}{2}k^2(1-2x_2) - \frac{k^3}{4n} + \frac{ck}{2} + O(k_{\max}^3/M)\right)$$

and

(5.6)

$$\begin{aligned}J(1)^{\binom{n}{2}} &= \exp\left(\binom{n}{2}p - \frac{(1-2x_2)p^2n^2}{4} + \frac{(1-3x_2+3x_3)p^3n^2}{6} + O(k_{\max}^3/M)\right)\\ &= \exp\left(\frac{1}{2}M - \frac{1}{4}k(2x_2k + 2 - k) + \frac{ck}{2} - \frac{k^3(2+3x_2-3x_3)}{6n} + O(k_{\max}^3/M)\right).\end{aligned}$$

To calculate the product in (5.2) we consider two cases, depending on the value of x_2. First assume that $x_2 = 0$ (that is, the value 2 is not permitted off the diagonal). Then

$$J(z)^{n-1}J^*(z^2) = (1+pz)^{n-1}H_0(z),$$

where

$$H_0(z) = \left(1 + \frac{\sum_{i\geq 3} x_i p^i z^i}{1+pz}\right)^{n-1}\left(1 + \sum_{i\geq 1} y_i p^i z^{2i}\right).$$

Consequently, for any integer r with $0 \leq r \leq k_{\max}$ we have

$$\begin{aligned}(5.7)\quad [z^r]J(z)^{n-1}J^*(z^2) &= \sum_{j=0}^{r}\binom{n-1}{r-j}p^{r-j}\,[z^j]H_0(z)\\ &= \binom{n-1}{r}p^r + \binom{n-1}{r-2}y_1p^{r-1} + \binom{n-1}{r-3}x_3(n-1)p^r\\ &\quad + \binom{n-1}{r}p^r\Delta_0,\end{aligned}$$

where

$$\varDelta_0 = \sum_{j=4}^{r} \binom{n-1}{r-j}\binom{n-1}{r}^{-1} p^{-j}\,[z^j]H_0(z).$$

Now define

$$H_0^+(z) = \biggl(1 + \frac{\sum_{i\ge 3} p^i z^i}{1-pz}\biggr)^{n-1} \biggl(1 + \sum_{i\ge 1} p^i z^{2i}\biggr) = \biggl(1 + \frac{p^3z^3}{(1-pz)^2}\biggr)^{n-1} (1-pz^2)^{-1}.$$

Notice that the coefficients of $H_0^+(z)$, when expanded as a Taylor series in z, are nonnegative and dominate those of $H_0(z)$. Also notice that $\binom{n-1}{r-j}\binom{n-1}{r}^{-1} = O(1)(r/n)^j$ uniformly for $4 \le j \le r$, since $r = o(n^{1/2})$. Therefore, setting $\alpha = r/pn$,

$$\begin{aligned} |\varDelta_0| &= O(1) \sum_{j=4}^{r} \alpha^j\,[z^j]H_0^+(z) \\ &= O\bigl(H_0^+(\alpha) - 1 - p\alpha^2 - (n-1)p^3\alpha^3\bigr). \end{aligned}$$

Since $p\alpha = o(1)$, by Taylor's theorem we can write

$$\biggl(1 + \frac{p^3\alpha^3}{(1-p\alpha)^2}\biggr)^{n-1} = 1 + (n-1)\frac{p^3\alpha^3}{(1-p\alpha)^2} + O(n^2p^6\alpha^6).$$

Substituting this into the definition of $H_0^+(\alpha)$, we find that

$$H_0^+(\alpha) - 1 - p\alpha^2 - (n-1)p^3\alpha^3 = \frac{p^2\alpha^4(1+o(1))}{(1-p\alpha)^2(1-p\alpha^2)} = O(p^2\alpha^4),$$

and so

$$\varDelta_0 = O(rk_{\max}^3/M^2).$$

Applying this bound to (5.7), we have

$$\begin{aligned} &[z^r]J(z)^{n-1}J^*(z^2) \\ &\quad= \binom{n-1}{r} p^r \exp\biggl(\frac{y_1 r(r-1)}{kn} + \frac{x_3 r^3}{n^2} + O(rk_{\max}^3/M^2)\biggr) \\ &\quad= \frac{n^r p^r}{r!} \exp\biggl(-\frac{r(r+1)}{2n} + \frac{y_1 r(r-1)}{kn} + \frac{r^3(6x_3-1)}{6n^2} + O(rk_{\max}^3/M^2)\biggr). \end{aligned} \tag{5.8}$$

Now we assume that $x_2 = 1$ and perform a similar calculation. We have

$$J(z)^{n-1}\,J^*(z^2) = (1-pz)^{-n+1} H_1(z),$$

where

$$H_1(z) = \biggl((1-pz)\sum_{i\ge 0} x_i p^i z^i\biggr)^{n-1} \biggl(1 + \sum_{i\ge 1} y_i p^i z^{2i}\biggr).$$

We find that for any r with $0 \le r \le k_{\max}$,

$$[z^r]\, J(z)^{n-1} J^*(z^2) = \binom{n+r-2}{r} p^r + \binom{n+r-4}{r-2} y_1 p^{r-1} + \binom{n+r-5}{r-3} (x_3-1)(n-1)\, p^r + \binom{n+r-2}{r}\, p^r \varDelta_1,$$

where

$$\varDelta_1 = \sum_{j=4}^{r} \binom{n+r-j-2}{r-j} \binom{n+r-2}{r}^{-1} p^{-j} [z^r] H_1(z).$$

The coefficients of $H_1(z)$ are dominated by those of $H_1^+(z)$, where

$$H_1^+(z) = \Bigl(1 + \sum_{i\ge 3} p^i z^i\Bigr)^{n-1} \Bigl(1 + \sum_{i\ge 1} p^i z^{2i}\Bigr) = \Bigl(1 + \frac{p^3 z^3}{1-pz}\Bigr)^{n-1} (1-pz^2)^{-1}.$$

Arguing as before, this implies that $\varDelta_1 = O(rk_{\max}^3/M^2)$, and so we have

(5.9)

$$\begin{aligned}[z^r]\, J(z)^{n-1} J^*(z^2) &= \binom{n+r-2}{r} p^r \exp\Bigl(\frac{y_1 r(r-1)}{kn} + \frac{(x_3-1)r^3}{n^2} + O(rk_{\max}^3/M^2)\Bigr)\\ &= \frac{n^r p^r}{r!} \exp\Bigl(\frac{r(r-3)}{2n} + \frac{y_1 r(r-1)}{kn} + \frac{r^3(6x_3-7)}{6n^2} + O(rk_{\max}^3/M^2)\Bigr).\end{aligned}$$

Combining the cases (5.8) and (5.9) we have, for $i = 1, \ldots, n$,

$$[z^{k_i}]\, J(z)^{n-1} J^*(z^2) = \frac{n^{k_i}\, p^{k_i}}{k_i!} \exp\Bigl(-\Bigl(\frac{1+2x_2}{2n} + \frac{y_1}{kn}\Bigr) k_i - \Bigl(\frac{1-2x_2}{2n} - \frac{y_1}{kn}\Bigr) k_i^2 - \frac{1+6x_2-6x_3}{6n^2}\, k_i^3 + O(k_i k_{\max}^3/M^2)\Bigr).$$

Multiplying these n equations together gives

$$\prod_{i=1}^{n} [z^{k_i}] J(z)^{n-1} J^*(z^2) = \frac{n^M p^M}{\prod_{i=1}^n k_i!}\, \exp\Bigl(\frac{M_2(2x_2-1) - 2M}{2n} + \frac{y_1 M_2}{kn} - \frac{(6x_2 - 6x_3 + 1)M_3}{6n^2} + O(k_{\max}^3/M)\Bigr). \tag{5.10}$$

We now substitute (5.5), (5.6), and (5.10) into (5.2). Writing the result in terms of μ_2, μ_3, M, and n, using the identities

$$\begin{aligned} M_2 &= M\mu_2 + (k-1)M,\\ M_3 &= M\mu_3 + 3(k-1)M\,\mu_2 + (k-1)(k-2)M \end{aligned}$$

we find that

$$G_{\text{naive}}(\boldsymbol{k}, J, J^*) = \frac{1}{\prod_{i=1}^{n} k_i!}\left(\frac{M}{e}\right)^{M/2} \exp\biggl(-y_1(1-\mu_2) + \frac{(2y_1 + 2x_2(\mu_2 - 1) - \mu_2)k}{2} + \frac{(2x_2-1)k^2}{4} - \frac{k(6x_2 - 6x_3 + 1)(2\mu_3 + 6\mu_2 k + k^2)}{12n} + O(k_{\max}^3/M)\biggr).$$

(At this point we can observe that no remaining terms depend on c, which verifies our earlier claim that the exact value of the $O(k/n^2)$ term in (5.3) does not affect the statement of this theorem.) The proof is completed by applying Theorem 1.1 and using Stirling's formula. ∎

COROLLARY 5.2. *Under the conditions of Theorem* 5.1, *if* $\mu_2 = O(M^{1/6})$, *then*

$$G(\boldsymbol{k}, J, J^*) = \sqrt{2}\, G_{\text{naive}}(\boldsymbol{k}, J, J^*) \exp\biggl(\frac{1}{4}(1-\mu_2)\bigl(1 + 2x_2 + \mu_2(1-2x_2)\bigr) + O\bigl(k_{\max}^2/M^{2/3}\bigr)\biggr).$$

If the stronger bound $\mu_2 = O(k_{\max}^{1/2})$ *holds, then*

$$G(\boldsymbol{k}, J, J^*) = \sqrt{2}\, G_{\text{naive}}(\boldsymbol{k}, J, J^*) \exp\biggl(\frac{1}{4}(1-\mu_2)\bigl(1 + 2x_2 + \mu_2(1-2x_2)\bigr) + O\bigl(k_{\max}^3/M\bigr)\biggr).$$

Finally, if $\boldsymbol{k} = (k, k, \ldots, k)$ *is a regular degree sequence with* kn *even, then*

$$G(\boldsymbol{k}, J, J^*) = \sqrt{2}\, \exp\biggl(\frac{1}{4}(1 + 2x_2) + O(k^2/n)\biggr) G_{\text{naive}}(\boldsymbol{k}, J, J^*). \tag{5.11}$$

Proof. For the first two statements, we just need to check that all additional terms inside the exponential factor in Theorem 5.1 are covered by the claimed error bounds. For this it is useful to note that $\mu_2 \le k_{\max}$ and $|\mu_3| \le k_{\max}\mu_2$. Finally, when $\boldsymbol{k}$ is a regular degree sequence we have $\mu_2 = \mu_3 = 0$ and (5.11) follows. ∎

The constant $\frac{1}{4}(1 + 2x_2)$ in (5.11) was previously noted for simple sparse regular graphs in [17] and in [5] for simple sparse regular graphs with loops allowed. Interestingly, with different error terms, the same constant was observed in [17] for dense simple regular graphs, in [5] for dense simple regular graphs with loops allowed, and in [13] for dense symmetric integer matrices with zero diagonal. In these three cases, we conjectured that (5.11) holds (with some vanishing error term) for all degree sequences except for the two extreme cases of graphs with no edges and graphs with all possible edges. The corresponding conjectures may be less likely to be true in full generality for all J, J^*.

From the first expression of Corollary 5.2, we see that $G(\boldsymbol{k}, J, J^*)$ is closely approximated by $\sqrt{2}\, G_{\text{naive}}(\boldsymbol{k}, J, J^*)$ whenever μ_2 is close to 1. It appears that in the random matrix model described at the start of this section, μ_2 will be concentrated around 1 with high probability whenever p tends to zero slowly enough to ensure that $M \to \infty$ with high probability, but quickly enough to ensure that $k_{\max}^3 = o(M)$ with high probability. If so, this will lead to a model for the degree sequences of sparse multigraphs analogous to that obtained by McKay and Wormald [18] for graphs and by McKay and Skerman [14] for bipartite graphs and directed graphs. Details will be given in a future paper.

The authors are grateful to the referees for their careful reading, which led to several improvements.

REFERENCES

[1] J. R. Banavar, F. Colaiori, A. Flammini, A. Maritan, and A. Rinaldo, *Topology of the fittest transportation network*, Phys. Rev. Lett., 84 (2000), pp. 4745–4748.

[2] A. Barvinok and J. A. Hartigan, *An asymptotic formula for the number of non-negative integer matrices with prescribed row and column sums*, Trans. Amer. Math. Soc., 364 (2012), pp. 4323–4368.

[3] E. A. Bender and E. R. Canfield, *The asymptotic number of labeled graphs with given degree sequences*, J. Combin. Theory Ser. A, 24 (1978), pp. 296–307.

[4] C. Greenhill and B. D. McKay, *Asymptotic enumeration of sparse nonnegative integer matrices with specified row and column sums*, Adv. in Appl. Math., 41 (2008), pp. 459–481.

[5] C. Greenhill and B. D. McKay, *Counting loopy graphs with given degrees*, Linear Algebra Appl., 436 (2012), pp. 901–926.

[6] H. H. Gan, D. Fera, J. Zorn, N. Shiffeldrim, M. Tang, U. Laserson, N. Kim, and T. Schlick, *RAG: RNA-As-Graphs database—concepts, analysis, and features*, Bioinformatics, 20 (2004), pp. 1285–1291.

[7] E. A. J. Godehardt, *Probability models for random multigraphs with applications in cluster analysis*, Ann. Discrete Math., 55 (1993), pp. 93–108.

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

[9] M. Hasheminezhad and B. D. McKay, *Combinatorial estimates by the switching method*, Contemp. Math., 531 (2010), pp. 209–221.

[10] D. Knisley, J. Knisley, C. Ross, and A. Rockney, *Classifying Multigraph models of secondary RNA structure using graph-theoretic descriptors*, ISRN Bioinformatics, (2012), 157135.

[11] B. D. McKay, *Asymptotics for* 0-1 *matrices with prescribed line sums*, in Enumeration and Design, S. A. Vanstone and D. M. Jackson, eds., Academic Press, Toronto, ON, Canada, 1984, pp. 225–238.

[12] B. D. McKay, *Subgraphs of dense random graphs with specified degrees*, Combin. Probab. Comput., 20 (2011), pp. 413–433.

[13] B. D. McKay and J. C. McLeod, *Asymptotic enumeration of symmetric integer matrices with uniform row sums*, J. Aust. Math. Soc., 92 (2012), pp. 367–384.

[14] B. D. McKay and F. Skerman, *Degree sequences of random digraphs and bipartite graphs*, arXiv:1302.2446, 2012.

[15] B. D. McKay and N. C. Wormald, *Asymptotic enumeration by degree sequence of graphs of high degree*, European J. Combin., 11 (1990), pp. 565–580.

[16] B. D. McKay and N. C. Wormald, *Uniform generation of random regular graphs of moderate degree*, J. Algorithms, 11 (1990), pp. 52–67.

[17] B. D. McKay and N. C. Wormald, *Asymptotic enumeration by degree sequence of graphs with degrees* $o(n^{1/2})$, Combinatorica, 11 (1991), pp. 369–382.

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

[19] R. C. Read, *Some Enumeration Problems in Graph Theory*, Ph.D. thesis, University of London, 1958.