ASYMPTOTIC ENUMERATION BY DEGREE SEQUENCE OF GRAPHS WITH DEGREES $o(n^{1/2})$

Brendan D. McKay

Computer Science Department, Australian National University,

GPO Box 4, ACT 2601, Australia

Nicholas C. Wormald

Department of Mathematics, University of Melbourne,

Parkville, Vic 3052, Australia

Abstract.

We determine the asymptotic number of labelled graphs with a given degree sequence for the case where the maximum degree is $o(|E(G)|^{1/3})$. The previously best enumeration, by the first author, required maximum degree $o(|E(G)|^{1/4})$. In particular, if $k = o(n^{1/2})$, the number of regular graphs of degree k and order n is asymptotically

$$\frac{(nk)!}{(nk/2)!\,2^{nk/2}(k!)^n}\exp\Bigl(-\frac{k^2-1}{4}-\frac{k^3}{12n}+O(k^2/n)\Bigr).$$

Under slightly stronger conditions, we also determine the asymptotic number of unlabelled graphs with a given degree sequence. The method used is a switching argument recently used by us to uniformly generate random graphs with given degree sequences.

1. Introduction.

Where it is suitable, we will use the notation of [1] or [4]. For any integer $y \geq 0$, define $[x]_y = x(x-1)\cdots(x-y+1)$. Let $\boldsymbol{k} = \boldsymbol{k}(n) = (k_1, k_2, \ldots, k_n)$ be a sequence of nonnegative integers with even sum. Define $k_{\max} = \max_{i=1}^n k_i$ and $\bar{k} = (k_1 + k_2 + \cdots + k_n)/n$. For $r \geq 0$, further define $M_r = M_r(\boldsymbol{k}) = \sum_i [k_i]_r$ and $\nu_r = \nu_r(\boldsymbol{k}) = \sum_i k_i^r/(\bar{k}^r n)$. It is easy to see that $1 = \nu_0 = \nu_1 \leq \nu_2 \leq \nu_3 \leq \cdots$, with the inequalities being equalities if and only if $k_1 = k_2 = \cdots = k_n$. For simplicity, write $M = M_1$.

Let $\mathcal{G}(\boldsymbol{k})$ be the set of all labelled simple graphs with degree sequence $\boldsymbol{k}$, and define $G(\boldsymbol{k}) = |\mathcal{G}(\boldsymbol{k})|$. We are concerned with the asymptotic value of $G(\boldsymbol{k})$ as $n \to \infty$. Many authors have obtained results by restricting the growth of the maximum degree. Work prior to [1] can be found summarised there. More recently, a completely different approach [3] has born fruit for high degrees. Interestingly, the result in both these extreme cases can be cast in a common form.

Theorem 1.1. *Define* $\lambda = \bar{k}/(n-1)$ *and* $\gamma_2 = n\bar{k}^2(\nu_2 - 1)/(n-1)^2$. *Suppose that either of the following is true:*

(i) $1 \leq k_{\max} = o(M^{1/4})$, $M \to \infty$.

(ii) $|k_i - \bar{k}| = O(n^{1/2+\epsilon})$ *and* $\min\{\bar{k}, n-\bar{k}-1\} > cn/\log n$ *for sufficiently small* $\epsilon > 0$ *and any* $c > \frac{2}{3}$.

Then

$$G(\boldsymbol{k}) \sim \sqrt{2}\big(\lambda^{\lambda}(1-\lambda)^{1-\lambda}\big)^{\binom{n}{2}} \prod_{i=1}^{n} \binom{n-1}{k_i} \exp\Big(\frac{1}{4} - \frac{\gamma_2^2}{4\lambda^2(1-\lambda)^2}\Big). \quad \blacksquare$$

In this paper we will determine the asymptotic value of $G(\boldsymbol{k})$ when $k_{\max} = o(M^{1/3})$. The result will match Theorem 1.1 if some extra restrictions are imposed on the amount of variation amongst the degrees. The method used closely resembles that of [1], the major improvement being the use of switching operations which lend themselves to easier analysis. In Section 6, we will consider the case of unlabelled graphs.

2. The model.

Consider a set of M *points* arranged in *cells* $v_1, v_2, \ldots, v_n$ of size $k_1, k_2, \ldots, k_n$, respectively. Take a partition P (called a *pairing*) of the M points into $M/2$ parts (called *pairs*) of size 2. The *degree* of cell i is k_i.

The *multigraph* $G(P)$ *associated with* P has vertices $v_1, v_2, \ldots, v_n$. The edges of $G(P)$ are in correspondence with the pairs of P; the pair (x, y) corresponds to an edge (v_i, v_j) if $x \in v_i$ and $y \in v_j$. A *loop* of P is a pair whose two points lie in the same vertex, while a *link* is one involving two distinct vertices. Two pairs are *parallel* if they involve the same cells. The *multiplicity* of a pair is the number of pairs (including itself) parallel to it. A *single* pair is a pair of multiplicity one. A *double* pair is a set of two parallel pairs of multiplicity two, whilst a *triple* pair is a set of three parallel pairs of multiplicity three.

If p is a point, then $v(p)$ is the cell containing that point.

For $l, d, t \geq 0$, define $\mathcal{C}_{l,d,t} = \mathcal{C}_{l,d,t}(\boldsymbol{k})$ be the set of all pairings with degrees $\boldsymbol{k}$, and exactly l loops, d double pairs, and t triple pairs, but no loops of multiplicity greater than one nor pairs of multiplictity greater than three.

We will make use of the following three operations on a pairing: the first two were introduced in [4].

I **ℓ-switching:** Take a loop $\{p_1, p_1'\}$ and two links $\{p_2, p_2'\}$ and $\{p_3, p_3'\}$, such that five distinct cells are involved. Replace these three pairs by $\{p_1, p_2\}$, $\{p_1', p_3\}$ and $\{p_2', p_3'\}$. It is required that all of the pairs created or destroyed be single. (See Figure 1.)

Figure 1. An ℓ-switching.

II d-switching: Take a double link $\{\{p_1, p_1'\}, \{p_2, p_2'\}\}$, where $v(p_1) = v(p_2)$, and two links $\{p_3, p_3'\}$ and $\{p_4, p_4'\}$, such that six distinct cells are involved. Replace these four pairs by $\{p_1, p_3\}$, $\{p_2, p_4\}$, $\{p_1', p_3'\}$ and $\{p_2', p_4'\}$. Other than the original double link, all of the pairs created or destroyed must be single. (See Figure 2.)

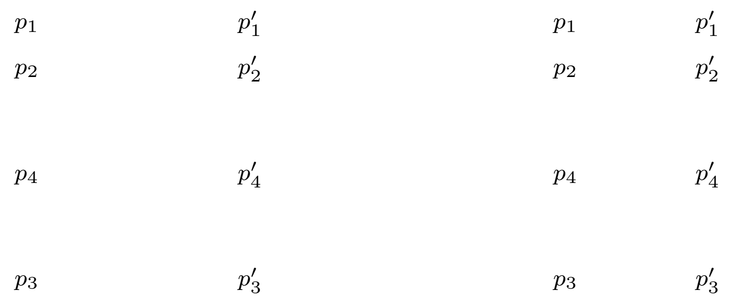

Figure 2. A d-switching.

III t-switching: Take a triple link $\{\{p_1, p_1'\}, \{p_2, p_2'\}, \{p_3, p_3'\}\}$, where $v(p_1) = v(p_2) = v(p_3)$, and three other links $\{p_4, p_4'\}$, $\{p_5, p_5'\}$ and $\{p_6, p_6'\}$, such that eight distinct cells are involved. Replace these six pairs by $\{p_1, p_4\}$, $\{p_2, p_5\}$, $\{p_3, p_6\}$, $\{p_1', p_4'\}$, $\{p_2', p_5'\}$ and $\{p_3', p_6'\}$. Other than the original triple link, all of the pairs created or destroyed must be single. (See Figure 3.)

The inverse of an ℓ-switching will be termed an *inverse ℓ-switching*, and similarly with the other switching types. Note that an ℓ-switching reduces by one the number of loops, without affecting the number of double or triple pairs. We will use this fact to estimate the relative cardinalities of $\mathcal{C}_{l,d,t}$ and $\mathcal{C}_{l-1,d,t}$. The numbers of double and triple links are similarly affected by d-switchings and t-switchings, respectively.

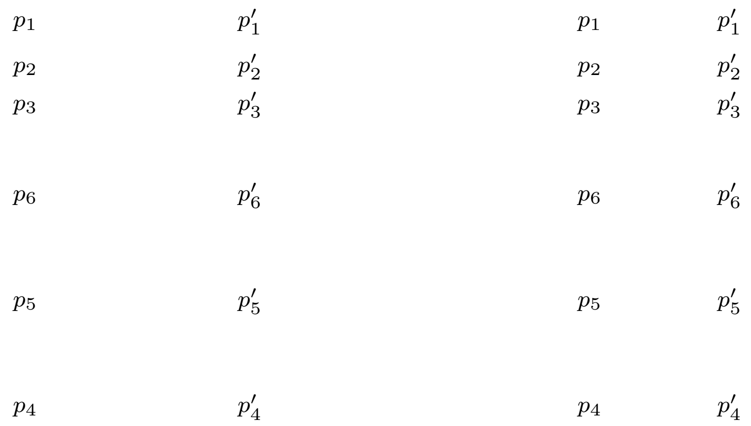

Figure 3. A t-switching.

3. Preliminary results.

Let P be a random pairing with degrees $k_1, k_2, \ldots, k_n$, where $1 \le k_{\max} = o(M^{1/3})$. We will begin with some elementary bounds on the probability that P has certain substructures. The first follows from a simple count.

Lemma 3.1. *The probability of t given pairs occurring in P is*

$$\frac{[M/2]_t 2^t}{[M]_{2t}} \le (M-2t)^{-t}. \quad \blacksquare$$

Define $P(\boldsymbol{k})$ to be the probability that P contains no loops, and no links of multiplicity greater than one. Since each labelled simple graph with degree sequence $k_1, k_2, \ldots, k_n$ corresponds to exactly $k_1!\, k_2! \cdots k_n!$ pairings, we have

$$G(\boldsymbol{k}) = \frac{M!}{(M/2)!\, 2^{M/2} k_1! \cdots k_n!} P(\boldsymbol{k}). \tag{1}$$

Our task is thus reduced to computing $P(\boldsymbol{k})$. We first show that we can ignore pairs of high multiplicity, and bound the number of loops and other non-single links. Define

$$N_1 = \max\bigl(\lceil \log M \rceil, \lceil 4M_2/M \rceil\bigr)$$
$$N_2 = \max\bigl(\lceil \log M \rceil, \lceil 2M_2^2/M^2 \rceil\bigr)$$

and

$$N_3 = \max\bigl(\lceil \log M \rceil, \lceil M_3^2/M^3 \rceil\bigr).$$

In the following lemma, and for the remainder of the paper, the notations "$O(\,)$" and "$o(\,)$" refer to the passage of M to infinity within the constraint that $k_{\max}^3 = o(M)$. The implied constants will be uniform over all free variables unless otherwise stated.

Lemma 3.2.

$$\frac{1}{P(\boldsymbol{k})} = \big(1 + O(k_{\max}^3/M)\big) \sum_{l=0}^{N_1} \sum_{d=0}^{N_2} \sum_{t=0}^{N_3} \frac{|\mathcal{C}_{l,d,t}|}{|\mathcal{C}_{0,0,0}|}.$$

Proof. By considering all the possibilities and applying Lemma 3.1, we find that the probability that P contains a loop of multiplicity greater than one is $O(k_{\max}^3/M)$. Similarly, the probability of a link of multiplicity greater than three is $O(k_{\max}^6/M^2) = o(k_{\max}^3/M)$.

Consider the probability that there are more than N_1 single loops. By Lemma 3.1, the expected number of sets of $l = N_1 + 1$ single loops is $O\big((M_2/(2M))^l/l!\big) = o\big((e/8)^{\log M}\big) = o(1/M)$. Similarly, the probability that there are more than N_2 double links or more than N_3 triple links is $o(1/M)$. The lemma follows. ∎

We will estimate $|\mathcal{C}_{l,d,t}|/|\mathcal{C}_{0,0,0}|$ via estimates on the terms of the expansion

$$\frac{|\mathcal{C}_{l,d,t}|}{|\mathcal{C}_{0,0,0}|} = \frac{|\mathcal{C}_{l,d,t}|}{|\mathcal{C}_{l,d,t-1}|} \cdots \frac{|\mathcal{C}_{l,d,1}|}{|\mathcal{C}_{l,d,0}|} \frac{|\mathcal{C}_{l,d,0}|}{|\mathcal{C}_{l-1,d,0}|} \cdots \frac{|\mathcal{C}_{1,d,0}|}{|\mathcal{C}_{0,d,0}|} \frac{|\mathcal{C}_{0,d,0}|}{|\mathcal{C}_{0,d-1,0}|} \cdots \frac{|\mathcal{C}_{0,1,0}|}{|\mathcal{C}_{0,0,0}|}.$$

Each of these terms can be estimated by means of one of the three switchings.

In the case of d-switchings, the analysis will be considerably more exacting, and we will need the following result adapted from [1].

If K is a multigraph, let $e(K)$ denote its number of edges (including loops, counting multiplicities), and let $e_1(K)$ denote its number of loops. If xx' is an edge of K, then $\mu_K(xx')$ denotes its multiplicity, i.e., the number of edges parallel to xx' including itself. If K and K' are multigraphs with the same vertex set, then $K + K'$ is the multigraph with the same vertex set such that $\mu_{K+K'}(xx') = \mu_K(xx') + \mu_{K'}(xx')$ for all $\{x, x'\}$. Similarly, $2K$ means $K + K$ and $K + xx'$ is K with the multiplicity of xx' increased by one.

Let L be a graph on n vertices which is simple apart from a loop on each vertex, and let H be a multigraph on the same vertex set with the restriction that if any edge xx' has $\mu_H(xx') \geq 1$, then xx' is an edge of L. Let $l_{\max}$ denote the maximum degree of L. Define $\mathcal{C}(L,H) = \mathcal{C}(L,H;\boldsymbol{k})$ to be the set of all pairings P with degrees $\boldsymbol{k}$ such that the following are true for all $\{x, x'\}$:

(a) If xx' is an edge of L, then $\mu_{G(P)}(xx') = \mu_H(xx')$.

(b) If xx' is not an edge of L, then $\mu_{G(P)}(xx') \leq 1$.

In other words, $G(P)$ must be simple outside L and match H inside L.

Lemma 3.3. *Suppose that L is as defined above, and H and $H + J$ satisfy the requirements given above for H. Let $h_1, h_2, \ldots, h_n$ be the degrees of H, and let $j_1, j_2, \ldots, j_n$ be the degrees of J. Then, if $k_{\max}(k_{\max} + l_{\max})e(J) = o(M)$, $e(H) = o(M)$, and $\mathcal{C}(L,H) \neq \emptyset$, we have*

$$\frac{|\mathcal{C}(L,H+J)|}{|\mathcal{C}(L,H)|} = \frac{\prod_{i=1}^n [k_i - h_i]_{j_i}}{2^{e_1(J)+e(J)} [M/2 - e(H)]_{e(J)} \prod_{\{x,x'\}} [\mu_{H+J}(xx')]_{\mu_J(xx')}} \Big(1 + O\Big(\frac{k_{\max}(k_{\max} + l_{\max})e(J)}{M}\Big)\Big).$$

Proof. This is a special case of the combination of Theorems 3.3 and 3.6 of [1]. ∎

We will use Lemma 3.3 to analyse the structure of $\mathcal{C}_{0,d,0}$. For a pairing $P \in \mathcal{C}_{0,d,0}$, let $D(P)$ be the simple graph with vertices $v_1, v_2, \ldots, v_n$ and just those edges which correspond in position to the d double links of P.

Lemma 3.4. *Let $D = D(P)$ for some $P \in \mathcal{C}_{0,d,0}$, where $0 \le d \le N_2$. Let S be a simple graph on $v_1, v_2, \ldots, v_n$ which is edge-disjoint from D. Let $d_1, d_2, \ldots, d_n$ be the degrees of D, $s_1, s_2, \ldots, s_n$ be the degrees of S, and suppose that $e(S)k_{\max}^2 = o(M)$. Then the probability that $S \subseteq G(P)$ when P is chosen at random from those $P \in \mathcal{C}_{0,d,0}$ such that $D(P) = D$ is*

$$\frac{\prod_{i=1}^n [k_i - 2d_i]_{s_i}}{2^{e(S)}[M/2]_{e(s)}} \exp\Big(O\Big(\frac{e(S)(k_{\max}^2 + d)}{M}\Big)\Big).$$

Proof. The lemma is clearly true if $s_i > k_i - 2d_i$ for any i, so suppose that that is not the case. Define the graph L which has the edges of D and S as well as a loop on each vertex. Then, for any $J \subseteq S$, Lemma 3.3 tells us that

$$\frac{|\mathcal{C}(L, 2D + J)|}{|\mathcal{C}(L, 2D)|} = f(J),$$

where

$$f(J) = \frac{\prod_{i=1}^n [k_i - 2d_i]_{j_i}}{2^{e(J)}[M/2 - 2d]_{e(J)}} \exp\Big(O\Big(\frac{k_{\max}^2 e(J)}{M}\Big)\Big),$$

and $j_1, j_2, \ldots, j_n$ are the degrees of J. Now, the required probability can be written as

$$\frac{f(S)}{\sum_{J \subseteq S} f(J),}$$

and since the denominator is $1 + O\big(e(S)k_{\max}^2/M\big)$, the lemma follows. ∎

In the following, we will find it convenient to write k_v in place of k_i if $v = v_i$.

Lemma 3.5. *Suppose that $0 \le d \le N_2$ and $M_2 \ge M$. Choose $v \in \{v_1, v_2, \ldots, v_n\}$ and $r \ge 0$. Then, if P is chosen at random from $\mathcal{C}_{0,d,0}$, cell v is incident with exactly r double links with probability $Q_v(r)/\sum_{i \ge 0}^{\lfloor k_v/2 \rfloor} Q_v(i)$, where*

$$Q_v(i) = \frac{2^i [d]_i [k_v]_{2i}}{i!\, M_2^i} \exp\Big(O\Big(\frac{ik_{\max}^2}{M} + \frac{i^2 k_{\max}^2 + idk_{\max}}{M_2}\Big)\Big).$$

Proof. Suppose that $D = D(P)$ for some $P \in \mathcal{C}_{0,d,0}$, and let w be a neighbour of v in D. Let x and x' be distinct vertices other than v, such that xx' is not an edge of D. Let L be the graph with the edges of D, plus xx', plus a loop on each vertex. Let $R = D - vw$, $0 \le \alpha \le 2$ and $0 \le \beta \le 2$, and, for vertex u, let r_u denote the degree of u in R. Then

$$\frac{|\mathcal{C}(L, 2R + \alpha vw + \beta xx')|}{|\mathcal{C}(L, 2R)|} = \frac{f_R(\alpha, \beta; v, w, x, x')}{2^{\alpha+\beta}\alpha!\, \beta!\, [M/2 - 2d + 2]_{\alpha+\beta}} \Big(1 + O\Big(\frac{k_{\max}^2}{M}\Big)\Big),$$

where, by Lemma 3.3,

$$f_R(\alpha,\beta;v,w,x,x') = \begin{cases} [k_v-2r_v]_\alpha[k_w-2r_w]_\alpha[k_x-2r_x]_\beta[k_{x'}-2r_{x'}]_\beta, & \text{if } w\notin\{x,x'\},\\ [k_v-2r_v]_\alpha[k_w-2r_w]_{\alpha+\beta}[k_{x'}-2r_{x'}]_\beta, & \text{if } w=x, \text{ and}\\ [k_v-2r_v]_\alpha[k_w-2r_w]_{\alpha+\beta}[k_x-2r_x]_\beta, & \text{if } w=x'. \end{cases}$$

For any simple graph X, let $N[X]$ denote the number of pairings $P\in\mathcal{C}_{0,d,0}$ such that $D(P)=X$. Then $N[R+vw]=|\mathcal{C}(L,2R+2vw)\cup\mathcal{C}(L,2R+2vw+xx')|$, and similarly for $N[R+xx']$. Thus, when $N[R+vw]\neq 0$,

$$\begin{aligned}\frac{N[R+xx']}{N[R+vw]} &= \frac{f_R(0,2;v,w,x,x')}{f_R(2,0;v,w,x,x')}\Big(1+O\Big(\frac{k_{\max}^2}{M}\Big)\Big)\\ &= \frac{[k_x-2r_x]_2[k_{x'}-2r_{x'}]_2}{[k_v-2r_v]_2[k_w-2r_w]_2}\Big(1+O\Big(\frac{k_{\max}^2}{M}\Big)\Big), \qquad (2)\end{aligned}$$

since the terms involving $f_R(1,2;v,w,x,x')$ and $f_R(2,1;v,w,x,x')$ are small enough to be incorporated into the error term.

Suppose $1\le i\le\lfloor k_v/2\rfloor$. Define $\mathcal{R}(i)$ to be the set of all simple graphs on $V=\{v_1,v_2,\ldots,v_n\}$ with exactly $d-1$ edges, of which exactly $i-1$ are incident with v. For $R\in\mathcal{R}(i)$, let $\mathcal{X}(R)$ denote the set of all distinct pairs $\{x,x'\}$ such that $x\neq v$, $x'\neq v$ and xx' is not an edge of R. Similarly, let $\mathcal{W}(R)$ denote the set of all $w\in V$ such that $w\neq v$ and vw is not an edge of R.

If n_i denotes the number of pairings $P\in\mathcal{C}_{0,d,0}$ such that exactly i double links are incident with v, then

$$n_{i-1}=\frac{1}{d-i+1}\sum_{R\in\mathcal{R}(i)}\sum_{xx'\in\mathcal{X}(R)}N[R+xx']$$

and

$$n_i=\frac{1}{i}\sum_{R\in\mathcal{R}(i)}\sum_{w\in\mathcal{W}(R)}N[R+vw].$$

From (2) we find that, for any w and $R\in\mathcal{R}(i)$ for which the denominator is non-zero,

$$\frac{\sum_{xx'\in\mathcal{X}(R)}N[R+xx']}{N[R+vw]}=\frac{M_2^2}{2[k_v-2r_v]_2[k_w-2r_w]_2}\Big(1+O\Big(\frac{k_{\max}^2}{M}+\frac{dk_{\max}+k_{\max}^2}{M_2}\Big)\Big).$$

(To see this, express the numerator as a sum over all ordered pairs $xx'\in V\times V$ and subtract those not in $\mathcal{X}(R)$.) We can sum over w in a similar way to obtain, for any $R\in\mathcal{R}(i)$ for which the denominator is non-zero,

$$\frac{\sum_{xx'\in\mathcal{X}(R)}N[R+xx']}{\sum_{w\in\mathcal{W}(R)}N[R+vw]}=\frac{M_2}{2[k_v-2(i-1)]_2}\Big(1+O\Big(\frac{k_{\max}^2}{M}+\frac{ik_{\max}^2+dk_{\max}}{M_2}\Big)\Big).$$

(Note that for M sufficiently large, both the numerator and denominator must be non-zero since $d=o(M^{2/3})$.) Since this is uniform over R, we conclude that

$$\frac{n_i}{n_{i-1}}=\frac{2(d-i+1)[k_v-2(i-1)]_2}{iM_2}\Big(1+O\Big(\frac{k_{\max}^2}{M}+\frac{ik_{\max}^2+dk_{\max}}{M_2}\Big)\Big).$$

Identifying n_r/n_0 as the quantity $Q_v(r)$, we now obtain the lemma on taking the product over i. ∎

4. Analysis of the switchings.

We now analyse each of the switching types in turn, under the assumptions of Section 3.

Lemma 4.1. *Suppose* $0 \le l \le N_1$, $0 \le d \le N_2$ *and* $1 \le t \le N_3$. *Then, if* $M_3 > 0$,

$$\frac{|\mathcal{C}_{l,d,t}|}{|\mathcal{C}_{l,d,t-1}|} = \frac{M_3^2}{12tM^3}\Big(1 + O\Big(\frac{k_{\max}^2(k_{\max}^2 + l + d + t)}{M_3}\Big)\Big).$$

Proof. To simplify the consideration of equivalences, we will consider each of the points involved in the t-switching to be separately labelled (with the labels p_i and p_i' used above in the definition of a t-switching).

Choose an arbitrary $P \in \mathcal{C}_{l,d,t}$, and let $N = N(P)$ be the number of t-switchings which can be applied to P. We can choose a triple link and its labelling in $12t$ ways, and choose three distinct labelled single links $\{p_4, p_4'\}$, $\{p_5, p_5'\}$ and $\{p_6, p_6'\}$ in $[M - 2l - 4d - 6t]_3$ ways. Of these choices, some are not satisfactory. Unwanted coincidences like $v(p_1) = v(p_4)$ account for $O(tk_{\max}M^2)$ choices , while those like $v(p_4) = v(p_5)$ account for $O(tMM_2)$. The forbidden cases where, for example, P already has a pair involving $v(p_1)$ and $v(p_4)$ account for $O(tk_{\max}^2M^2)$. Overall, we find that

$$N = 12tM^3\Big(1 + O\Big(\frac{k_{\max}^2 + l + d + t}{M}\Big)\Big).$$

Now choose an arbitrary $P' \in \mathcal{C}_{l,d,t-1}$, and let $N' = N'(P)$ be the number of *inverse* t-switchings which can be applied to it. We can choose two distinct 3-stars in $[M_3]_2$ ways. Of these choices, we must eliminate those not permitted. Unwanted coincidences, like $v(p_1) = v(p_1')$, $v(p_4) = v(p_5')$ or $v(p_4) = v(p_1')$ account for $O(k_{\max}^3M_3)$ choices. Cases where P' already has a pair involving $v(p_1)$ and $v(p_1')$ or $v(p_4)$ and $v(p_4')$, for example, account for $O(k_{\max}^4M_3)$ choices. Finally, cases where either of the 3-stars include loops or non-single pairs account for $O\big(k_{\max}^2(l + d + t)M_3\big)$ choices. Overall, we find that

$$N' = M_3^2\Big(1 + O\Big(\frac{k_{\max}^2(k_{\max}^2 + l + d + t)}{M_3}\Big)\Big).$$

The error term for N' dominates that for N, so the lemma follows on considering the ratio N'/N. ∎

Lemma 4.2. *Suppose* $1 \le l \le N_1$ *and* $0 \le d \le N_2$. *Then*

$$\frac{|\mathcal{C}_{l,d,0}|}{|\mathcal{C}_{l-1,d,0}|} = \frac{M_2}{2lM}\Big(1 + O\Big(\frac{k_{\max}^2}{M} + \frac{k_{\max}d + k_{\max}^2 l}{M_2}\Big)\Big).$$

Proof. We use the same method as for Lemma 4.1.

Let P be an arbitrary member of $\mathcal{C}_{l,d,0}$, and let $N = N(P)$ be the number of ℓ-switchings which can be applied to it. We can choose and label the loop in $2l$ ways, then choose two single labelled links in $[M-2l-4d]_2$ ways. This overcount needs to be corrected for unwanted coincidences like $v(p_1)=v(p_2)$ ($O(lk_{\max}M)$ choices), and unwanted adjacencies like a pair involving $v(p_1)$ and $v(p_2)$ ($O(lk_{\max}^2M)$). Thus,

$$N = 2lM^2\Big(1+O\Big(\frac{k_{\max}^2+l+d}{M}\Big)\Big).$$

Conversely, let P' be an arbitrary member of $\mathcal{C}_{l-1,d,0}$, and let $N'=N'(P)$ be the number of *inverse* ℓ-switchings which can be applied to it. We can choose the 2-star in M_2 and $\{p_2', p_3'\}$ in M ways. This overcount needs to be corrected for unwanted coincidences like $v(p_2)=v(p_2')$ ($O(k_{\max}M_2)$ choices), unwanted adjacencies like a pair involving $v(p_2)$ and $v(p_2')$ ($O(lk_{\max}^2M+k_{\max}^2M_2)$ choices), and unwanted involvement of loops or non-single pairs ($O((l+d)k_{\max}M)$ choices). Thus

$$N' = MM_2\Big(1+O\Big(\frac{k_{\max}^2}{M}+\frac{k_{\max}d+k_{\max}^2l}{M_2}\Big)\Big).$$

The lemma now follows on comparing N to N'. ∎

Lemma 4.3. *Suppose* $1\le d\le N_2$. *Then, if* $M_2>0$,

$$\frac{|\mathcal{C}_{0,d,0}|}{|\mathcal{C}_{0,d-1,0}|} = \frac{M_2^2}{4dM^2}\Big(1+O\Big(\frac{k_{\max}(k_{\max}^2+d)}{M_2}\Big)\Big).$$

Proof. This can be proved by precisely the same method used for the previous two lemmas. Details can be found in [4]. ∎

Whilst we will use Lemma 4.3 in one special case, it is not sufficiently accurate for us in general. The reason is that the number of double links in a random pairing is in general much higher than the numbers of loops or triple links. However, Lemma 4.3 is the best that can be done using uniform counts over arbitrary members of $\mathcal{C}_{0,d,0}$ and $\mathcal{C}_{0,d-1,0}$. In order to do better, we need to consider averages over $\mathcal{C}_{0,d,0}$ and $\mathcal{C}_{0,d-1,0}$.

Lemma 4.4. *Suppose* $1\le d\le N_2$ *and* $M_2\ge M$. *Then*

$$\frac{|\mathcal{C}_{0,d,0}|}{|\mathcal{C}_{0,d-1,0}|} = \frac{M_2^2}{4dM^2}\Big(1+\frac{4M_3}{M^2}+\frac{8d}{M}-\frac{M_3^2}{MM_2^2}-\frac{2M_2^2}{M^3}-\frac{16dM_3}{M_2^2}+O\Big(\frac{k_{\max}^2+d}{M_2}\Big)\Big).$$

Proof. Define N to be the average number of possible d-switchings, where the average is over all $P\in\mathcal{C}_{0,d,0}$. We can choose $\{p_1,p_1',p_2,p_2'\}$ in $4d$ ways and then $\{p_3,p_3',p_4,p_4'\}$ in at most $[M-4]_2$ ways. This gives us the initial overcount $N\le N^*=4d[M-4]_2=4dM^2\big(1+O(1/M)\big)$. However, some of these choices are not legal. We can divide the set of illegal choices into three families:

X_1: These are choices involving too few vertices, for example if $v(p_1)=v(p_3)$ or $v(p_3)=v(p_4')$.

X_2: These are the choices for which the pairing already has a link involving $v(p_1)$ and $v(p_3)$ or the three other similar cases. However, we exclude any choice which belongs to X_1.

X_3: These are choices for which either $\{p_3, p_3'\}$ or $\{p_4, p_4'\}$ has multiplicity two. However, we exclude any choice which belongs to X_1.

In each of these three cases, we will consider the probability that randomly choosing one of the N^* possibilities described above gives that case, where the probability is taken over random P. We will also bound the probability that X_2 and X_3 occur together.

Case X_1: The probabilty of landing in X_1 is easily seen to be at most $O(k_{\max}/M)$, by just counting the cases.

Case X_2: Let P_i $(i = 1, 2)$ be the probability that there is a pair $\{x, x'\}$ of multiplicity i such that $v(x) = v(p_1)$ and $v(x') = v(p_3)$. Note that our conditions on d, $k_{\max}$, M and M_2 imply that $dk_{\max}^2/M_2 = o(1)$. From Lemma 3.5, the expected number of pairs of adjacent double links is $O(d^2M_4/M_2^2) = O(d^2k_{\max}^2/M_2)$. Allowing $k_{\max}$ for the choices of p_3' and M for the choice of $\{p_4, p_4'\}$, we find that $P_2 = O\big(dk_{\max}^3/(MM_2)\big) = O(k_{\max}/M)$.

P_1 is more involved. For any choice of $D(P)$, p_1, p_1', p_2 and p_2', there are on average

$$M_2(k_v - 2r)\Big(1 + O\Big(\frac{k_{\max}^2}{M} + \frac{dk_{\max}}{M_2}\Big)\Big)$$

choices for p_3, p_3', p_4 and p_4', where $v = v(p_1)$ and r is the number of double links incident with v. This follows from Lemma 3.4 on summing over all the possibilities. If K denotes the expected number of configurations included in the value of P_1, then by Lemma 3.5,

$$\begin{aligned} K &= 2M_2\Big(1 + O\Big(\frac{k_{\max}^2}{M} + \frac{dk_{\max}}{M_2}\Big)\Big) \sum_v \sum_{r\ge 1} r(k_v - 2r) \frac{Q_v(r)}{\sum_{i\ge 0} Q_v(i)} \\ &= 4d\Big(1 + O\Big(\frac{k_{\max}^2}{M} + \frac{dk_{\max}}{M_2}\Big)\Big) \sum_v [k_v]_3 \frac{\sum_{i\ge 0} S_v(i)}{\sum_{i\ge 0} Q_v(i)}, \end{aligned}$$

where

$$S_v(i) = \frac{2^i[d-1]_i[k_v-3]_{2i}}{i!\,M_2^i} \exp\Big(O\Big(\frac{ik_{\max}^2}{M} + \frac{i^2k_{\max}^2 + idk_{\max}}{M_2}\Big)\Big).$$

Since $\sum_{i\ge 0} Q_v(i) \ge 1$,

$$\frac{\sum_{i\ge 0} S_v(i)}{\sum_{i\ge 0} Q_v(i)} = 1 + O\big(\sum_{i\ge 0}(Q_v(i) - S_v(i))\big).$$

Using the inequality $|e^x - 1| < |x|e^{|x|}$, we find that

$$Q_v(i) - S_v(i) = \frac{2^i[d-1]_{i-1}[k_v-3]_{2i-3}}{i!\,M_2^i} O\big(e^{O(z)}(ik_v^3 + dik_v^2 + dk_v^3 z)\big),$$

where $z = ik_{\max}^2/M + (i^2k_{\max}^2 + idk_{\max})/M_2$. Hence the terms of the series $\sum_{i\ge 1}\big(Q_v(i) - S_v(i)\big)$ are bounded in magnitude by those of a geometric series with ratio $o(1)$, and thus by a constant multiple of the bound on the first term, that is by $O\big(k_{\max}(k_{\max} + d)/M_2\big)$. Overall, we find that

$$P_1 = \frac{M_3}{M^2} + O\Big(\frac{k_{\max}}{M} + \frac{dk_{\max}^2}{M^2}\Big).$$

Any two of the eight events counted in X_2 (single or double link in any of four positions) occur together with probability $O(k_{\max}/M)$, so altogether we find that X_2 occurs with probability

$$\frac{4M_3}{M^2} + O\Big(\frac{k_{\max}}{M}\Big).$$

Case X_3: With the help of Lemmas 3.4 and 3.5, a routine calculation gives the probability of this case as $8d/M + O(k_{\max}/M)$.

Events X_2 and X_3 occur together with probability $O(k_{\max}/M)$, by similar reasoning. Thus we have altogether that

$$N = 4dM^2\Big(1 - \frac{4M_3}{M^2} - \frac{8d}{M} + O\Big(\frac{k_{\max}}{M}\Big)\Big).$$

Conversely, define N' to be the average number of possible inverse d-switchings, where the average is over all $P \in \mathcal{C}_{0,d-1,0}$. For each choice of $v = v(p_1)$, there are at most $[k_v]_2$ ways to choose p_1 and p_2. A similar bound holds for $v(p'_1)$, and so we an initial overcount $N' \le M_2^2$. However, some of these choices are not legal and, as before, we divide these into a number of cases:

Y_1: These are choices involving too few vertices, for example $v(p_1) = v(p'_1)$ or $v(p_3) = v(p'_3)$.

Y_2: These are choices where there is already a link involving $v(p_1)$ and $v(p'_1)$, excluding anything in Case Y_1.

Y_3: These are choices where there is already a link involving $v(p_3)$ and $v(p'_3)$, or $v(p_4)$ and $v(p'_4)$. Again, we exclude anything in case Y_1.

Y_4: These are choices for which one or more of the pairs chosen have multiplicity two, except any choice in case Y_1.

These four cases can be analysed using the same method used for X_1–X_3. For cases Y_2 and Y_3, we can simply sum over all the possibilities using Lemma 3.4. For Case Y_4, we need Lemma 3.5. We will merely state the probability in each case, leaving the details to the reader.

Case Y_1: $O\Big(\frac{k_{\max}^2}{M_2}\Big)$.

Case Y_2: $\frac{M_3^2}{MM_2^2} + O\Big(\frac{k_{\max}^2}{M_2}\Big)$.

Case Y_3: $\frac{2M_2^2}{M^3} + O\Big(\frac{k_{\max}^2}{M_2}\Big)$.

Case Y_4: $\frac{16dM_3}{M_2^2} + O\Big(\frac{d}{M_2}\Big)$.

The conjunction of any two of these cases gives no new error terms, so overall we have

$$N' = M_2^2\Big(1 - \frac{M_3^2}{MM_2^2} - \frac{2M_2^2}{M^3} - \frac{16dM_3}{M_2^2} + O\Big(\frac{k_{\max}^2 + d}{M_2}\Big)\Big).$$

The lemma now follows on comparing N to N'. ∎

5. Consolidation.

With the aid of the lemmas in Section 4, we can now apply Lemma 3.2 to estimate $P(\boldsymbol{k})$. As before, we assume that $1 \le k_{\max}^3 = o(M)$.

Lemma 5.1.

$$P(\boldsymbol{k}) = \exp\Bigl(-\frac{M_2}{2M} - \frac{M_2^2}{4M^2} - \frac{M_2^2 M_3}{2M^4} + \frac{M_2^4}{4M^5} + \frac{M_3^2}{6M^3} + O\Bigl(\frac{k_{\max}^3}{M}\Bigr)\Bigr).$$

Proof. Let $0 \le l \le N_1$, $0 \le d \le N_2$ and $0 \le t \le N_3$.

If $M_3 = 0$, then clearly $|\mathcal{C}_{l,d,t}| = 0$ if $t > 0$. Suppose instead that $M_3 > 0$. By Lemma 4.1,

$$\frac{|\mathcal{C}_{l,d,t}|}{|\mathcal{C}_{l,d,0}|} = \frac{M_3^{2t}}{12^t t!\, M^{3t}} \exp\Bigl(O\Bigl(\frac{tk_{\max}^2(k_{\max}^2 + l + d) + t^2 k_{\max}^2}{M_3}\Bigr)\Bigr).$$

Summing over t, we obtain

$$\sum_{t=0}^{N_3} \frac{|\mathcal{C}_{l,d,t}|}{|\mathcal{C}_{l,d,0}|} = \exp\Bigl(\frac{M_3^2}{12M^3} + O\Bigl(\frac{k_{\max}(k_{\max}^2 + l + d)}{M}\Bigr)\Bigr), \tag{3}$$

which just happens to be true also for $M_3 = 0$.

Similarly, from Lemma 4.2, we have

$$\frac{|\mathcal{C}_{l,d,0}|}{|\mathcal{C}_{0,d,0}|} = \frac{M_2^l}{2^l l!\, M^l} \exp\Bigl(O\Bigl(\frac{lk_{\max}^2}{M} + \frac{lk_{\max} d + k_{\max}^2 l^2}{M_2}\Bigr)\Bigr).$$

Combining this with (3) and summing over l, we obtain

$$\sum_{l=0}^{N_1}\sum_{t=0}^{N_3} \frac{|\mathcal{C}_{l,d,t}|}{|\mathcal{C}_{0,d,0}|} = \exp\Bigl(\frac{M_3^2}{12M^3} + \frac{M_2}{2M} + O\Bigl(\frac{k_{\max}(k_{\max}^2 + d)}{M}\Bigr)\Bigr). \tag{4}$$

Now suppose that $M_2 \ge M$. From Lemma 4.4, we have that

$$\frac{|\mathcal{C}_{0,d,0}|}{|\mathcal{C}_{0,0,0}|} = \frac{M_2^{2d}}{4^d d!\, M^{2d}} \exp\Bigl(\frac{4dM_3}{M^2} + \frac{4d^2}{M} - \frac{dM_3^2}{MM_2^2} - \frac{2dM_2^2}{M^3} - \frac{8d^2 M_3}{M_2^2} + O\Bigl(\frac{d(k_{\max}^2 + d)}{M_2}\Bigr)\Bigr).$$

Combining this with (4) and summing over d with the help of the approximation $d^2 \approx dM_2^2/(4M^2)$, we obtain

$$\sum_{d=0}^{N_2}\sum_{l=0}^{N_1}\sum_{t=0}^{N_3} \frac{|\mathcal{C}_{l,d,t}|}{|\mathcal{C}_{0,0,0}|} = \exp\Bigl(\frac{M_2}{2M} + \frac{M_2^2}{4M^2} + \frac{M_2^2 M_3}{2M^4} - \frac{M_2^4}{4M^5} - \frac{M_3^2}{6M^3} + O\Bigl(\frac{k_{\max}^3}{M}\Bigr)\Bigr). \tag{5}$$

In the case where $0 < M_2 < M$, Lemma 4.3 gives the same result to within the same error. In the trivial case $M_2 = 0$ (which implies $M_3 = 0$) Equation (5) again holds.

The desired estimate now follows from Lemma 3.2. ∎

We now have the result we have been seeking.

Theorem 5.2. *If* $1 \le k_{\max} = o(M^{1/3})$, *then*

$$G(\boldsymbol{k}) = \frac{M!}{(M/2)!\,2^{M/2}k_1!\cdots k_n!}\exp\Bigl(-\frac{M_2}{2M} - \frac{M_2^2}{4M^2} - \frac{M_2^2M_3}{2M^4} + \frac{M_2^4}{4M^5} + \frac{M_3^2}{6M^3} + O\Bigl(\frac{k_{\max}^3}{M}\Bigr)\Bigr)$$
$$= \frac{M!}{(M/2)!\,2^{M/2}k_1!\cdots k_n!}\exp\Bigl(-\frac{\bar{k}^2\nu_2^2-1}{4} - \frac{\bar{k}^3(6\nu_2^2\nu_3 - 3\nu_2^4 - 2\nu_3^2)}{12n} + O\Bigl(\frac{k_{\max}^3}{M}\Bigr)\Bigr),$$

uniformly as $M \to \infty$.

Proof. The first expression follows from Lemma 5.1 and Equation (1), and the second from the first using the simple bound $\bar{k}\nu_2 = O(k_{\max})$. ∎

Corollary 5.3. *If* $1 \le k = o(n^{1/2})$, *the number of labelled regular graphs of degree* k *and order* n *is asymptotically*

$$\frac{(nk)!}{(nk/2)!\,2^{nk/2}(k!)^n}\exp\Bigl(-\frac{k^2-1}{4} - \frac{k^3}{12n} + O(k^2/n)\Bigr)$$

as $n \to \infty$. ∎

Corollary 5.4. *Define* $k_{\min} = \min_{i=1}^n k_i$. *Then Theorem 1.1 holds in the additional case (iii)* $1 \le k_{\max} = o(M^{1/3})$ *and* $|k_{\max} - k_{\min}| = o\bigl(\min(n^{1/8}\bar{k}^{5/8}, n^{1/6}\bar{k}^{1/2})\bigr)$.

Proof. For $k_{\max} = o(M^{1/3})$, the estimate of $G(\boldsymbol{k})$ in Theorem 1.1 can be expanded as

$$\frac{M!}{(M/2)!\,2^{M/2}k_1!\cdots k_n!}\exp\Bigl(-\frac{\bar{k}^2\nu_2^2-1}{4} - \frac{\bar{k}^3(5+2\nu_3+6\nu_2^2-12\nu_2)}{12n} + O\Bigl(\frac{k_{\max}^3}{M}\Bigr)\Bigr).$$

Since $\nu_2 = 1 + \sigma_2$ and $\nu_3 = 1 + 3\sigma_2 + \sigma_3$, where $\sigma_r = \frac{1}{n}\sum_{i=1}^n (k_i/\bar{k} - 1)^r$, we find that Theorem 1.1 holds provided

$$\frac{\bar{k}^3(6\sigma_2^3 + 6\sigma_2^2\sigma_3 - 3\sigma_2^4 - 2\sigma_3^2)}{n} = o(1).$$

Since $|\sigma_r| \le ((k_{\max} - k_{\min})/\bar{k})^r$, the claim follows after a routine calculation. ∎

Corollary 5.4 adds additional support to the following conjecture, which first appeared in [3]. In fact, for $k_{\max} = o(M^{1/3})$ we can take $\epsilon = 1/8$.

Conjecture. *There is some absolute constant* $\epsilon > 0$ *such that the conclusion of Theorem 1.1 holds whenever* $0 < \bar{k} < n-1$, $|k_{\max} - k_{\min}| = o\bigl(n^\epsilon \min\{\bar{k}, n-\bar{k}-1\}^{1/2}\bigr)$ *and* $\min\{M, \binom{n}{2} - M\} \to \infty$ *as* $n \to \infty$. ∎

6. Unlabelled graphs.

Under some additional constraints, Theorem 5.2 can be applied to estimate the number of unlabelled graphs with a given degree sequence.

Theorem 6.1. *Let $\boldsymbol{k}$ be a graphical degree sequence with no entries of value 0, n_1 entries of value 1, and n_2 entries of value 2. Then, under any one of the following conditions, the number of unlabelled simple graphs with degree sequence $\boldsymbol{k}$ is asymptotically $G(\boldsymbol{k})/n!$.*

(i) $n_1 = O(n^{1/3})$, $n_2 = O(n^{2/3})$, *and* $k_{\max} \le \frac{1}{3}\log n/\log\log n$;

(ii) $k_{\min} \ge 4$ *and* $k_{\max} = o(\bar{k}^{1/2}n^{1/12})$;

(iii) $k_{\min} \ge 5$ *and* $k_{\max} = o(\bar{k}^{1/2}n^{2/15})$;

(iv) $k_{\min} \ge 6$ *and* $k_{\max} = o(\bar{k}^{1/2}n^{1/4-1/(2k_{\min})})$.

Proof. Under each of the conditions given, a slight weakening of Theorem 2.4 and Corollary 3.4 of [2] shows that the expected number of non-trivial automorphisms of a random member of $\mathcal{G}(\boldsymbol{k})$ is $o(1)$, which implies the theorem. ∎

Note that Theorem 6.1 covers the case of regular graphs of degree k for all $k = k(n)$ such that $3 \le k = o(n^{1/2})$.

Acknowledgement.

We wish to thank Wang Xiaoji for checking all the calculations.

References.

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

[2] B. D. McKay and N. C. Wormald, Automorphisms of random graphs with specified degrees, *Combinatorica*, **4** (1984) 325-338.

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

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