The asymptotic number of labeled graphs with n vertices, q edges, and no isolated vertices

Edward A. Bender
Center for Communications Research
4320 Westerra Court
San Diego, CA 92121
ed@ccrwest.org

E. Rodney Canfield *
Department of Computer Science
University of Georgia
Athens, GA 30602
erc@pollux.cs.uga.edu

Brendan D. McKay
Department of Computer Science
Australian National University
Canberra, ACT 0200, Australia
bdm@cs.anu.edu.au

August 17, 1996

Abstract

Let $d(n,q)$ be the number of labeled graphs with n vertices, $q \le N = \binom{n}{2}$ edges, and no isolated vertices. Let $x = q/n$ and $k = 2q - n$. We determine functions $w_k \sim 1$, $a(x)$, and $\varphi(x)$ such that $d(n,q) \sim w_k \binom{N}{q} e^{n\varphi(x)+a(x)}$ uniformly for all n and $q > n/2$.

Suggested Running Head: Graphs with no isolated vertices.

*Research supported by the National Security Agency.

Please address communications to:

E. Rodney Canfield
Department of Computer Science
University of Georgia
Athens, GA 30602

1 Introduction and statement of results.

For integers n and q, an (n,q)-*graph* is a labeled graph having n vertices and q edges. In a recent paper [1] we studied $c(n,q)$, the number of connected (n,q)-graphs. We proved the following asymptotic formula, with error bound uniform in q,

$$c(n,q) = u_k \binom{N}{q} F(x)^n A(x)(1+o(1)), \tag{1.1}$$

wherein $k = q-n$, $N = \binom{n}{2}$, $x = q/n$, and u_k is a known function with $u_k = 1+O(1/k)$. The functions $F(x)$ and $A(x)$ appearing in (1.1) may be obtained by substituting the expression $\binom{N}{q}F(x)^n A(x)$ for $c(n,q)$ into an exact recursion for $c(n,q)$, rearranging to obtain 1 on one side of the equation, expanding the other side as an asymptotic series, and then "equating coefficients." The last step leads to differential equations involving $F(x)$ and $A(x)$ which turn out to have exact solutions. One may say that (1.1) is the formal asymptotic solution of the recursion satisfied by $c(n,q)$. The proof that the formula so obtained provides a uniformly good estimate of $c(n,q)$ is long and messy.

It is of interest to see if this method of "formal solution" can succeed on other classes of graphs, and also to see if the general form of (1.1) holds for other classes of labeled graphs. The present paper begins this further study. The class of graphs singled out for investigation are the (n,q)-graphs having *no isolated vertices*. The number of such graphs will be denoted $d(n,q)$, "*d*" being both the next letter after "*c*" and also the first letter of the word "dumbbell," which is the typical component for small q. (See Lemma 3.1.) This class is interesting for two reasons. First, the recursion satisfied by $d(n,q)$, (see (1.2) below), is simpler than the nonlinear recursion satisfied by $c(n,q)$, (see [1, (1.11)]). Hence, it may be easier to gain insight into the method from the results on $d(n,q)$ than from those on $c(n,q)$. Second, the functions $F(x)$ and $A(x)$ in (1.1) reduce when $q = \frac{1}{2}n\ln n + \mu n$, to an expression for $c(n,q)$ equivalent to a famous theorem of Erdős and Rényi featured in the classic paper [3]. As is well known in the study of random graphs, the proof of the latter theorem begins by showing that, for the stated range of q, "connected" and "no isolated vertices" are roughly equivalent properties. With a uniform estimate of $d(n,q)$ we can compare these two properties for the entire range of q.

Here is the recursion satisfied by $d(n,q)$, the number of (n,q)-graphs having no isolated vertices, with $N = \binom{n}{2}$

$$qd(n,q) = (N-q+1)d(n,q-1) + n(n-1)d(n-1,q-1) + Nd(n-2,q-1). \tag{1.2}$$

With the boundary conditions $d(0,q) = \delta_{q,0}$ and $d(n,0) = \delta_{n,0}$, the above determines $d(n,q)$. The proof of (1.2) is immediate: the removal of an edge from a graph counted by $d(n,q)$ creates either zero, one, or two isolated vertices, respectively.

In the remainder of this paper, we will use the following notation:

$$\begin{aligned} n &= \text{number of vertices} \\ N &= \binom{n}{2} = \text{number of possible edges} \end{aligned}$$

$$
\begin{aligned}
q &= \text{number of edges} \\
k &= 2q - n \\
x &= q/n \\
y = y(x) &= \begin{cases} \text{the positive solution of } 2xy = -\ln(1-y), & \text{if } x > \frac{1}{2}, \\ \\ 0, & \text{if } x = \frac{1}{2}, \end{cases}
\end{aligned} \quad (1.3)
$$

Thus, associated with the pair (n, q) is a triple of values (k, x, y), and given n, any one of q, k, x, or y determines the other three. If k and n are given rather than q and n, it is always understood that $q = (n+k)/2$ is an integer, that is, *we assume that* $k \equiv n \pmod 2$. Similarly, if x or y is given rather than k or q, it is understood that they are such that q is an integer. By expanding $x = -\ln(1-y)/2y$ as a power series in y, it is clear that $x \mapsto y(x)$ is an increasing bijection from $[1/2, \infty)$ to $[0, 1)$. We use the notation $(n)_s$ for n falling factorial s), that is, the product $n(n-1)\cdots(n-s+1)$.

For $k > 0$ we define

$$
\begin{aligned}
w_k &= \sqrt{2\pi k}\,(k/e)^k/k! \\
d^*(n,q) &= w_k \binom{N}{q} \left(\frac{e^{-2x}y^{1-2x}}{1-y}\right)^n \sqrt{\frac{1-y}{1-2x(1-y)}}\; e^{x+x^2(1-y^2)}.
\end{aligned}
$$

Note that, by Stirling's formula, $w_k = 1 + O(1/k)$. The easily derived alternative expression

$$
\left(\frac{e^{-2x}y^{1-2x}}{1-y}\right)^n = y^{-k}e^{-2q(1-y)}
$$

may be useful, but is not used here. For even n, we define

$$
d^*(n, n/2) = \binom{N}{n/2} e^{3/4-n}\sqrt{2\pi n},
$$

which, in fact, is the limit of $d^*(n,q)$ as $y \downarrow 0$.

Our main goal is to prove

Theorem 1. *Let $\epsilon > 0$ be a real constant and let $n/2 \le q \le N$. Uniformly in q as $n \to \infty$ we have*

$$
d(n,q) = d^*(n,q)\left(1 + O(1/n^{1/7-\epsilon})\right).
$$

Remark. Experimental evidence suggests that the estimate in Theorem 1 has an actual relative error of $O(1/q)$ uniformly over n; by direct computation we have found

$$
\left|\frac{d(n,q)}{d^*(n,q)} - 1\right| < \frac{1.35}{q} \quad \text{for} \quad n \le 160.
$$

We obtain Theorem 1 from the following three theorems, which give better estimates for the error in $d(n,q)/d^*(n,q)$ for various ranges of x.

Theorem 2. *Let $k \geq 0$ and $k = o(n^{2/3})$. Then, uniformly in k as $n \to \infty$,*

$$\begin{aligned} d\left(n, \frac{n+k}{2}\right) &= \sqrt{2}\,\frac{(1/2)^k}{k!} n^{(n+3k)/2} \exp\left\{-\frac{n}{2} + \frac{5k^2}{12n} + O\left(\frac{k+1}{n}\right) + O\left(\frac{k^3}{n^2}\right)\right\} \\ &= d^*(n,q)\left(1 + O\left(\frac{k+1}{n}\right) + O\left(\frac{k^3}{n^2}\right)\right). \end{aligned}$$

Theorem 3. *Let $k \geq 0$. Then, uniformly in q as $n \to \infty$,*

$$d(n,q) = \binom{N}{q}\left(1 + O(ne^{-2x})\right).$$

Uniformly in $x > 3\ln n$ as $n \to \infty$,

$$d(n,q) = d^*(n,q)\left(1 + O(1/n)\right).$$

Theorem 4. *Let $\epsilon > 0$ be a real constant and let $n/2 < q \leq N$. Then, uniformly in q as $n \to \infty$*

$$d(n,q) = d^*(n,q)\left(1 + O(1/k) + O(k^{1/7}/n^{2/7-\epsilon})\right).$$

To obtain Theorem 1, use Theorem 2 for $k \leq n^{2/5-\epsilon}$, Theorem 3 for $k > 6n\ln n$, and Theorem 4 for the remaining range.

Once isolated points are forbidden, there are only finitely many graphs with q edges. We will prove the following two theorems. As for Theorem 1, the relative error in Theorem 5 appears to be $(1/q)$.

Theorem 5. *For $q \geq 1$, denote the number of labeled graphs with q edges and no isolated vertices by*

$$d(q) = \sum_n d(n,q),$$

where the sum is over all n such that $n/2 \leq q \leq N$. For any $\epsilon > 0$,

$$d(q) = C_0(C_1 q)^q(1 + q^{-1/7+\epsilon}),$$

where

$$\begin{aligned} C_0 &= \frac{1}{2^{1+\ln 2/4}\ln 2} \approx 0.6397054049 \\ C_1 &= \frac{2}{(\ln 2)^2 e} \approx 1.5313857152. \end{aligned}$$

Theorem 6. *The number of vertices in a random labeled graph with q edges and no isolated vertices has an asymptotic distribution which is normal with mean $q/\ln 2$ and variance*

$$\frac{1-\ln 2}{2(\ln 2)^2}q.$$

The rest of the paper is organized as follows. Section 2 develops a few facts about the function $y = y(x)$, and some other related functions. Sections 3, 4, and 5 are devoted to Theorems 2, 3, and 4, respectively. We prove Theorem 2 by a combinatorial argument, Theorem 3 by computing the expected number of isolated vertices, and Theorem 4 by induction based on (1.2), using the results of Theorems 2 and 3 for extreme ranges of x. Theorems 5 and 6 are proved in Section 6. In Section 7 we discuss further avenues for exploration.

2 Some analytic facts.

We want our asymptotic estimate of $d(n,q)$ to be in the form $\binom{N}{q}\exp\{n\varphi(x)+a(x)\}$, and so we introduce the functions $\varphi(x)$ and $a(x)$, defined for $x > 1/2$ by

$$\varphi(x) = -2x + (1-2x)\ln y - \ln(1-y) \tag{2.1}$$

and

$$a(x) = x^2(1-y^2) + x + \tfrac{1}{2}\ln(1-y) - \tfrac{1}{2}\ln(1-2x(1-y)). \tag{2.2}$$

In this notation,

$$d^*(n,q) = w_k\binom{N}{q}\exp\{n\varphi(x)+a(x)\}. \tag{2.3}$$

It is clear that as $x \downarrow 1/2$, $\varphi(x) \to -1$ and $a(x) \to \infty$. Our first two lemmas concern relations satisfied by these functions.

Lemma 2.1. *With $\varphi(x)$ defined by (2.1), we have the two relations*

$$y^2 = e^{-\varphi'(x)} \tag{2.4}$$

$$y\left(1+\exp\{-2x-\varphi(x)+x\varphi'(x)\}\right) = 1. \tag{2.5}$$

Proof. We have

$$\begin{aligned}
\varphi(x) &= -2x + (1-2x)\ln y - \ln(1-y) \\
\varphi'(x) &= -2 + \frac{1-2x}{y}\frac{dy}{dx} - 2\ln y + \frac{1}{1-y}\frac{dy}{dx} \\
&= \frac{1}{y}\left(-2y - 2x\frac{dy}{dx} + \frac{1}{1-y}\frac{dy}{dx}\right) - 2\ln y \\
&= -2\ln y,
\end{aligned}$$

because when we differentiate the relation (1.3) with respect to x, we find that

$$-2y - 2x\frac{dy}{dx} + \frac{1}{1-y}\frac{dy}{dx} = 0. \qquad \square$$

We now introduce three functions $g_0(x)$, $g_1(x)$, and $g_2(x)$. How these three functions arise is clarified later in Lemma 5.1. For now our purpose is to record the fact that the function $a(x)$ given in (2.2) satisfies a certain differential equation. We define

$$g_0(x) = \left(\tfrac{1}{2}\varphi''(x) - a'(x)\right) y^2 \tag{2.6}$$

$$g_1(x) = \left(2 - 2x^2 + \tfrac{1}{2}(x-1)^2\varphi''(x) + (x-1)a'(x)\right) 2y(1-y) \tag{2.7}$$

$$g_2(x) = \left(4 - 4x - 4x^2 + \tfrac{1}{2}(2x-1)^2\varphi''(x) + (2x-1)a'(x)\right)(1-y)^2. \tag{2.8}$$

Lemma 2.2. *With $\varphi(x)$, $a(x)$, and $g_i(x)$ defined as above, we have*

$$g_0(x) + g_1(x) + g_2(x) = 0.$$

Proof. We have, since $e^{-\varphi'(x)} = y^2$ by the previous lemma,

$$\varphi''(x) = \frac{-2\,dy/dx}{y} \tag{2.9}$$

and, from (2.2),

$$a'(x) = 2x(1-y^2) - 2x^2y\frac{dy}{dx} + 1 - \frac{dy/dx}{2(1-y)} + \frac{1-y-x\,dy/dx}{1-2x(1-y)}. \tag{2.10}$$

Substitution of these formulas, along with the fact from (1.3) that

$$\frac{dy}{dx} = \frac{2y(1-y)}{1-2x(1-y)},$$

reduces the lemma to a calculation within the field of rational functions of x and y. $\square$

The next two lemmas obtain upper bounds which will be useful later.

Lemma 2.3. *We have, uniformly for $0 \le y < 1$,*

$$1 - y = O(e^{-2x})$$

Proof. Since $y \to 1$ as $x \to \infty$, the function $x(1-y) = xe^{-2xy}$ is uniformly bounded. Hence,

$$1 - y = e^{-2x}e^{2x(1-y)} = O(e^{-2x}). \qquad \square$$

Lemma 2.4. *With $\varphi(x)$ and $a(x)$ defined by (2.1) and (2.2), we have, uniformly for $0 \le y < 1$,*

$$\varphi''(x) = O(1/y), \quad \varphi'''(x) = O(1/y^2), \quad a'(x) = O(1/y), \quad \text{and} \quad a''(x) = O(1/y^2).$$

Proof. During this proof let $Z = 1 - 2x(1-y)$. Because $dy/dx = (1-y)2y/Z$, we find that the class $\mathcal{R}$ of all functions of the form

$$(1-y)\frac{p(x,y)}{Z^m},$$

in which $p(x,y)$ is a polynomial and m is a nonnegative integer, is closed under $\frac{d}{dx}$. From (2.9) and (2.10) we see, since dy/dx and dZ/dx are in $\mathcal{R}$, that $a''(x)$, $\varphi''(x)$, and all higher derivatives of both functions belong to $\mathcal{R}$. Any function $h(x)$ belonging to the class $\mathcal{R}$ will satisfy, in view of Lemma 2.3, $h(x) \to 0$ as $x \to \infty$. Although $a'(x)$ does not qualify for membership in $\mathcal{R}$, it is clearly bounded for $y \ge 1/2$. Hence, $a'(x), a''(x), \varphi''(x)$, and $\varphi'''(x)$ are all bounded for $y \ge \frac{1}{2}$. Note that in the range $y \le 1/2$ each of $\varphi''(x)$ and $a'(x)$ is expressable as $1/y$ times a power series in y convergent for $y < 1$. The lemma follows. □

The final lemma of this section will play an important role later. It is a bit different from the other four lemmas in that the variables k and n are again involved.

Lemma 2.5. *Let A and B be real constants with $1 \ge A > B/2 \ge 0$. There is a constant c_1 such that, uniformly in A, B, and $k \ge 1$,*

$$\frac{Ay}{k} - \frac{B(1-y)}{n} \ge \begin{cases} \dfrac{c_1(A - B/2)}{n}, & \text{if } y \le 1/2, \\[2mm] \dfrac{c_1(A - B/2)}{k}, & \text{if } y \ge 1/2. \end{cases}$$

Proof. Since $2x - 1 = k/n$, the quantity in question may be written

$$\frac{Ay - B(1-y)(2x-1)}{k},$$

which may be expanded as a power series in y:

$$\frac{1}{k}\left(\left(A - \frac{B}{2}\right)y + B\sum_{m=2}^{\infty}\frac{y^m}{m(m+1)}\right).$$

This proves the lemma for the case $y \ge 1/2$. For $y \le 1/2$ we observe from the expansion of $2x - 1 = k/n$ as a power series in y that y must be greater than some constant times k/n. The lemma follows. □

We remark, but will not use, that c_1 in the previous lemma can be taken equal to $1/2$.

3 The proof of Theorem 2.

The proof of Theorem 2 appears after we state and prove five preliminary lemmas. Throughout this section we let $\mathcal{D}(n, q)$ be the class of graphs with n vertices and q edges having no isolated vertices; thus,

$$|\mathcal{D}(n, q)| = d(n, q).$$

We shall see that when $k = o(n^{2/3})$, most graphs in $\mathcal{D}(n, q)$ contain only four types of components: a single edge, a path with three vertices and two edges, a star with a central vertex joined to three others, and a path with four vertices and three edges. These are the four possible trees with four or less vertices, and we shall refer to them by the names $K_2, P_3, K_{1,3}$, and P_4, respectively. Lemmas 3.2 and 3.4 below are examples of "switching arguments." Switching has proven to be a useful enumerative tool, especially in asymptotic enumeration where it eliminates hard to estimate sums with alternating signs from inclusion/exclusion. No survey exposition has appeared yet; see however [4] and [5] for early examples.

Lemma 3.1. *Any graph G belonging to the class $\mathcal{D}(n, q)$ has at least $n - 3k$ vertices in K_2 components.*

Proof. Letting N_1 denote the number of vertices in question, and N_2 the rest, we have

$$N_1 + N_2 = n.$$

Every component containing vertices of the second class has an edge/vertex ratio of 2/3 or more; hence,

$$\frac{N_1}{2} + \frac{2N_2}{3} \leq \frac{n + k}{2}.$$

The lemma follows easily. □

The *corank* of a graph $G = (V, E)$ having c components is $|E| - |V| + c$, which is the dimension of the cycle space of G [2, p.36]. In particular, a graph is a forest if and only if its corank is zero. We now write

$$\mathcal{D}(n, q) = \mathcal{M}_0 \cup \mathcal{M}_1 \cup \cdots,$$

$\mathcal{M}_h$ being the class of graphs in $\mathcal{D}(n, q)$ having corank equal to h.

Lemma 3.2. *Let the classes $\mathcal{M}_0, \mathcal{M}_1, \ldots$ be defined as above, $k = o(n^{2/3})$, and $n \to \infty$. Then, uniformly in h and k,*

$$|\mathcal{M}_h| = O(k^3/(hn^2))\, |\mathcal{M}_{h-1}|.$$

Consequently, all but $O(k^3/n^2)$ of the graphs in $\mathcal{D}(n, q)$ are cycle-free.

Proof. Given a graph in $\mathcal{M}_h$, remove an edge which belongs to a cycle and use it to join two K_2's into a P_4. The resulting graph belongs to $\mathcal{M}_{h-1}$. Since the corank is the dimension of the cycle space, there are at least h edges which belong to a cycle. Thus the operation of removing an edge from a cycle and joining two K_2's may be done in at least

$$\tfrac{1}{2}h \cdot (n-3k)(n-3k-2)$$

ways. A given graph in $\mathcal{M}_{h-1}$ is obtained by such an operation in at most $(3k/4)\binom{3k}{2}$ ways. In the latter estimate, the first factor bounds the number of P_4 components and the second factor bounds the choices of two vertices in the same component which are not joined by an edge. The first assertion of the lemma now follows easily, and the second assertion is obtained by summing. □

The next lemma is an easy consequence of the Prüffer algorithm [6, p. 229].

Lemma 3.3. *Let $\mathcal{L}$ be an ordered set of labels with $L = |\mathcal{L}| \geq 5$. Then there is an injection from $\mathcal{R}_1$ to $\mathcal{R}_2$, where*

$\mathcal{R}_1 = \{T : T$ *is a tree on the set* $\mathcal{L}\}$,

$\mathcal{R}_2 = \Big\{(T, X_1, X_2, \dots, X_{L-4}) : T$ *is a rooted tree with three vertices from the set* $\mathcal{L}$, *and each* $X_i \in \mathcal{L}\Big\}$.

Proof. As in the usual Prüfer algorithm, prune the given tree T of one leaf at a time, always pruning the leaf with the smallest label. Each time a leaf is removed, write down in sequence the vertex to which it was attached, except for the $(L-3)$-rd, which is the last. When the $(L-3)$-rd vertex is removed, let the point of its attachment become the root of the remaining tree of size 3. That this process is injective follows from the usual Prüfer bijection. □

The algorithm for realizing the injection of Lemma 3.3 will be referred to as the "partial Prüfer algorithm," since it amounts to applying the usual algorithm and stopping just a few steps early.

We now write

$$\mathcal{M}_0 = \mathcal{N}_0 \cup \mathcal{N}_1 \cup \cdots,$$

$\mathcal{N}_h$ being those cycle-free graphs in the class $\mathcal{D}(n, q)$ having h components of size 5 or greater.

Lemma 3.4. *Let $\mathcal{N}_0, \mathcal{N}_1, \dots$ be the class of graphs defined above, $k = o(n^{2/3})$, and $n \to \infty$. Then, uniformly in k and h,*

$$|\mathcal{N}_h| = O\left(k^3/(hn^2)\right)\, |\mathcal{N}_{h-1}|.$$

Consequently, all but $O(k^3/n^2)$ of the graphs in $\mathcal{D}(n, q)$ are forests whose components belong to the set $\{K_2, P_3, K_{1,3}, P_4\}$.

Proof. First, we claim there is a injection from $\mathcal{S}_1$ to $\mathcal{S}_2$ where

$\mathcal{S}_1 = \Big\{(G, C, X_1, X_2, \ldots, X_{L-3}) : G$ is a graph in $\mathcal{N}_h$, C is a component of G having $L \geq 5$ vertices, each X_i is an endpoint of a K_2 component of G, and no two X_i belong to the same K_2 component$\Big\}$

$\mathcal{S}_2 = \Big\{(G, T, \{Y_1, Y_2, \ldots, Y_{L-3}\}, Z_1, Z_2, \ldots Z_{L-4}) : G$ is a graph in $\mathcal{N}_{h-1}, T$ is a component of G which is a tree of size 3, T has been rooted, the set $\{Y_i\}$ is an unordered collection of endpoints of P_3 components of G, no two Y_i belonging to the same P_3 component, each Z_j is either a vertex of T or one of the Y_i, and repitition is allowed among the $Z_j\Big\}$.

Although the sets $\mathcal{S}_1$ and $\mathcal{S}_2$ are lengthy in description, the bijection is not: Given $(G, C, \ldots)$ in $\mathcal{S}_1$, apply the partial Prüfer algorithm to the tree C, obtaining a rooted tree T of size 3 and an $(L-4)$-tuple $(Z_1, Z_2, \ldots, Z_{L-4})$ of points of attachment; the set $Y_1, Y_2, \ldots, Y_{L-3}$ is the set of leaves removed from C; each is attached, in turn, to the corresponding ordered X_i yielding a P_3 component of which Y_i is an endpoint; the order of removal is then forgotten.

As a consequence of this injection we have, with $\mathcal{N}_h^{(L)}$ denoting the class of graphs in $\mathcal{N}_h$ having a component of size $L \geq 5$ distinguished,

$$\begin{aligned} \left|\mathcal{N}_h^{(L)}\right| \prod_{i=0}^{L-4}(n-3k-2i) &\leq |\mathcal{S}_1| \leq |\mathcal{S}_2| \\ &\leq \frac{1}{(L-3)!}\Big(\prod_{i=0}^{L-4}(2k-2i)\Big)(3k)\cdot L^{L-4}\cdot|\mathcal{N}_{h-1}| \end{aligned}$$

The leftmost inequality follows from the fact that, given $(G, C) \in \mathcal{N}_h^{(L)}$, we have by Lemma 3.1 at least $n - 3k$ vertices from which the $X_1, X_2, \ldots, X_{L-3}$ may be selected. The rightmost inequality follows from the fact that, given $G \in \mathcal{N}_{h-1}$, we have by Lemma 3.1 at most $2k$ endpoints of P_3's from which to choose Y_i, at most $3k$ choices for a root of a tree of size 3, and of course L^{L-4} choices for the Z_j. In all of the above, L can be at most $3k$. Hence, uniformly in k, h, and L,

$$\left|\mathcal{N}_h^{(L)}\right| = O\left(\frac{L^{L-4}}{(L-3)!}\left(\frac{2k}{n-3k}\right)^{L-3} 3k\right)|\mathcal{N}_{h-1}|. = O\left(\frac{L^{L-4}}{(L-3)!}\,\frac{(3k)^{L-2}}{n^{L-3}}\right)|\mathcal{N}_{h-1}|.$$

When we sum the above for $L \geq 5$ we obtain

$$h\,|\mathcal{N}_h| = O(k^3/n^2)\,|\mathcal{N}_{h-1}|.$$

This is equivalent to the first assertion of the lemma. The second follows by summing on h, and using Lemma 3.2. □

Lemma 3.5. *Let $k \geq 1, k = o(n^{2/3})$, and $n \to \infty$. Then, uniformly in k,*

$$\sum_{s > 4k^{1/2}} \frac{(8k^2/3n)^s}{s!} = O(k^3/n^2).$$

Proof. Considering ratios of consecutive terms, the sum is $O(1)$ times the first term. Using $s! > (s/e)^s$ and $4k^{1/2} \geq 2$ completes the proof of the lemma. $\square$

We are now ready for the proof of Theorem 2.

Proof. (of Theorem 2) When $k = 0$, and n is even, we have

$$d(n, n/2) = \frac{n!}{(n/2)!\,2^{n/2}} = \sqrt{2}\,n^{n/2}e^{-n/2}(1 + O(1/n)),$$

which is consistent with the first equality in the theorem. The second equality follows easily.

Henceforth in the proof we assume $k \geq 1$, and note, uniformly in k,

$$\begin{aligned} \frac{n!}{\left(\frac{n-3k}{2}\right)!\,k!\,2^{(n-k)/2}} &= \sqrt{2}\,\frac{(1/2)^k}{k!}n^{(n+3k)/2}\exp\{-n/2 - 9k^2/4n\} \\ &\quad \times \left(1 + O(k/n) + O(k^3/n^2)\right), \end{aligned} \tag{3.1}$$

where we have used $k = o(n^{2/3})$. Consider a graph G whose every component is one of K_2, P_3, $K_{1,3}$, or P_4. If the graph G contains s components of size 4, then it must contain $k - 2s$ of size 3 and $s + (n-3k)/2$ of size 2. In view of Lemma 3.4 and the fact that there are n^{n-2} unrooted, labeled n-vertex trees, we have

$$\begin{aligned} d(n, q) &= \sum_{0\leq s\leq k/2} \frac{n!\,(3^{3-2})^{k-2s}(4^{4-2})^s\,(1 + O(k^3/n^2))}{(2!)^{s+(n-3k)/2}\,\left(s + \frac{n-3k}{2}\right)!\,(3!)^{k-2s}\,(k-2s)!\,(4!)^s\,s!} \\ &= \frac{n!\ (1 + O(k^3/n^2))}{\left(\frac{n-3k}{2}\right)!\,k!\,2^{(n-k)/2}} \sum_{0\leq s\leq k/2} t_s, \end{aligned} \tag{3.2}$$

where

$$t_s = \frac{(k)_{2s}(4/3)^s}{(s + (n-3k)/2)_s\,s!}.$$

Since

$$\frac{t_s}{t_{s-1}} \leq \frac{8k^2}{3s(n-3k)},$$

it is readily seen that $t_s \leq (8k^2/3(n-3k))^s/s!$ and so

$$\sum_{k/2\geq s>4k^{1/2}} t_s = O(1)\left(\frac{2ek^{3/2}}{3(n-3k)}\right)^{4k^{1/2}} = O(k^3/n^2),$$

the first bound following from the facts that the sum is $O(1)$ times its first term ($k = o(n^{2/3})$) and that $s! \geq (s/e)^s$, and the second bound from the fact that $4k^{1/2} > 2$. Because $t_0 = 1$,

$$\sum_{0\leq s\leq k/2} t_s = \left(\sum_{0\leq s\leq 4k^{1/2}} t_s\right)\left(1 + O(k^3/n^2)\right). \tag{3.3}$$

Uniformly for $0 \le s \le 4k^{1/2}$, we have

$$\begin{aligned} (k)_{2s} &= k^{2s}\left(1+O(s^2/k)\right), \\ (s+(n-3k)/2)_s &= ((n-3k)/2)^s\left(1+O(s^2/n)\right) \\ &= (n/2)^s\left(1+O(ks/n)\right), \end{aligned}$$

and

$$t_s = (8k^2/3n)^s(s!)^{-1}\left(1+O(s^2/k)+O(ks/n)\right).$$

The sum of the right side over $s \ge 0$ is $\exp\{8k^2/3n\}\left(1+O(k^3/n^2)\right)$. Invoking Lemma 3.5, we find

$$\sum_{0\le s\le 4k^{1/2}} t_s = \exp\{8k^2/3n\}\left(1+O(k^3/n^2)\right). \tag{3.4}$$

The first equality of the theorem, for $k \ge 1$, now follows by combining (3.1), (3.2), (3.3), and (3.4), noting $-\frac{9}{4}+\frac{8}{3}=\frac{5}{12}$. We have the following uniform estimates, which follow from Stirling's formula and the definition of y:

$$\begin{aligned} \binom{N}{\frac{n+k}{2}} &= \frac{n^{(n+k)/2}}{\sqrt{\pi n}}\exp\{n/2-3/4-k^2/4n+O(k/n)+O(k^3/n^2)\}, \\ y &= (2k/n)(1-4k/3n)\left(1+O(k^2/n^2)\right), \\ y^{(1-2x)n} &= y^{-k} = (n/2k)^k\exp\left\{4k^2/3n+O(k^3/n^2)\right\}, \\ (1-y)^{-n} &= \exp\left\{2k-2k^2/3n+O(k^3/n^2)\right\}, \\ 1-2x(1-y) &= (k/n)\left(1+O(k/n)\right), \end{aligned}$$

and

$$x^2(1-y^2)+x = 3/4+O(k/n).$$

Putting the above together yields the second equality in Theorem 2. □

4 The proof of Theorem 3.

The probability that there is an isolated vertex in a randomly chosen (n,q)-graph is no greater than the expected number of isolated vertices. With X denoting the random variable which counts isolated vertices, we calculate

$$\begin{aligned} E(X) &= \frac{n\binom{\binom{n-1}{2}}{q}}{\binom{N}{q}} = n\frac{\left(\binom{n-1}{2}\right)_q}{(N)_q} \\ &\le n\frac{\binom{n-1}{2}^q}{N^q} = n\left(1-\frac{2}{n}\right)^q \le ne^{-2x}, \end{aligned}$$

and the first part of Theorem 3 follows. The following lemma completes the proof of the theorem and provides a tighter error bound.

Lemma 4.1. *Let $x > 3\ln n$, and $n \to \infty$. Then, uniformly in q,*

$$d(n,q) = \binom{N}{q} \exp\{n\varphi(x) + a(x)\} \left(1 + O(1/n^4)\right).$$

Proof. For large x we have the following, using Lemma 2.3 and (1.3),

$$\begin{aligned} y &= 1 + O(e^{-2x}), \\ y^{(1-2x)n} &= y^{-k} = 1 + O(ke^{-2x}), \\ 1 - y &= e^{-2xy} = e^{-2x}e^{2x(1-y)} = e^{-2x}\left(1 + O(xe^{-2x})\right), \\ \left(\frac{e^{-2x}}{1-y}\right)^n &= 1 + O(nxe^{-2x}), \\ \sqrt{1-y} &= e^{-x}\left(1 + O(xe^{-2x})\right), \end{aligned}$$

and

$$e^{x^2(1-y^2)+x}\sqrt{\frac{1-y}{1-2x(1-y)}} = 1 + O(x^2e^{-2x}),$$

leading to

$$\exp\{n\varphi(x) + a(x)\} = 1 + O(qe^{-2x}).$$

Comparing with the first part of Theorem 3, the lemma follows. □

5 The proof of Theorem 4.

Let the two dimensional array $b(n,k)$ be defined by the equation

$$d(n,q) = \binom{N}{q} e^{n\varphi(x)+a(x)} \left(1 + b(n,k)\right), \tag{5.1}$$

where $\varphi(x)$ and $a(x)$ are given by (2.1) and (2.2). Our object is to establish an upper bound on $b(n,k)$. Throughout this section we shall use the three inequalities

$$y \le 1/2, \quad x \le \ln 2, \quad \text{and} \quad k \le (2\ln 2 - 1)n$$

interchangeably, without repeatedly remarking on the equivalence. We define the function $\Lambda = \Lambda(n,q)$ by

$$\Lambda = \begin{cases} 1/k, & \text{if } y \le 1/2, \\ 1/n, & \text{if } y > 1/2. \end{cases}$$

Substituting (5.1) into the recurrence (1.2) and dividing through by $q\binom{N}{q}\exp\{n\varphi(x)+a(x)\}$, we find

$$\begin{aligned} 1 + b(n,k) &= W_0(1 + b(n,k-2)) \\ &\quad + W_1(1 + b(n-1,k-1)) \\ &\quad + W_2(1 + b(n-2,k)), \end{aligned} \tag{5.2}$$

where, for example,

$$W_1 = \frac{n(n-1)\binom{\binom{n-1}{2}}{q-1}}{q\binom{N}{q}} \exp\left\{(n-1)\varphi\left(\tfrac{q-1}{n-1}\right) + a\left(\tfrac{q-1}{n-1}\right) - n\varphi(x) - a(x)\right\},$$

and similar quotients may be written for W_0 and W_2. The object of the next lemma is to estimate each quotient W_i.

Lemma 5.1. *Let the three quotients W_0, W_1, and W_2 be defined as above in (5.2), let the functions $g_i(x)$ be as introduced in (2.6) – (2.8), and let the "error terms" $e_i = e_i(n,q)$ be defined so that*

$$\begin{aligned} W_0 &= y^2 + \frac{g_0(x)}{n} + e_0(n,q) \\ W_1 &= 2y(1-y) + \frac{g_1(x)}{n} + e_1(n,q) \\ W_2 &= (1-y)^2 + \frac{g_2(x)}{n} + e_2(n,q). \end{aligned}$$

Finally, let $n^{2/5} \leq k = o(n^{3/2})$, and $n \to \infty$. Then, uniformly in k, the error terms e_i satisfy

$$e_i = O(x^4\Lambda^2).$$

Proof. We shall write down details for W_1. The proofs for W_0 and W_2 are very similar. Starting from the identity $q\binom{N}{q} = N\binom{N-1}{q-1}$, it follows that

$$\begin{aligned} \frac{n(n-1)\binom{\binom{n-1}{2}}{q-1}}{q\binom{N}{q}} &= 2\left(\frac{(n-1)(n-2)}{n(n-1)-2}\right)^{q-1} \prod_{i=0}^{q-2} \frac{1-i\big/\binom{n-1}{2}}{1-i/(N-1)} \\ &= 2\exp\left\{-2x + \tfrac{2-2x^2}{n} + O(q^4/n^6)\right\}. \end{aligned} \tag{5.3}$$

Since $(q-1)/(n-1) = x + (x-1)/(n-1)$, we have, by Taylor's formula with remainder,

$$\begin{aligned} &\exp\left\{(n-1)\varphi\left(\tfrac{q-1}{n-1}\right) + a\left(\tfrac{q-1}{n-1}\right) - n\varphi(x) - a(x)\right\} \\ &\quad = \exp\left\{-\varphi(x) + (x-1)\varphi'(x) + \tfrac{(x-1)^2\varphi''(x)}{2n} + \tfrac{(x-1)a'(x)}{n} + E_1\right\}, \end{aligned} \tag{5.4}$$

where

$$\begin{aligned} E_1 &= \frac{\frac{1}{2}(x-1)^2\varphi''(x) + (x-1)a'(x)}{n(n-1)} \\ &\quad + \frac{(x-1)^3\varphi'''(\xi)}{6(n-1)^2} + \frac{(x-1)^2 a''(\zeta)}{2(n-1)^2}, \end{aligned}$$

with ξ and ζ known to be between x and $x + (x-1)/(n-1)$. From

$$(2x-1) = \frac{y}{2} + \frac{y^2}{3} + \cdots < \frac{y/2}{1-y}$$

we see that

$$y \geq (2x - 1) = k/n \quad \text{for } y \leq 1/2. \tag{5.5}$$

If y is divided by 2 the effect on the corresponding x is to reduce it by more than $y/4$. Using this, (5.5), and $k \geq n^{2/5}$ when $y \leq 1/2$, and easier reasoning when $y \geq 1/2$, we see that the y values associated with ξ and ζ are at least $y/2$; by Lemma 2.4 then

$$\varphi'''(\xi) = O(1/y^2), \quad a''(\zeta) = O(1/y^2).$$

Applying Lemma 2.4 to the other two terms in E_1 we conclude

$$E_1 = O(x^3\Lambda^2).$$

Again by Lemma 2.4, $\left(\frac{1}{2}(x-1)^2\varphi''(x) + (x-1)a'(x)\right)\big/ n = O(x^2\Lambda)$. We may thus expand the "$\exp\{\cdots\}$" term in (5.4). Recalling that $y^2 = e^{-\varphi'(x)}$, we find

$$\begin{aligned} &\exp\Big\{(n-1)\varphi\Big(\tfrac{q-1}{n-1}\Big) + a\Big(\tfrac{q-1}{n-1}\Big) - n\varphi(x) - a(x)\Big\} \qquad (5.6)\\ &\quad = y^2 \exp\{-\varphi(x) + x\varphi'(x)\}\left(1 + \frac{(x-1)^2\varphi''(x)}{2n} + \frac{(x-1)a'(x)}{n} + O(x^4\Lambda^2)\right). \end{aligned}$$

Combining (5.3) and (5.6) yields the desired estimate of W_1. The quotients W_0 and W_2 may be handled similarly, and the lemma is complete. □

The next lemma defines and bounds five additional error terms which are needed in the proof of Theorem 4.

Lemma 5.2. *Let A and B be real constants with $0 < A, B < 1$ and let W_0, W_1 and W_2 be as in (5.2). Let $e_i = e_i(n, q, A, B)$, $3 \leq i \leq 7$, be defined by the equations*

$$\begin{aligned} \frac{(k-2)^A}{n^B} &= \frac{k^A}{n^B}\left(1 - \frac{2A}{k} + e_3\right)\\ \frac{(k-1)^A}{(n-1)^B} &= \frac{k^A}{n^B}\left(1 - \frac{A}{k} + \frac{B}{n} + e_4\right)\\ \frac{k^A}{(n-2)^B} &= \frac{k^A}{n^B}\left(1 + \frac{2B}{n} + e_5\right)\\ W_0 + W_1 + W_2 &= 1 + e_6\\ W_0\left(1 - \frac{2A}{k} + e_3\right) &+ W_1\left(1 - \frac{A}{k} + \frac{B}{n} + e_4\right) + W_2\left(1 + \frac{2B}{n} + e_5\right)\\ &= 1 - \frac{2Ay}{k} + \frac{2B(1-y)}{n} + e_7. \end{aligned}$$

Finally, let $n^{2/5} \leq k = o\left(n^{3/2}\right)$ and $n \to \infty$. Then, uniformly in A, B, and q,

$$e_i = O\left(x^4\Lambda^2\right), \quad \textit{for } 3 \leq i \leq 7.$$

Proof. The assertions about e_3, e_4, and e_5 are very simple, and that about e_6 follows from the preceding lemma and the fact (Lemma 2.2) that $g_0(x)+g_1(x)+g_2(x) = 0$. There remains e_7. Expand the left side of the equation which defines e_7 to obtain the four quantities displayed here:

$$\begin{aligned} W_0\left(1-\frac{2A}{k}+e_3\right) &+ W_1\left(1-\frac{A}{k}+\frac{B}{n}+e_4\right)+W_2\left(1+\frac{2B}{n}+e_5\right) \\ &= (W_0+W_1+W_2) - \frac{A}{k}(2W_0+W_1) \\ &+ \frac{B}{n}(W_1+2W_2)+(W_0e_3+W_1e_4+W_2e_5). \end{aligned}$$

The first term on the right is $1+e_6$. For the second and third terms, we note first by Lemma 5.1

$$2W_0+W_1 = 2y+\frac{2g_0(x)+g_1(x)}{n}+2e_0(n,q)+e_1(n,q)$$

$$W_1+2W_2 = 2(1-y)+\frac{g_1(x)+2g_2(x)}{n}+e_1(n,q)+2e_2(n,q).$$

Using Lemma 2.4 we check that all three of $g_i(x)$ are uniformly bounded, as are the W_i by Lemma 5.1. Since both $1/k$ and $1/n$ are $O(\Lambda)$, we obtain the desired bound on e_7 from the known bounds on $e_0, \dots e_6$. This concludes the proof of Lemma 5.2. □

Proof of Theorem 4. Let $0 < \epsilon < 2/7$ be given. Define $A = 1/7$ and $B = 2/7 - \epsilon$. Since $1/k \le k^A/n^B$ for $k \ge n^{2/5}$, it suffices to exhibit a constant C sufficiently large that

$$|b(n,k)| \le C\left(\frac{1}{k+1}+\frac{k^A}{n^B}\right) \quad \text{for } k \le n^{2/5} \tag{5.7}$$

$$|b(n,k)| \le C\frac{k^A}{n^B} \quad \text{for } n^{2/5} \le k \le n^2-2n. \tag{5.8}$$

Let c_1 be the constant given in Lemma 2.5. By Lemmas 5.1 and 5.2, there is a sufficiently large c_2 such that $|e_i| \le c_2x^4\Lambda^2$ for $0 \le i \le 7$ and $n^{2/5} \le k = o(n^{3/2})$. With these two constants and ϵ known in advance, we claim that C may be chosen as follows:

C1. Choose n_0 sufficiently large that, for all $n \ge n_0$,

- **(a)** $n^{3/5} \ge n^{2/5}+2$,
- **(b)** $(324c_2/c_1\epsilon)(\ln n)^4 \le n^{1/5}$, and
- **(c)** $(1944c_2/c_1\epsilon)(\ln n)^5 \le n$.

C2. Choose C sufficiently large that (5.7) and (5.8) hold for the finitely many pairs (n,q) with $n < n_0$.

C3. Choose C sufficiently large, by Theorem 2, that (5.7) and (5.8) hold provided $k \le n^{3/5}$.

C4. Choose C sufficiently large, by Lemma 4.1, that (5.8) holds provided $x \ge 3\ln n$.

C5. Choose C sufficiently large that

(a) $(324c_2/c_1\epsilon)((\ln n)^4/n^{\epsilon}) \le C$ for $n \ge n_0$.

(b) $(1944c_2/c_1\epsilon)((\ln n)^5/n^{4/5+\epsilon}) \le C$ for $n \ge n_0$.

We now prove that (5.7) and (5.8) hold for this choice of C, using proof by contradiction. Assume that the set of pairs (n,q) for which one of (5.7) or (5.8) fails is nonempty, and choose one such pair which is minimal with respect to the product partial order on $\mathbb{N}\times\mathbb{N}$. By Conditions C2, C3, and C4 we must have $n \ge n_0, k > n^{3/5}$, and $x \le 3\ln n$. Because $(n, q-1)$, $(n-1, q-1)$, and $(n-2, q-1)$ are all smaller than (n,q) in the product partial order, and because Condition C1(a) implies $k-2 \ge n^{2/5}$, $k-1 \ge (n-1)^{2/5}$, and $k \ge (n-2)^{2/5}$, we will have, by minimality of our counterexample,

$$\begin{aligned} b(n,k-2) &\le C\frac{(k-2)^A}{n^B} \\ b(n-1,k-1) &\le C\frac{(k-1)^A}{(n-1)^B} \\ b(n-2,k) &\le C\frac{k^A}{(n-2)^B}. \end{aligned}$$

Denoting the three quantities $b(n,k-2)$, $b(n-1,k-1)$, and $b(n-2,k)$ by b_0, b_1, and b_2 respectively, we have from (5.2)

$$|b(n,k)| \le \left|\sum W_i - 1\right| + \left|\sum W_i b_i\right| \le \left|\sum W_i - 1\right| + \sum W_i\,|b_i|\,,$$

since $W_i \ge 0$. By the definitions of e_i in Lemma 5.2, $\sum W_i - 1 = e_6(n,q)$ and

$$\begin{aligned} \sum W_i|b_i| &\le W_0\cdot C\frac{k^A}{n^B}\left(1-\frac{2A}{k}+e_3\right) + W_1\cdot C\frac{k^A}{n^B}\left(1-\frac{A}{k}+\frac{B}{n}+e_4\right) \\ &\quad + W_2\cdot C\frac{k^A}{n^B}\left(1+\frac{2B}{n}+e_5\right) \\ &= C\frac{k^A}{n^B}\left(1-\frac{2Ay}{k}+\frac{2B(1-y)}{n}+e_7(n,q)\right), \end{aligned}$$

and so, by Lemma 5.2,

$$|b(n,k)| \le c_2x^4\Lambda^2 + C\frac{k^A}{n^B}\left(1-\frac{2Ay}{k}+\frac{2B(1-y)}{n}+c_2x^4\Lambda^2\right). \tag{5.9}$$

In terms of Λ and ϵ the conclusion of Lemma 2.5 may be expressed

$$\frac{2Ay}{k} - \frac{2B(1-y)}{n} \ge \frac{c_1\epsilon}{\Lambda nk}.$$

When $y \le \frac{1}{2}$, $\Lambda = 1/k$, and $k \ge n^{3/5}$, Condition C1(b) implies

$$c_2x^4\Lambda^2 \le \frac{c_1\epsilon}{4\Lambda nk}$$

and Condition C5(a) implies

$$c_2 x^4 \Lambda^2 \le C \frac{k^A}{n^B} \frac{c_1 \epsilon}{4\Lambda n k}.$$

When $y > 1/2$, $\Lambda = 1/n$, and $x \le 3 \ln n$, Condition C1(c) implies

$$c_2 x^4 \Lambda^2 \le \frac{c_1 \epsilon}{4\Lambda n k},$$

and Condition C5(b) implies

$$c_2 x^4 \Lambda^2 \le C \frac{k^A}{n^B} \frac{c_1 \epsilon}{4\Lambda n k}.$$

Thus, from (5.9),

$$\begin{aligned} |b(n,k)| &\le C \frac{k^A}{n^B} \frac{c_1 \epsilon}{4\Lambda n k} + C \frac{k^A}{n^B} \left(1 - \frac{c_1 \epsilon}{\Lambda n k} + \frac{c_1 \epsilon}{4\Lambda n k}\right) \\ &= C \frac{k^A}{n^B} \left(1 - \frac{c_1 \epsilon}{2\Lambda n k}\right), \end{aligned}$$

contradicting the assumption that $|b(n,k)| > C k^A / n^B$. This completes the proof. □

6 The proof of Theorems 5 and 6.

We require the following estimate for $d(n+t, q)$.

Lemma 6.1. *Fix $\epsilon > 0$. Then uniformly for $\frac{1}{2} + \epsilon \le x = O(1)$ and $|t| \le q^{2/3}$, we have*

$$\ln \frac{d(n+t,q)}{\binom{N}{q} e^{n\varphi(x)+a(x)}} = t(2x + \varphi(x) - x\varphi'(x)) + t^2 x^2 q^{-1}(-1 + \tfrac{1}{2} x \varphi''(x)) + O(q^{-1/7+\epsilon} + |t|^3 q^{-2}).$$

Proof. By writing

$$(K)_q = \prod_{i=-r}^{r} (K - r + i) = \prod_{i=-r}^{r} ((K-r)^2 - i^2)^{1/2}$$

with $r = (q-1)/2$ and routine expansion, we find that

$$\ln \frac{\binom{\binom{n+t}{2}}{q}}{\binom{\binom{n}{2}}{q}} = 2tx - \frac{t x^2 (t - 2x - 1)}{q} + O\left(\frac{|t|^3 + 1}{q^2}\right).$$

From Theorem 1, (2.4), (2.9) and the fact that $e^{a(x)}$ and its derivative are bounded for $x \in [\frac{1}{2} + \epsilon, \infty]$, we have

$$\begin{aligned} \ln \frac{d(n+t,q)}{\binom{N}{q} e^{n\varphi(x)+a(x)}} &= 2tx - t^2 x^2 / q + u(n+t)\varphi'(x) + \tfrac{1}{2} u^2 (n+t) \varphi''(x) + t\varphi(x) \\ &\quad + O(q^{-1/7+\epsilon} + n|u|^3 + (1+|t|)q^{-1} + |t|^3 q^{-2}), \end{aligned}$$

where $u = q/(n+t) - q/n$. Substituting $u = -t x^2 / q + O(t^2 x^3 q^{-2})$, we have the required expansion. □

We now prove Theorems 5 and 6.

Note first that $x_0 = \ln 2$ is a solution to the equation $2x + \varphi(x) - x\varphi'(x) = 0$. Define $n_0 = q/x_0$. Substituting into Lemma 6.1, we find that for $|t| \le q^{2/3}$ and integer $n_0 + t$, we have

$$\begin{aligned} d(n_0+t,q) &= \binom{\binom{n_0}{2}}{q} \\ &\quad e^{n_0\varphi(x_0)+a(x_0)} \exp\left(-\frac{(\ln 2)^2 t^2}{(1-\ln 2)q}\right)(1+O(q^{-1/7+\epsilon}+|t|^3 q^{-2})). \end{aligned} \tag{6.1}$$

Applying Euler-Maclaurin summation, we find that

$$\sum_{|t|\le q^{2/3}} d(n_0+t,q) = \binom{\binom{n_0}{2}}{q} e^{n_0\varphi(x_0)+a(x_0)} \frac{\sqrt{\pi(1-\ln 2)q}}{\ln 2}(1+O(q^{-1/7+\epsilon})).$$

Finally, we substitute the values of x_0 and n_0 with the aid of the expansion

$$\binom{\binom{n_0}{2}}{q} = \left(\frac{eq}{2x_0^2}\right)^q (2\pi q)^{-1/2} e^{-x_0-x_0^2}(1+O(q^{-1}))$$

to obtain

$$\sum_{|t|\le q^{2/3}} d(n_0+t,q) = C_0(C_1 q)^q (1+O(q^{-1/7+\epsilon})). \tag{6.2}$$

It remains to be shown that larger values of t do not significantly contribute. We begin by establishing a log-concavity result. For any $q > 0$ and all n, define

$$\alpha(n,q) = \begin{cases} \binom{N}{q} e^{n\varphi(q/n)}, & q/n > \frac{1}{2}; \\ 0, & \text{otherwise.} \end{cases}$$

Note that $\alpha(2q,q) = 0$. Since

$$\left(\binom{n}{2} - i\right)^2 \ge \left(\binom{n-1}{2} - i\right)\left(\binom{n+1}{2} - i\right),$$

we have

$$\binom{\binom{n}{2}}{q}^2 \ge \binom{\binom{n-1}{2}}{q}\binom{\binom{n+1}{2}}{q}.$$

Also, if $g(z) = z\varphi(1/z)$ then $g''(z) = \varphi''(1/z)/z^3$, which is negative for $z < 2$. Hence,

$$e^{2n\varphi(q/n)} \ge e^{(n-1)\varphi(q/(n-1))} e^{(n+1)\varphi(q/(n+1))}$$

for $2q > n+1$. If $\alpha(n,q) = 0$, then $\alpha(n+1,q) = 0$ and so we have

$$\alpha(n,q)^2 \ge \alpha(n-1,q)\alpha(n+1,q) \tag{6.3}$$

for all n.

When $t < -q^{2/3}$, the ratio $q/(n_0+t)$ is strictly larger than $q/n_0 = \log 2 > \frac{1}{2}$, and so $a(q/(n_0+t))$ is bounded. By Theorem 1 we have

$$\sum_{t<-q^{2/3}} d(n_0+t,q) = O(1) \sum_{t<-q^{2/3}} \alpha(n_0+t,q). \tag{6.4}$$

Define $n_1 = \lfloor n_0 \rfloor$, $n_2 = \lfloor n_0 - q^{2/3} \rfloor$. Using Lemma 6.1 we find that for sufficiently large q and some $c > 0$,

$$\frac{\alpha(n_2,q)}{\alpha(n_1,q)} \leq e^{-cq^{1/3}}.$$

By (6.3), this implies that for $i \geq 0$,

$$\alpha(n_2 - i, q) \leq \alpha(n_2,q)\exp(-cq^{1/3}i/(n_1-n_2))$$

and so, summing a geometric series,

$$\sum_{i\geq 0} \alpha(n_2+i,q) \;=\; O(q^{1/3})\alpha(n_2,q) \;=\; O(q^{1/3}e^{-cq^{1/3}})\alpha(n_1,q) \;=\; O(q^{1/3}e^{-cq^{1/3}})d(n_1,q).$$

Since $d(n_1,q)$ is a term of the sum (6.2), we see that the terms for $t < -q^{2/3}$ contribute to $d(q)$ less than the error terms of (6.2).

For the upper tail (the sum over $t > q^{2/3}$), almost the same argument applies. The range of the sum is $2q - n_0 \geq t > q^{2/3}$. Since $\alpha(2q,q)$ had been defined to be 0, it is necessary (and easy) to account for $d(2q,q)$ separately. In the remaining range, namely $2q-n_0-1 \geq t > q^{2/3}$, we have

$$\frac{q}{n_0+t} \geq \frac{q}{2q-1} \geq \frac{1}{2} + \frac{1}{4q}.$$

Noting that $a(x)$ is decreasing, that

$$2x(1-y) = \frac{-\ln(1-y)}{y}(1-y) = 1 - \left(\frac{y}{2} + \frac{y^2}{6} + \cdots\right),$$

and that $y \geq c_1(2x-1)$ for some $c_1 > 0$, we find that for $x = q/(n_0+t)$ and $2q - n_0 - 1 \geq t > q^{2/3}$, we have

$$e^{a(x)} = \sqrt{\frac{1-y}{1-2x(1-y)}}\, e^{x+x^2(1-y^2)} > \frac{c_2}{y^{1/2}} = O(q^{1/2})$$

for some $c_2 > 0$. Replacing the $O(1)$ term in (6.4) by $O(q^{1/2})$, we may now follow the same argument used for $t < q^{-2/3}$.

Theorem 6 follows from Equation (6.1) and the tail bounds established above. □

7 Some unexplored trails.

The above results leave a lot of unanswered questions. Here are a few in what may be an increasing order of difficulty.

1. The probability that a random graph with n vertices, q edges has exactly t isolated vertices is
$$\binom{n}{t} d(n-t,q) \Big/ \binom{N}{q}.$$
When $q = \frac{1}{2} n \log n + \mu n$, routine analysis shows that this is asymptotically a Poisson distribution with parameter $e^{-2\mu}$, as noted by Erdős and Rényi [3]. One might explore the entire range from this Poisson to the normal distribution that occurs when q is small.

2. Let G be a connected labeled graph with s vertices and t edges, and define $X = X(G,n,q)$ to be the expected number of components isomorphic to G in a random labeled graph with n vertices, q edges, and no isolated vertices. Then,
$$E(X) = \frac{(n)_s}{|\operatorname{Aut}(G)|} d(n-s, q-t)/d(n,q).$$
In this equation $\operatorname{Aut}(G)$ is the graph G's automorphism group. Using Theorem 1, we can estimate $E(X)$ uniformly. Similarly, any moment of the distribution of X can be estimated. In many cases, this would allow us to infer a Poisson or normal asymptotic distribution for X. A more challenging project would be to consider deeper questions such as the point at which a large component appears when q increases.

3. For what range of q is it true that almost all (n,q) graphs without isolated vertices have only tree and unicyclic connected components ? Preliminary calculations indicate that the boundary is near the point $x = \frac{1}{2} e/(e-1)$.

4. Having no isolated vertices is the same as requiring that the minimum degree be at least 1. Can one obtain similar results when the minimum degree is 2? (Requiring the minimum degree to be at least $t > 2$ may bring in new difficulties.)

5. It might be possible to prove the stronger relative error estimate $O(1/q)$ mentioned in the Remark after Theorem 1 by applying our method to
$$\binom{N}{q} \exp\{n\varphi(x) + a(x) + \frac{\beta_1(x)}{n}\}.$$
Presumably formal expansion yields a differential equation for $\beta_1(x)$. Theorem 3 probably suffices for large x, but Theorem 2 may need to be extended to a larger range of k with an explicit term of the form ck^3/n^2 in the exponential. On the other hand perhaps there is a different and better method awaiting discovery.

6. Can our results be generalized to the case in which each component has at least t vertices? If so, do the functions corresponding to $\varphi(x)$ and $a(x)$ converge to those found in [1] for connected graphs as $t \to \infty$? It seems likely that $\varphi(x)$ will converge but it may be too much to ask the same of $a(x)$.

References

[1] Bender, E. A., E. R. Canfield, and B. D. McKay, The asymptotic number of labeled connected graphs with a given number of vertices and edges, *Random Structures and Algorithms* 1 (1990) 127–169.

[2] Bollobás, B. *Graph Theory. An Introductory Course*, Springer-Verlag (1979).

[3] Erdős, P., and A. Rényi, On random graphs I, *J. Publ. Math. Debrecen* 6 (1959) 290–297.

[4] McKay, B. D., Subgraphs of random graphs with specified degrees, *Congressus Numerantium* 33 (1981) 213–223.

[5] Stein, C. M., Asymptotic evaluation of the number of latin rectangles, *J. Combin. Theory, Series A* 25 (1978) 38–49.

[6] Williamson, S. G., *Combinatorics for Computer Science*, Computer Science Press (1985).