Decomposing random regular graphs into stars

Michelle Delcourt * Catherine Greenhill† Mikhail Isaev‡
Bernard Lidický§ Luke Postle ¶

August 30, 2023

Abstract

We study k-star decompositions, that is, partitions of the edge set into disjoint stars with k edges, in the uniformly random d-regular graph model $\mathcal{G}_{n,d}$. We prove an existence result for such decompositions for all d, k such that $d/2 < k \leq d/2 + \max\{1, \frac{1}{6}\log d\}$. More generally, we give a sufficient existence condition that can be checked numerically for any given values of d and k. Complementary negative results are obtained using the independence ratio of random regular graphs. Our results establish an existence threshold for k-star decompositions in $\mathcal{G}_{n,d}$ for all $d \leq 100$ and $k > d/2$, and strongly suggest the a.a.s. existence of such decompositions is equivalent to the a.a.s. existence of independent sets of size $(2k-d)n/(2k)$, subject to the necessary divisibility conditions on the number of vertices.

For smaller values of k, the connection between k-star decompositions and β-orientations allows us to apply results of Thomassen (2012) and Lovász, Thomassen, Wu and Zhang (2013). We prove that random d-regular graphs satisfy their assumptions with high probability, thus establishing a.a.s. existence of k-star decompositions (i) when $2k^2 + k \leq d$, and (ii) when k is odd and $k < d/2$.

1 Introduction

A *k-star* is a complete bipartite graph $K_{1,k}$, and a k-star decomposition of a given graph G is a partition of the edge set of G into disjoint k-stars. The problem of decomposing graphs into disjoint stars has been well-studied in the design community: see, for example, [4, 12, 20, 26] and references therein.

*Department of Mathematics, Toronto Metropolitan University, Toronto, Ontario M5B 2K3, Canada `mdelcourt@torontomu.ca` .

†School of Mathematics and Statistics, UNSW Sydney, NSW 2052, Australia. `c.greenhill@unsw.edu.au`

‡School of Mathematics, Monash University, VIC 3800, Australia. `mikhail.isaev@monash.edu` .

§Department of Mathematics, Iowa State University, Ames, IA, USA. `lidicky@iastate.edu` .

¶Combinatorics and Optimization Department, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada `lpostle@uwaterloo.ca` .

We consider the problem of existence of k-star decompositions in the uniformly random d-regular graph $\mathcal{G}_{n,d}$, where k, d are constants and $n \to \infty$. A necessary condition is that k divides $dn/2$, since there are exactly $dn/(2k)$ disjoint k-stars in any such decomposition. Let $\mathcal{N}_{d,k}$ be defined by

$$\mathcal{N}_{d,k} := \{n \in \mathbb{Z}^+ \mid 2k \text{ divides } dn\}.$$

All our asymptotic statements are with respect to $n \to \infty$ along $\mathcal{N}_{d,k}$. If an event holds with probability which tends to 1 then we say that the event holds *asymptotically almost surely* (abbreviated as a.a.s.).

In 2018 the first and last authors [7] proved that $\mathcal{G}_{n,4}$ a.a.s. has a 3-star decomposition; they also asserted that the work of Lovász, Thomassen, Wu and Zhang [15, Theorem 3.1] implies that if $k \le \lceil d/2 \rceil$ then a.a.s. $\mathcal{G}_{n,d}$ admits a k-star decomposition. In this paper, we would like to clarify that the situation is not as straightforward, since [15, Theorem 3.1] assumes that k is odd and we require a stronger assumption than d-edge-connectivity. Thus we must perform a more careful application of [15, Theorem 3.1], together with some known properties of random regular graphs, to prove the following in Section 2.

Theorem 1.1. *Let k, d be positive integers with $k < d$. Then a.a.s. $\mathcal{G}_{n,d}$ has a k-star decomposition in the following cases:*

(i) *$k = 2$ or d is an even multiple of k.*

(ii) *$d \ge 2k^2 + k$.*

(iii) *k is odd and $3 \le k < d/2$.*

Now we discuss the case $k > d/2$. We can view a k-star decomposition of a graph G as an orientation of its edges, by directing all edges of each copy of $K_{1,k}$ towards the centre. When $k > d/2$, each vertex can be the centre of at most one k-star, and so every vertex has in-degree 0 or k in this orientation. A vertex is called a *centre* if it has in-degree k, otherwise it is a *leaf*. There are $\frac{dn}{2k}$ centres and $n - \frac{dn}{2k} = \frac{(2k-d)n}{2k}$ leaves in any decomposition. Observing that the leaves form an independent set, we establish another necessary condition:

$$\begin{aligned}&\text{if a } d\text{-regular graph } G \text{ admits a } k\text{-star decomposition}\\ &\text{then } G \text{ contains an independent set of size } \tfrac{(2k-d)n}{2k}.\end{aligned} \tag{1}$$

For fixed $d \ge 3$ let $k_{\text{ind}}(d)$ be the largest positive integer k such that a.a.s. $\mathcal{G}_{n,d}$ contains an independent set of size $(2k-d)n/(2k)$. We make the following conjecture.

Conjecture 1.2. *For d, k fixed positive integers with $d \ge 3$,*

$$\Pr\left(\mathcal{G}_{n,d} \text{ has a } k\text{-star decomposition}\right) \to \begin{cases} 1 & \text{if } k \le k_{\text{ind}}(d), \\ 0 & \text{if } k > k_{\text{ind}}(d). \end{cases}$$

For $d \geq 3$, let $k^+_{\text{ind}}(d)$ be the largest positive integer $d/2 < k < d$ such that

$$d^{d-1} > (2k-d)^{(2k-d)/d}\, 2^{(d^2-2k)/d}\, k^{k(d-2)/d}\, (d-k)^{d-k}. \tag{2}$$

Note that $k^+_{\text{ind}}(d)$ is well defined as (2) holds when $k = \lceil d/2 \rceil$. In fact, $k^+_{\text{ind}}(d)$ corresponds to the expectation threshold for existence of independent sets of size $(2k-d)n/(2k)$ in $\mathcal{G}_{n,d}$. We will see this in Section 3.1, where we prove the following theorem which establishes the negative side of Conjecture 1.2.

Theorem 1.3. *If $d \geq 3$ then $k_{\text{ind}}(d) \leq k^+_{\text{ind}}(d)$. Furthermore, if $(d,k) = (5,4)$ or $k > k_{\text{ind}}(d)$ then*

$$\Pr\big(\mathcal{G}_{n,d} \textit{ has a } k\textit{-star decomposition}\big) \to 0.$$

Frieze and Łuczak [9] proved that the independence number of $\mathcal{G}_{n,d}$ is a.a.s. bounded above by $\frac{2n\log d}{d}$. Using (1) it follows that

$$k_{\text{ind}}(d) \leq \frac{d^2}{2(d-2\log d)} \approx d/2 + \log d,$$

where the additive error in the final approximation tends to zero as d grows. Our final result, which does not rely on numerical computations, gives an infinite family of existence results. Note that if Conjecture 1.2 is true then $k_{\text{ind}}(d) - d/2$ grows logarithmically with d, and Theorem 1.4 covers one-sixth of the range $(d/2, k_{\text{ind}}(d))$.

Theorem 1.4. *Let $d \geq 4$ and $k \leq d/2 + \max\left\{1, \frac{\log d}{6}\right\}$. Then*

$$\Pr\big(\mathcal{G}_{n,d} \textit{ has a } k\textit{-star} \text{ decomposition}\big) \to 1.$$

We prove Theorem 1.4 using a more general result which we state in the next section.

1.1 Numerically-verifiable sufficient condition for existence

Define the polynomial f and function η on $[0,\infty)$ by

$$f(x) = \frac{1}{\binom{d}{k}} \sum_{i=0}^{d-k} \binom{k}{i}\binom{d-k}{i} x^{k-i}, \tag{3}$$

$$\eta(x) = \frac{2k-d}{d-k+\sqrt{(d-k)^2 + d(2k-d)\,f(x)}}. \tag{4}$$

Note that $f(x)$ is the PGF of the hypergeometric distribution with parameters (d,k,k). For $d \geq 3$ we define $k_{\text{exist}}(d)$ to be the largest integer such that for k satisfying $d/2 < k < k_{\text{exist}}(d)$ we have that

$$(2k-d)^2 < 4k-d-2 \tag{5}$$

and

$$(1+\eta(x))f(x) = \frac{(x+1)\,f'(x)}{k} \quad \text{has unique solution } x=1$$
$$\text{on} \quad \left(\left(1 + \frac{(2k-d)^2 d}{k(d-k)(4k-d-2-(2k-d)^2)}\right)^{-1}, \frac{5k-2d}{d-k} \right). \tag{6}$$

The condition (6) can be checked numerically for any specific pair (d,k); see Subsection 1.2 for an example. Direct computations show that $k = \lceil \frac{d+1}{2} \rceil$ satisfies (5), and the estimates of Section 6.4 show that (6) also holds for this value of k. On the other hand, (5) fails for all $k \geq d$ which shows that $k_{\text{exist}}(d)$ is well-defined.

The most technical part of the paper is the proof of the following theorem using the small subgraph conditioning method.

Theorem 1.5. *If $d \geq 3$ and $d/2 < k \leq k_{\text{exist}}(d)$ then*

$$\Pr\big(\mathcal{G}_{n,d} \textit{ has a } k\textit{-star decomposition}\big) \to 1.$$

As an immediate consequence of (1), Theorem 1.3 and Theorem 1.5 we obtain the following.

Corollary 1.6. *If $k_{\text{exist}}(d) = k^+_{\text{ind}}(d)$ then $k_{\text{exist}}(d) = k_{\text{ind}}(d) = k^+_{\text{ind}}(d)$.*

We remark that Theorem 1.5 implies an a.a.s. lower bound of $\frac{2k_{\text{exist}}(d)-d}{2\,k_{\text{exist}}(d)}$ on the independence ratio of $\mathcal{G}_{n,d}$, using (1). For example, Corollary 1.6 and data from Table 1, presented in the next subsection, show that $k_{\text{ind}}(16) = 10$. This leads to a lower bound of $1/5$ on the independence ratio of $\mathcal{G}_{n,16}$. To our surprise, this improves on the best-known explicit lower bound of 0.1985 given by Wormald [24, Table 1], obtained using the differential equations method. More recent results can be found in [8, 16], but these papers do not report the resulting lower bound on the independence ratio when $d = 16$. Ding, Sly and Sun [6] proved that the independence ratio of $\mathcal{G}_{n,d}$ is highly concentrated around a constant $\alpha^*(d)$ when $d \geq d_0$, where d_0 is a sufficiently large constant. However the constant $\alpha^*(d)$ seems hard to compute, and the value of d_0 is not stated explicitly, so it is not clear whether result applies to $d = 16$. Harangi [11] used replica symmetry breaking formulas to obtain improved upper bounds for the indepence ratio when $d = 3, \ldots, 19$.

1.2 Small values of d and k

In Table 1 we record the values of $k_{\text{exist}}(d)$, $k_{\text{ind}}(d)$ (where known) and $k^+_{\text{ind}}(d)$ for $d = 3, 4, \ldots, 20$. The values for $k_{\text{exist}}(d)$ and $k^+_{\text{ind}}(d)$ were computed from their definition, and then Corollary 1.6 determined the values of $k_{\text{ind}}(d)$ for all d other than 5, 10, 12, 14. The value of $k_{\text{ind}}(10)$ was computed from (1) using the a.a.s. lower and upper bounds on the independence ratio for $\mathcal{G}_{n,10}$ given by Wormald [24, Theorem 5] and McKay [17], respectively.

d	3	4	5	6	7	8	9	10	11	**12**	13	**14**	15	16	17	18	19	20
$k_{\text{exist}}(d)$	2	3	3	4	5	5	6	6	7	7	8	8	9	10	10	11	11	12
$k_{\text{ind}}(d)$	2	3	?	4	5	5	6	6	7	?	8	?	9	10	10	11	11	12
$k^+_{\text{ind}}(d)$	2	3	4	4	5	5	6	7	7	8	8	9	9	10	10	11	11	12

Table 1: The values of $k_{\text{exist}}(d)$, $k_{\text{ind}}(d)$ (where known) and $k^+_{\text{ind}}(d)$ for $d = 4, 5, \ldots, 20$.

From Table 1 together with Theorems 1.3 and 1.5, we see that Conjecture 1.2 holds for all $k \geq d/2$ and $d \in \{4, \ldots, 20\}$ except possibly for the pairs $(d, k) \in \{(5, 4),\ (12, 8),\ (14, 9)\}$. We have checked numerically that for $d = 3, \ldots, 100$,

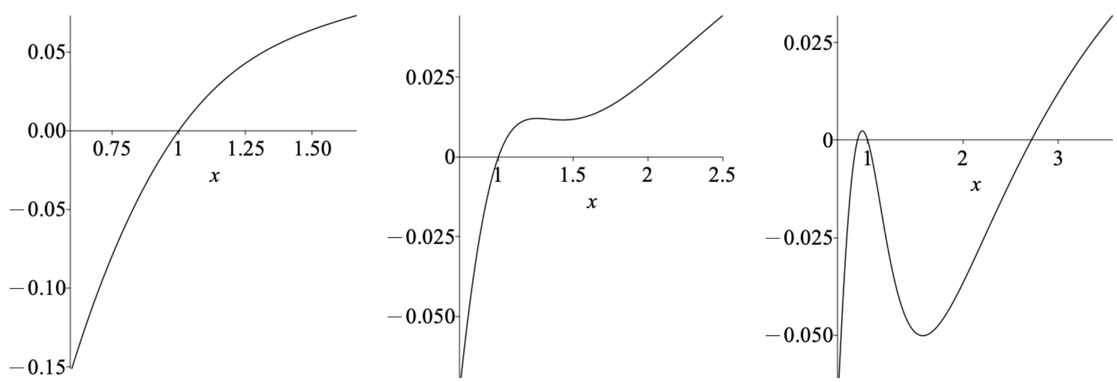
$$k_{\text{exist}}(d) \in \{k^+_{\text{ind}}(d) - 1,\ k^+_{\text{ind}}(d)\}.$$

Thus we observe an existence threshold (possibly non-sharp) for the property of having a k-star decomposition for these values of d.

Before proceeding, we give an example showing how to compute Figure 1 contains plots of the function

$$\widehat{f}(x) := \frac{k(1 + \eta(x))\, f(x)}{(x + 1)\, f'(x)} - 1 \tag{7}$$

on the range specified in (6) for $d = 20$ and $k = 11, 12, 13$. $k_{\text{exist}}(d)$ when $d = 20$.



$d = 20,\ k = 11$ $\quad$ $d = 20,\ k = 12$ $\quad$ $d = 20,\ k = 13$

Figure 1: Plots of $\widehat{f}(x)$, defined in (7), for $d = 20$ and $k = 11, 12, 13$

Since the function $\widehat{f}$ is zero if and only if (6) holds, and since (5) holds for all these values of (d, k), these plots prove that $k_{\text{exist}}(20) = 12$. Direct computations show that $k^+_{\text{ind}}(20) = 12$, therefore Conjecture 1.2 holds when $d = 20$ and $k \geq 10$.

Next, in Table 2 we summarise what is known about the a.a.s. existence or non-existence of k-star decompositions of $\mathcal{G}_{n,d}$ when $3 \leq d \leq 20$ and $2 \leq k < d$. The highlighted cell in each row is $(d, k_{\text{exist}}(d))$. Every cell strictly to the right of $(d, k^+_{\text{ind}}(d))$ is zero: columns for $k = 17, 18, 19$ are not shown but contain only zero entries.

d \ k	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
9	1	1	?	1	1	0	0	0	0	0	0	0	0	0	0
10	1	1	?	1	1	0	0	0	0	0	0	0	0	0	0
11	1	1	?	1	1	1	0	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	?	0	0	0	0	0	0	0	0
13	1	1	1	1	?	1	1	0	0	0	0	0	0	0	0
14	1	1	1	1	?	1	1	?	0	0	0	0	0	0	0
15	1	1	1	1	?	1	1	1	0	0	0	0	0	0	0
16	1	1	1	1	?	1	1	1	1	0	0	0	0	0	0
17	1	1	?	1	?	1	?	1	1	0	0	0	0	0	0
18	1	1	1	1	1	1	?	1	1	1	0	0	0	0	0
19	1	1	1	1	1	1	?	1	1	1	0	0	0	0	0
20	1	1	1	1	1	1	?	1	1	1	1	0	0	0	0

Table 2: Limit of $\Pr(\mathcal{G}_{n,d}$ has a k-star decomposition$)$ for $3 \le d \le 20$, where known.

Let $(\mathcal{A}_n)$ and $(\mathcal{B}_n)$ be sequences of probability spaces on the same sequence of finite underlying sets (Ω_n). We say that $(\mathcal{A}_n)$ and $\mathcal{B}_n)$ are *contiguous* if for any sequence of events $(\mathcal{E}_n)$ we have

$$\lim_{n\to\infty} \Pr_{\mathcal{A}_n}(\mathcal{E}_n) = 1 \quad \text{if and only if} \quad \lim_{n\to\infty} \Pr_{\mathcal{B}_n}(\mathcal{E}_n) = 1.$$

The *superposition* (also called *graph-restricted sum* in [25]) of two random graph models G_n and H_n on the vertex set $[n]$, denoted $G_n \oplus H_n$, is the probability space obtained by randomly choosing G from G_n and H from H_n repeatedly until $E(G) \cap E(H) = \emptyset$, then returning $([n], E(G) \cup E(H))$. Using contiguity arithmetic [25, Corollary 4.17] we know that if $d \ge 3$ and $a_1, \ldots, a_s$ are positive integers such that $d = \sum_{i=1}^{s} a_i$ then

$$\mathcal{G}_{n,d} \approx \mathcal{G}_{n,a_1} \oplus \mathcal{G}_{n,a_2} \oplus \cdots \oplus \mathcal{G}_{n,a_s}. \tag{8}$$

If two graphs with disjoint edge sets have k-star decompositions, then their union also has a k-star decomposition. We used this fact, combined with (8), to fill in some entries in Table 2 with $k \leq d/2$ which were not covered by Theorem 1.1. For example, $\mathcal{G}_{n,6}$ a.a.s. has a 4-star decomposition, and $\mathcal{G}_{n,12} \approx \mathcal{G}_{n,6} \oplus \mathcal{G}_{n,6}$, which implies that $\mathcal{G}_{n,12}$ a.a.s. has a 4-star decomposition. Similarly $\mathcal{G}_{n,20}$ a.a.s. has a 4-star decomposition since $\mathcal{G}_{n,20} \approx \mathcal{G}_{n,7} \oplus \mathcal{G}_{n,7} \oplus \mathcal{G}_{n,6}$.

We believe that the a.a.s. existence for remaining unknown pairs with $k \leq d/2$ can be resolved by extending the result of [15] to even values of k. In particular, Lovász et al. [15, Section 5.2] claim (without proof) that their approach can be extended to $(3k-2)$-edge-connected graphs for even k, which would confirm our conjecture for the cases

$$(d,k) \in \{(10,4)\,(11,4),\ (17,4),\ (16,6),\ (17,6)\}.$$

We believe that the $(d,k) = (12,8)$ cell in Table 2 should be 0. To prove this, it would suffice to show that the independence ratio of $\mathcal{G}_{n,12}$ is a.a.s. strictly less than 1/4. On the other hand, we believe that the $(d,k) = (14,9)$ cell in Table 2 should be 1. Condition (6) ensures that a certain function has a unique local maximum at $x = 1$. In fact, our approach would still work if this other local maxima exist but have smaller values than the value at $x = 1$. The function is difficult to analyse, which is why we restrict our attention to the simpler situation when there is only one local maximum. This more complicated analysis will work for $(d,k) = (14,9)$ but we do not include those details here. If, as claimed, $\mathcal{G}_{n,14}$ a.a.s. has a 9-star decomposition then by (1) we conclude that the independence ratio of $\mathcal{G}_{n,14}$ is at least $2/9 = 0.222\ldots$, which would improve on the best-known lower bound 0.2143 proved by Wormald [24, Theorem 5].

Finally we remark that Conjecture 1.2 for the pair $(d,k) = (5,4)$ is equivalent to the statement that the independence ratio of $\mathcal{G}_{n,5}$ is strictly less than $3/8 = 0.375$, using (1). However, the best-known upper bound [11] on the independence ratio of $\mathcal{G}_{n,5}$ is 0.379 (to 3 significant figures), which is not sufficient to verify our conjecture in this case.

1.3 Outline of the paper

In Section 2 we prove Theorem 1.1. Part (i) follows easily from known results about 2-star decompositions and Eulerian orientations, while parts (ii) and (iii) are proved using results of Lovász, Thomassen, Wu and Zhang [15] and Thomassen [21], respectively.

In Section 3 we prove Theorem 1.3 and state the small subgraph conditioning theorem from Janson [13, Theorem 1]. The rest of the paper is devoted to verifying the assumptions of this theorem applied to our problem. In particular, Theorems 1.4 and Theorem 1.5 are proved in Section 5.2, utilising calculations performed in Sections 4 and 5. The most technical part of the second moment calculations are postponed until Section 6.

2 Small values of k

In this section we prove Theorem 1.1. We will repeatedly use the following result from Wormald [23]:

$$\text{If } d \geq 3 \text{ is fixed then } \mathcal{G}_{n,d} \text{ is a.a.s. } d\text{-connected.} \tag{9}$$

It is well known that a connected graph with an even number of edges has a 2-star decomposition; see for example [14] or [5, Theorem 1]. Next, let G be a $2rk$-regular graph G with $r \in \mathbb{Z}^+$. Then G has an Eulerian orientation, where each vertex v has exactly rk in-edges. These in-edges can be partitioned to give exactly r distinct k-stars centred at v. These observations imply Theorem 1.1(i) since $\mathcal{G}_{n,d}$

Theorem 1.1(ii) follows immediately from Thomassen [21, Theorem 5] using (9).

To prove Theorem 1.1(iii) we use machinery from Lovász et al. [15, Definition 1.9]. Denote the set of integers modulo k by $\mathbb{Z}_k$. A function $\beta : V(G) \to \mathbb{Z}_k$ is called a $\mathbb{Z}_k$-*boundary* of G if $\sum_{v\in V(G)} \beta(v) \equiv 0 \pmod{k}$. Given a $\mathbb{Z}_k$-boundary of G, an orientation D of G is called a β-*orientation* if for every vertex $v \in V(G)$,

$$d_D^+(v) - d_D^-(v) \equiv \beta(v) \pmod{k}.$$

The following lemma establishes a connection with k-star decompositions.

Lemma 2.1. *Let G be a d-regular graph with n vertices such that $2k$ divides dn, where $k \geq 3$ is an odd positive integer. Then G has a k-star decomposition if and only if G has a β-orientation with $\beta(v) \equiv d \pmod{k}$ for all $v \in V(G)$.*

Proof. Since k is odd and $d_D^+(v) + d_D^-(v) = d$ for all $v \in V(G)$, the condition $d_D^+(v) - d_D^-(v) \equiv d \pmod{k}$ is equivalent to saying that k divides the in-degree of v. When this condition holds for $v \in V(G)$ we may partition incoming edges at v into k-stars centred at v, and vice-versa. □

Combining (9) and the above lemma with the result of Lovász et al. stated below, we establish Theorem 1.1(iii) when $k \leq (d + 1)/3$ and k odd.

Theorem 2.2 ([15, Theorem 1.12])**.** *Let $k \geq 3$ be an odd integer. Every $(3k - 3)$-edge-connected graph G has a β-orientation for every $\mathbb{Z}_k$-boundary β of G.*

To cover the remaining range, that is $(d+1)/3 < k < d/2$, we need a more technical result from [15]. In this range, we let $\beta(v) = d - 2k \in \{0, 1, \ldots, k-1\}$ for all $v \in V(G)$. Lovász et al. [15] use a function $\tau : V(G) \to \{0, \pm 1, \ldots, \pm k\}$ such that for each vertex $v \in V(G)$,

$$\tau(v) \equiv \begin{cases} \beta(v) \pmod{k}, \\ d \pmod{2}. \end{cases}$$

In our range we have $\tau(v) = \beta(v) = d - 2k$ for all $v \in V$. They extend τ to a function over subsets of $V(G)$ such that $|\tau(A)| \leq k$ for all $A \subseteq V(G)$.

We will state a special case of their theorem which suffices for our purposes, obtained by applying [15, Theorem 3.1] to the disjoint union of a d-regular graph and an isolated vertex z_0, using the fact that $\tau(v)$ is always nonzero and the observations above. Let $e(A, \bar{A})$ denote the number of edges from A to $\bar{A} = [n] \setminus A$, for all $A \subseteq [n]$.

Theorem 2.3 ([15, Theorem 3.1])**.** *Let G be a d-regular graph with $d \geq 3$, and let k be an odd integer such that $(d+1)/3 < k < d/2$. If $e(A, \bar{A}) \geq 3k - 2$ for each vertex subset A with $|A| \geq 2$ such that $|V(G) \setminus A| \geq 2$ then there exists a β-orientation D of G, where β is the $\mathbb{Z}_k$-boundary of G defined by $\beta(v) = d - 2k$ for all $v \in V(G)$.*

A graph is said to be *internally r-connected* if for every cut with at least two vertices on either side, the number of edges across the cut is at least r. The proof of Theorem 1.1(iii) is completed by combining Theorem 2.3 and the following lemma.

Lemma 2.4. *Let $d \geq 4$ be an integer. Then $\mathcal{G}_{n,d}$ is a.a.s. internally $2(d-1)$-edge connected.*

Proof. Let $A \subseteq [n]$ be a subset of vertices such that $2 \leq |A| \leq n/2$. If $|A| = 2$ then $e(A, \bar{A}) \geq 2d - 2$, with equality if the two vertices in A are adjacent.

More generally, a.a.s. for any vertex subset A with $3 \leq |A| \leq 11(d-1)$, the induced subgraph $\mathcal{G}_{n,d}[A]$ contains at most $|A|$ edges, for example using Wormald [25, Lemma 2.7]. Hence a.a.s., every A with $3 \leq |A| \leq 11(d-1)$ satisfies

$$e(A, \bar{A}) \geq |A|(d-2) \geq 2(d-1),$$

using the fact that $d \geq 4$.

Now suppose that $|A| > 11(d-1)$. It follows from Bollobás [3, Theorem 1 and Corollary 2] that in $\mathcal{G}_{n,d}$, the following a.a.s. holds: for all A with $|A| > 11(d-1)$ we have

$$e(A, \bar{A}) \geq \frac{2}{11}|A| > 2(d-1)$$

as required. □

3 Configuration model

Throughout the paper, we use log to denote the natural logarithm. By convention, $0^0 = 1$ and $0 \log 0 = 0$.

Our calculations will be performed in the *configuration model* (or *pairing model*) for d-regular graphs on n vertices, denoted by $\Omega_{n,d}$. In the configuration model there are n cells, each containing d points. A *pairing* is a partition of the dn points into $dn/2$ unordered pairs. Replacing cells by vertices and pairs by edges, each pairing P corresponds to a multigraph $G(P)$ which may have loops or multiple edges. The number of pairings on $2a$ points is denoted by

$$M(2a) = \frac{(2a)!}{a!\, 2^a}. \tag{10}$$

Bender and Canfield [1] proved that the probability that a randomly chosen configuration from $\Omega_{n,d}$ is simple is

$$\Pr(\text{Simple}) \sim \exp\big(-(d^2-1)/4\big). \tag{11}$$

3.1 Proof of Theorem 1.3

Our proof of Theorem 1.3 relies on the first moment approach. Similar calculations can be found in [6, 9]. Let Z_α denote the number of independent sets of size αn in the configuration model $\Omega_{n,d}$, where $\alpha \in (0,1)$ is fixed. Then

$$\mathbf{E}Z_\alpha = \binom{n}{\alpha n} (dn - d\alpha n)_{d\alpha n} \frac{M(dn - 2d\alpha n)}{M(dn)}.$$

Applying (10) and Stirling's approximation gives

$$\begin{aligned}\mathbf{E}Z_\alpha &= \frac{n!\,(dn-d\alpha n)!\,(dn/2)!\,2^{d\alpha n}}{(\alpha n)!\,\big((1-\alpha)n\big)!\,(dn/2-d\alpha n)!\,(dn!)} \\ &\sim \sqrt{\frac{1}{2\pi\alpha(1-2\alpha)\,n}}\left(\frac{(1-\alpha)^{(d-1)(1-\alpha)}}{\alpha^\alpha\,(1-2\alpha)^{d(1-2\alpha)/2}}\right)^n. \end{aligned} \tag{12}$$

Finally we substitute $\alpha = (2k-d)/(2k)$, writing Z instead of $Z_{(2k-d)/(2k)}$ for ease of notation, giving

$$\mathbf{E}Z \sim \frac{k}{\sqrt{\pi(2k-d)\,(d-k)\,n}}\left(\frac{d^{d-1}}{(2k-d)^{(2k-d)/d}\,(d-k)^{d-k}\,2^{(d^2-2k)/d}\,k^{k(d-2)/d}}\right)^{dn/(2k)}.$$

If $k > k^+_{\text{ind}}(d)$ then the base of the exponential factor is bounded above by 1, and hence $\mathbf{E}Z \to 0$. Using (11), we find that the expected number of independent sets in $\mathcal{G}_{n,d}$ of size $n - dn/(2k)$ also tends to zero, by (14). By Markov's inequality this implies that a.a.s. $\mathcal{G}_{n,d}$ has no independent set of size $(2k-d)n/(2k)$, so $k > k_{\text{ind}}(d)$. This proves the first statement. The second statement follows from (1) when $k > k_{\text{ind}}(d)$, and from Lemma 4.1 when $(d,k) = (5,4)$.

3.2 Small subgraph conditioning method

To prove Theorem 1.5 and Theorem 1.4 we will apply the small subgraph conditioning method, introduced by Robinson and Wormald [19]. This statement is taken from [13].

Theorem 3.1 (Janson [13, Theorem 1])**.** *Let $\lambda_j > 0$ and $\delta_j \geq -1$, $j = 1, 2, \ldots,$ be constants and suppose that for each n there are random variables $X_{j,n}$, $j = 1, 2, \ldots,$ and Y_n (defined on the same probability space) such that $X_{j,n}$ is a nonnegative integer valued and $\mathbf{E}Y_n \neq 0$ (at least for large n), and furthermore the following conditions are satisfied:*

(A1) *$X_{j,n} \xrightarrow{d} Z_j$ as $n \to \infty$ jointly for all j, where $Z_j \sim \text{Po}(\lambda_j)$ are independent Poisson random variables;*

(A2) *For any finite sequence* $x_1, \ldots, x_m$ *of nonnegative integers,*

$$\frac{\mathbf{E}(Y_n | X_{1,n} = x_1, \ldots, X_{m,n} = x_m)}{\mathbf{E}Y_n} \to \prod_{j=1}^{m} (1+\delta_j)^{x_j} e^{-\lambda_j \delta_j} \qquad \text{as } n \to \infty;$$

(A3) $\sum_{j \geq 1} \lambda_j \delta_j^2 < \infty$*;*

(A4) $\frac{\mathbf{E}Y_n^2}{(\mathbf{E}Y_n)^2} \to \exp\left(\sum_{j \geq 1} \lambda_j \delta_j^2\right) \qquad$ *as* $n \to \infty$*.*

Then

$$\frac{Y_n}{\mathbf{E}Y_n} \xrightarrow{d} W = \prod_{j=1}^{\infty} (1+\delta_j)^{Z_j} e^{-\lambda_j \delta_j} \qquad \text{as } n \to \infty;$$

moreover, this and the convergence in (A1) *hold jointly. In particular,* $W > 0$ *almost surely if and only if every* $\delta_j > -1$*.*

Janson remarks in [13] that the index set $\mathbb{Z}^+$ may be replaced by any other countably infinite set, and that 0^0 is defined to be 1. We will apply Theorem 3.1 to Y, and letting X_j be the number of j-cycles in a random pairing. Bollobás [2] proved that $X_j \to Z_j$ as $n \to \infty$, where Z_j are asymptotically independent Poisson random variables with mean

$$\lambda_j = \frac{(d-1)^j}{2j}. \tag{13}$$

In order to verify condition (A2), the following lemma is helpful.

Lemma 3.2 (Janson [13, Lemma 1])**.** *Let* $\lambda'_j \geq 0$*,* $j = 1, 2, \ldots$ *be constants. Suppose that* (A1) *holds, that* $Y_n \geq 0$ *and that*

$$\text{(A2}'\text{)} \qquad \frac{\mathbf{E}(Y_n (X_{1,n})_{x_1} \cdots (X_{m,n})_{x_m})}{\mathbf{E}Y_n} \to \prod_{j=1}^{m} (\lambda'_j)^{x_j} \qquad \text{as } n \to \infty,$$

for every finite sequence $x_1, \ldots, x_m$ *of nonnegative integers. Then condition* (A2) *holds with* $\lambda'_j = \lambda_j(1+\delta_j)$*.*

Let $Y = Y_{d,k}$ be the number of ways to orient all pairs in a uniformly random pairing from $\Omega_{n,d}$, such that every cell has in-degree k or 0. This corresponds to an edge-disjoint decomposition of the (multigraph corresponding to the) pairing into copies of k-stars. For $j \geq 1$ let X_j be the number of j-cycles in a uniformly random pairing from $\Omega_{n,d}$. The remainder of the paper is devoted to proving Theorem 1.5 and Theorem 1.4, using the small subgraph conditioning method applied to the random variables Y, X_j.

The first moment calculations are given at the start of Section 4, and the effect of short cycles is investigated in Section 4.1. The second moment calculations are presented in Section 5. These are the most technical, and rely on a certain function having a unique maximum in the interior of a given domain: this fact is proved in Section 6.

Successful application of the small subgraph conditioning method will allow us to conclude that a.a.s. $Y > 0$. Now let $Y_{\mathcal{G}} = Y_{\mathcal{G}}(d,k)$ be the number of k-star decompositions of $\mathcal{G}_{n,d}$. Since each element of $\mathcal{G}_{n,d}$ corresponds to the same number of pairings, it follows using (11) that

$$\Pr(Y_{\mathcal{G}} = 0) = \Pr(Y = 0 \mid \text{Simple}) \le \frac{\Pr(Y=0)}{\Pr(\text{Simple})} = O\big(\Pr(Y=0)\big). \tag{14}$$

In particular, if $\Pr(Y = 0) \to 0$ then a.a.s. $Y_{\mathcal{G}} > 0$.

4 First moment and the effect of short cycles

With notation as above,

$$\mathbf{E}Y = \binom{n}{\frac{dn}{2k}} \binom{d}{k}^{dn/(2k)} (dn/2)!\, \frac{1}{M(dn)}. \tag{15}$$

Here the first factor chooses the centres, the next factor identifies the in-points in each centre, the third factor fixes a pairing which pairs each in-point with an out-point, and the final factor is the probability that we observe this particular pairing.

Applying Stirling's formula, we find that

$$\mathbf{E}Y \sim \frac{k}{\sqrt{2k-d}} \left(\frac{\binom{d}{k}\, k^{2k/d}}{2^{k(d-2)/d}\, d\, (2k-d)^{(2k-d)/d}} \right)^{dn/(2k)}. \tag{16}$$

Before concluding this section we prove the following result, which will be needed in Section 5.

Lemma 4.1. *Suppose that $d \ge 4$ and $d/2 < k \le k^+_{\text{ind}}(d)$. If $(d,k) \ne (5,4)$ then*

$$\frac{\binom{d}{k}\, k^{2k/d}}{2^{k(d-2)/d}\, d\, (2k-d)^{(2k-d)/d}} > 1 \tag{17}$$

and hence $\mathbf{E}Y \to \infty$. If $(d,k) = (5,4)$ then $\mathbf{E}Y \to 0$.

Proof. Recall the asymptotic formula for $\mathbf{E}Z_\alpha$ from (12), where Z_α is the number of independent sets of size αn in $\Omega_{n,d}$. For $0 \le \alpha < 1/2$, let

$$h_d(\alpha) = (d-1)(1-\alpha)\log(1-\alpha) - \alpha\log\alpha - \frac{d}{2}(1-2\alpha)\log(1-2\alpha)$$

be the logarithm of the exponential part of the asymptotic formula for $\mathbf{E}Z_\alpha$. As noted by Ding, Sly and Sun [6, Lemma 2.1], $h_d(\alpha)$ is concave on $[0, 1/2)$, and $h'_d(\alpha) > 0$ for small values of α. Since $h_d(0) = 0$ and $h_d(1/2) < 0$, this implies that there is a unique value of $\alpha \in (0, 1/2)$ such that $h_d(\alpha) = 0$. The inequality (2) holds if and only if $h_d(\alpha^*) > 0$ holds for $\alpha^* = (2k-d)/(2k)$. Since $(2k-d)/(2k) = 1 - d/(2k)$ is an increasing function of k, it follows that (2) holds for all $d/2 < k \le k^+_{\mathrm{ind}}(d)$.

Next, using (2), to prove that (17) it is sufficient to prove that

$$1 < \frac{\binom{d}{k}\, k^{2k/d}}{2^{k(d-2)/d}\, d} \cdot \frac{2^{(d^2-2k)/d}\, k^{k(d-2)/d}\, (d-k)^{d-k}}{d^{d-1}} = \frac{\binom{d}{k}\, k^k\, (d-k)^{d-k}\, 2^{d-k}}{d^d}.$$

Rewrite this inequality as

$$\frac{\binom{d}{k}\, k^k\, (d-k)^{d-k}}{2^k} > \left(\frac{d}{2}\right)^d. \tag{18}$$

We start by dealing with the cases $d \in \{4, 5, 6, 7, 8, 9\}$ computationally. Let $c(d, k)$ denote the left hand side of (17). Recall that $k^+_{\mathrm{ind}}(4) = 3$, $k^+_{\mathrm{ind}}(5) = 4$, $k^+_{\mathrm{ind}}(6) = 4$, $k^+_{\mathrm{ind}}(7) = 5$, $k^+_{\mathrm{ind}}(8) = 5$, $k^+_{\mathrm{ind}}(9) = 6$. Computing $c(d, k)$ for relevant values to 3 decimal figures gives:

$$\begin{array}{lllll} c(4,3) = 1.299, & c(5,3) = 2.146, & c(5,4) = 0.901, & c(6,4) = 1.984, & \\ c(7,4) = 3.365, & c(7,5) = 1.571, & c(8,5) = 3.271, & c(9,5) = 5.651, & c(9,6) = 2.778. \end{array}$$

This shows that (17) holds when $d \in \{4, 5, 6, 7, 8, 9\}$ and $d/2 < k \le k^+_{\mathrm{ind}}(d)$, except for the special case $(d, k) = (5, 4)$ where $c(5, 4) < 1$ and hence $\mathbf{E}Y \to 0$.

For the remainder of the proof, suppose that $d \ge 10$. We claim that the left hand side of (18) is monotonically decreasing for $k \le d - 2$. This is true if

$$\frac{\binom{d}{k+1}\, (k+1)^{k+1}\, (d-k-1)^{d-k-1}}{\binom{d}{k}\, k^k\, (d-k)^{d-k}} = \frac{(1 + 1/k)^k}{\big(1 + 1/(d-k-1)\big)^{d-k-1}} < 2,$$

and the above inequality holds as

$$\left(1 + \frac{1}{k}\right)^k \le e < 4 \le 2\left(1 + \frac{1}{d-k-1}\right)^{d-k-1}$$

when $d/2 < k \le d - 2$.

Next, when $d \ge 10$ we have $36\big(1 - 3/d\big)^d > 1$ and hence

$$27\binom{d}{3}(d-3)^{d-3} > 2^{d-3}\left(\frac{d}{2}\right)^d.$$

This implies that (18) holds at $k = d - 3$ when $d \ge 10$. In other words, (2) implies (17) when $k = d - 3$ and $d \ge 10$. By monotonicity of the left hand side of (18), proved above, it follows that (2) implies (17) for all $d \ge 10$ and k such that $d/2 < k \le d - 3$.

Next, we show that (2) fails when $k = 2d/3$, and hence $k^+_{\mathrm{ind}}(d) < 2d/3$. Note that $2d/3 \le d - 3$ as $d \ge 10$. After substitution and cancellation, we see that (2) fails when $k = 2d/3$ if and only if

$$\left(\frac{3}{2^{5/3}}\right)^d \le d,$$

and since $3 < 2^{5/3}$, this inequality holds for all $d \geq 1$.

Putting this together, we see that if $d \geq 10$ and $k \leq k^+_{\text{ind}}(d)$ then (2) holds, and hence (17) holds. This completes the proof. $\square$

4.1 Effect of short cycles

A subpairing is a set of at most $dn/2$ pairs of points. If the corresponding multigraph $G(P)$ is a cycle of length j, then we say that the subpairing P is a cycle of length j.

Let $\boldsymbol{x} = (x_1, \ldots, x_m)$ be a sequence of nonnegative integers, for some fixed $m \geq 1$. Define $J = \sum_{j=1}^m jx_j$ and assume that $J > 0$. Write $r = \sum_{j=1}^m x_j$.

Let $\mathcal{S}(\boldsymbol{x})$ be the set of sequences $(P_1, \ldots, P_r)$ of subpairings such that $G(P_1), \ldots,$ $G(P_r)$ are distinct cycles so that the first x_1 have length 1, the next x_2 have length 2, etc, and the last x_m cycles have length m. Next, let $\mathcal{S}^*(\boldsymbol{x})$ be the set of all sequences $(P_1, \ldots, P_r)$ in $\mathcal{S}(\boldsymbol{x})$ so that the cycles $G(P_1), \ldots, G(P_r)$ are vertex-disjoint. Then

$$\frac{\mathbf{E}(Y(X_1)_{x_1} \cdots (X_m)_{x_m})}{\mathbf{E}Y} = \frac{1}{\mathbf{E}Y} \sum_{(P_1,\ldots,P_r)\in\mathcal{S}(\boldsymbol{x})} \sum_{P\supset P_1\cup\cdots\cup P_r} \frac{Y(P)}{M(dn)} \tag{19}$$

where the second sum is over all pairings $P \in \Omega_{n,d}$ which contain $P_1 \cup \cdots \cup P_r$. First we perform the summation over $(P_1, \ldots, P_r) \in \mathcal{S}^*(\boldsymbol{x})$, where all cycles are vertex-disjoint. Later we prove that the full sum is asymptotically equal to this restricted sum. For all positive integers j, define

$$\delta_j = \left(\frac{d - 2k + 1}{d - 1}\right)^j. \tag{20}$$

Observe that $\delta_j > -1$ for all $j \geq 1$ whenever $k < d$.

Lemma 4.2. *Let $\boldsymbol{x} = (x_1, \ldots, x_m)$ be a sequence of nonnegative integers, for some fixed $m \geq 1$. Then*

$$\frac{1}{\mathbf{E}Y} \sum_{(P_1,\ldots,P_r)\in\mathcal{S}^*(\boldsymbol{x})} \sum_{P\supset P_1\cup\cdots\cup P_r} \frac{Y(P)}{M(dn)} \sim \prod_{j=1}^m \big(\lambda_j(1 + \delta_j)\big)^{x_j}.$$

Proof. Define $\Sigma^*(\boldsymbol{x})$ by

$$\Sigma^*(\boldsymbol{x}) = \sum_{(P_1,\ldots,P_r)\in\mathcal{S}^*(\boldsymbol{x})} \sum_{P\supset P_1\cup\cdots\cup P_r} \frac{Y(P)}{M(dn)}.$$

We will choose distinct vertex-disjoint cycles $C_1, \ldots, C_r$ so that C_i has length j_i, where $(j_1, \ldots, j_r)$ is the vector

$$\big(\underbrace{1, 1, \ldots, 1}_{x_1}, \underbrace{2, 2, \ldots, 2}_{x_2}, \ldots, \underbrace{m, m \ldots, m}_{x_m}\big)$$

of lengths. Note that $\sum_{i=1}^r j_i = J$. There are

$$\frac{(n)_J}{2^r \prod_{i=1}^r j_i}$$

ways to choose these cycles. Next, we choose an orientation of the edges of cycles $C_1, \ldots, C_r$ in

$$2^r \prod_{i=1}^r \sum_{p_i=0}^{\lfloor j_i/2 \rfloor} \binom{j_i}{2p_i}$$

ways, where p_i is the number of sources on C_i, which equals the number of sinks on C_i. This follows as sources and sinks must alternate around C_i, so having chosen a subset of $2p_i$ vertices of C_i, there are 2 possible orientations of C_i which make these $2p_i$ vertices the sources/sinks. Then there are $(d(d-1))^J$ ways to assign points to the pairs corresponding to the J edges in $C_1 \cup \cdots \cup C_J$. Given the above orientation, we already know which point is the out-point of each pair of these J pairs.

Next, we will choose which vertices are going to be centres and which are going to be leaves. Suppose that ℓ_i sources on C_i are leaves, so that $p_i - \ell_i$ sources on C_i are centres. The $j_i - p_i$ non-sources on C_i must all be centres, as they have positive in-degree. Therefore there are $j_i - \ell_i$ centres on C_i, giving $J - L$ centres in $C_1 \cup \cdots \cup C_r$, where $L = \sum_{i=1}^r \ell_i$. There are

$$\prod_{i=1}^r \binom{p_i}{\ell_i}$$

ways to identify all leaves, and hence all centres, on $C_1 \cup \cdots \cup C_r$.

There are $\binom{n-J}{\frac{dn}{2k}-J+L}$ ways to choose the remaining centres, from the $n - J$ vertices not in $C_1 \cup \cdots \cup C_r$. Since J and L are bounded, we have

$$\binom{n-J}{\frac{dn}{2k} - J + L} \sim \left(\frac{d}{2k}\right)^{J-L} \left(\frac{2k-d}{2k}\right)^L \binom{n}{\frac{dn}{2k}}. \tag{21}$$

If v is one of these new centres then there are $\binom{d}{k}$ ways to select the in-points for v. If v is one of the $p_i - \ell_i$ sources on C_i which becomes a centre then there are $\binom{d-2}{k}$ ways to choose the in-points for v. If v is one of the ℓ_i sources on C_i which become a leaf then all points of v are out-points. If a vertex v of C_i is neither a source nor a sink then there are $\binom{d-2}{k-1}$ ways to select the remaining in-points for v. Finally, if v is a sink on C_i then there are $\binom{d-2}{k-2}$ ways to select the remaining in-points for v. Now that all in-points are identified, there are $(dn/2 - J)!$ ways to complete the oriented pairing, matching out-points to in-points, since the pairs in $C_1 \cup \cdots \cup C_r$ are already chosen and oriented. Again we divide by $M(dn)$ for the probability that we observe this pairing.

Using (21), this leads to the expression

$$\Sigma^*(\boldsymbol{x})$$
$$\sim \frac{\big(d(d-1)\,n\big)^J}{\big(\prod_{i=1}^r j_i\big)} \frac{(dn/2-J)!}{M(dn)} \binom{d}{k}^{dn/(2k)-J} \binom{n}{\frac{dn}{2k}} \left(\frac{d}{2k}\right)^J$$

$$\times \prod_{i=1}^{r} \sum_{p_i=0}^{\lfloor j_i/2 \rfloor} \binom{j_i}{2p_i} \sum_{\ell_i=0}^{p_i} \left(\frac{2k-d}{2k} \binom{d}{k} \right)^{\ell_i} \binom{p_i}{\ell_i} \binom{d-2}{k}^{p_i-\ell_i} \binom{d-2}{k-1}^{j_i-2p_i} \binom{d-2}{k-2}^{p_i}.$$

Next we divide by $\mathbf{E}(Y)$, using (15). Since J is bounded, after much simplification this gives

$$\begin{aligned}
&\frac{\Sigma^*(\boldsymbol{x})}{\mathbf{E}Y} \\
&\sim \prod_{i=1}^{r} \frac{(d-k)^{j_i}}{j_i} \sum_{p_i=0}^{\lfloor j_i/2 \rfloor} \binom{j_i}{2p_i} \left(\frac{(k-1)(d-k-1)}{k(d-k)} \right)^{p_i} \sum_{\ell_i=0}^{p_i} \binom{p_i}{\ell_i} \left(\frac{(2k-d)(d-1)}{(d-k)(d-k-1)} \right)^{\ell_i} \\
&= \prod_{i=1}^{r} \frac{(d-k)^{j_i}}{j_i} \sum_{p_i=0}^{\lfloor j_i/2 \rfloor} \binom{j_i}{2p_i} \left(\frac{(k-1)(d-k-1)}{k(d-k)} \right)^{p_i} \left(1 + \frac{(2k-d)d}{(d-k)(d-k-1)} \right)^{p_i} \\
&= \prod_{i=1}^{r} \frac{(d-k)^{j_i}}{j_i} \sum_{p_i=0}^{\lfloor j_i/2 \rfloor} \binom{j_i}{2p_i} \left(\frac{k-1}{d-k} \right)^{2p_i}.
\end{aligned}$$

This shows that asymptotically, the effect of each (disjoint) cycle is independent. The sum over only even values of N in the generating function $F(x) = \sum_N a_n x^N$ is given by $\frac{1}{2}\big(F(x) + F(-x)\big)$. Using this, we conclude that

$$\begin{aligned}
\frac{\Sigma^*(\boldsymbol{x})}{\mathbf{E}Y} &\sim \prod_{i=1}^{r} \frac{(d-k)^{j_i}}{2j_i} \left(\left(1 + \frac{k-1}{d-k}\right)^{j_i} + \left(1 - \frac{k-1}{d-k}\right)^{j_i} \right) \\
&= \prod_{i=1}^{r} \frac{(d-1)^{j_i}}{2j_i} \left(1 + \left(\frac{d-2k+1}{d-1} \right)^{j_i} \right).
\end{aligned}$$

Since exactly x_j values of i satisfy $j_i = j$, for $i = 1, \ldots, m$, this completes the proof. □

It remains to prove that the sum over $\mathcal{S}^*(\boldsymbol{x})$ dominates the whole sum.

Lemma 4.3. *Let $\boldsymbol{x} = (x_1, \ldots, x_m)$ be a sequence of nonnegative integers, for some fixed $m \geq 1$. Then, recalling (13) and (20),*

$$\frac{\mathbf{E}(Y(X_1)_{x_1} \cdots (X_m)_{x_m})}{\mathbf{E}Y} \sim \prod_{j=1}^{m} \big(\lambda_j(1+\delta_j)\big)^{x_j}.$$

Proof. Using (19), it suffices to prove that the sum over all $(P_1, \ldots P_r) \in \mathcal{S}(\boldsymbol{x}) \setminus \mathcal{S}^*(\boldsymbol{x})$ contributes negligibly. We adjust the argument given in Lemma 4.2. Suppose that we choose cycles $(C_1, \ldots, C_r)$ with the appropriate lengths which are not vertex-disjoint. If $C_1 \cup \cdots \cup C_r$ has a vertices and b edges, then there are $O(n^a)$ ways to choose and orient the cycles and choose points for the pairs corresponding to edges in the cycles. There are $O(1)$ ways to choose centres within these cycles, and $O\Big(\binom{n}{\frac{dn}{2k}}\Big)$ ways to choose the remaining centres. There are $O\Big(\binom{d}{k}^{dn/(2k)}\Big)$ ways to choose the remaining out-points for all vertices, and $(dn/2-b)!$ ways to complete the oriented pairing, matching out-points to in-points. Note that $a<b$ as the cycles are not vertex-disjoint. Using (16), it follows that the contribution of all $(P_1,\ldots,P_r)$ such that the corresponding cycles together have a vertices and b edges, and are not vertex-disjoint, is $O(n^{a-b})=o(1)$. The proof follows since there are only $O(1)$ choices for a and b. □

To complete this section we investigate the infinite sum needed for (A3).

Lemma 4.4. *Let λ_j, δ_j be defined in (13), (20), for all positive integers j. If k is an integer such that $5\le d<2k$ and $(2k-d)^2<4k-d-2$ then*

$$\sum_{j=1}^{\infty}\lambda_j\delta_j^2=\tfrac{1}{2}\log\left(\frac{d-1}{4k-d-2-(2k-d)^2}\right).$$

Proof. From the definition of λ_j and δ_j, we have

$$\sum_{j=1}^{\infty}\lambda_j\delta_j^2=\tfrac{1}{2}\sum_{j=1}^{\infty}\frac{(d+1-2k)^{2j}}{j\,(d-1)^j}.$$

Our assumptions on k imply that $(d+1-2k)^2<d-1$, so the series converges. Hence

$$\sum_{j=1}^{\infty}\lambda_j\delta_j^2=-\tfrac{1}{2}\log\left(1-\frac{(d+1-2k)^2}{d-1}\right)=\tfrac{1}{2}\log\left(\frac{d-1}{4k-d-2-(2k-d)^2}\right),$$

as claimed. □

5 Second moment

Following [7], a *signature* is a set S of t points such that each cell/vertex has either k or 0 points in S. The signature equals the set of in-points of the orientation.

Let (S_1,S_2) be an ordered pair of signatures, corresponding to an ordered pair of orientations (O_1,O_2). For $i=0,1,\ldots,d-k$, let B_i be the number of vertices v which are centres in both orientations such that exactly $k-i$ points corresponding to v belong to $S_1\cap S_2$. Hence the total number of vertices which are centres in both orientations is exactly $\sum_{i=0}^{d-k}B_i$. The number of vertices v such that v is a centre in O_1 but not in O_2 is $\frac{dn}{2k}-\sum_{i=0}^{d-k}B_i$, and the same number of vertices are centres in O_2 but not in O_1. Exactly $n-\frac{dn}{k}+\sum_{i=0}^{d-k}B_i$ vertices are leaves in both O_1 and O_2.

Therefore, the number of ways to "classify" all vertices is

$$\binom{n}{B_0,B_1,\ldots,B_{d-k},\,\frac{dn}{2k}-\sum B_i,\,\frac{dn}{2k}-\sum B_i,\,n-\frac{dn}{k}+\sum B_i}.$$

(Here, and throughout this section, any sum involving B_i is assumed to be over $i = 0, \dots, d-k$ unless otherwise specified.)

Next, we must choose (S_1, S_2) to match these parameters $(B_0, \dots, B_{d-k})$. If a vertex is a leaf in both orientation then it contributes no points to S_1 or S_2. If v is a leaf in O_1 and a centre in O_2 then there are $\binom{d}{k}$ ways to assign in-points to v in (S_1, S_2). The same is true if v is a centre in O_1 and a leaf in O_2.

Now suppose that v is a centre in both O_1 and O_2, with $k - i$ points from v in $O_1 \cap O_2$. There are

$$\tau_i := \binom{d}{k-i,\, i,\, i,\, d-k-i} = \frac{d!}{(k-i)!\,(d-k-i)!\,i!^2}$$

ways to assign in-points, since $k-i$ points are in-points in both O_1 and O_2 (meaning that that part has the same orientation in both orientations), $d-k-i$ points are in-points in both O_1 and O_2 (and again, that part has the same orientation in O_1 and O_2), and there are i points which are out-points in O_1 but in-points in O_2: the same is true for points which are out-points in O_2 and in-points in O_1. Overall, there are

$$\binom{d}{k}^{\frac{dn}{k} - 2\sum B_i} \prod_{i=0}^{d-k} \tau_i^{B_i}$$

ways to choose (S_1, S_2).

Finally, we must pair up all the points, which we can do in

$$\Big(\sum (k-i)B_i\Big)!\ \Big(dn/2 - \sum (k-i)B_i\Big)!$$

ways. The first factor matches all points in $S_1 \cap S_2$ to all points in the complement of $(S_1 \cup S_2)$, while the first factor matches all points in $S_1 \setminus S_2$ to the points in $S_2 \setminus S_1$. We multiply by $1/M(dn)$ as usual, to give the probability that this particular pairing is observed.

Define the domain

$$\mathcal{D} = \Big\{ \boldsymbol{B} = (B_0, \dots, B_{d-k}) \in \mathbb{Z}^{d-k+1} \;\Big|\; \tfrac{(d-k)n}{k} \le \sum_{i=0}^{d-k} B_i \le \tfrac{dn}{2k}, \quad B_0, \dots, B_{d-k} \ge 0 \Big\}.$$

Then

$$\begin{aligned} \mathbf{E}(Y^2) &= \frac{1}{M(dn)} \sum_{\boldsymbol{B} \in \mathcal{D}} \binom{n}{B_0, \dots, B_{d-k},\, \frac{dn}{2k} - \sum B_i,\, \frac{dn}{2k} - \sum B_i,\, n - \frac{dn}{k} + \sum B_i} \\ &\qquad \times \binom{d}{k}^{dn/k - 2\sum B_i} \Big(\sum (k-i)B_i\Big)! \Big(\frac{dn}{2} - \sum (k-i)B_i\Big)! \prod_{i=0}^{d-k} \tau_i^{B_i} \\ &= \sum_{\boldsymbol{B} \in \mathcal{D}} J_n(\boldsymbol{B}) \end{aligned} \tag{22}$$

where, for all $\boldsymbol{B} \in \mathcal{D}$,

$$J_n(\boldsymbol{B}) = \frac{n!\,(dn/2)!\,2^{dn/2}\,\big(\sum(k-i)B_i\big)!\,\big(\frac{dn}{2}-\sum(k-i)B_i\big)!}{(dn)!\,\big(\frac{dn}{2k}-\sum B_i\big)!^2\,\big(n-\frac{dn}{k}+\sum B_i\big)!}\binom{d}{k}^{dn/k-2\sum B_i}\prod_{i=0}^{d-k}\frac{\tau_i^{B_i}}{B_i!}.$$

Let $x \vee y$ denote $\max(x,y)$. Applying Stirling's formula in the form

$$\log(N!) = N\log N - N + \tfrac{1}{2}\log(N\vee 1) + \tfrac{1}{2}\log 2\pi + O(1/(N+1)),$$

valid for all nonnegative integers, to $J_n(\boldsymbol{B})$ shows that

$$\begin{aligned}
J_n(\boldsymbol{B}) = {} & \frac{\sqrt{n/2}\,\binom{d}{k}^{dn/k}}{(2\pi)^{(d-k+1)/2}\,d^{dn/2}\,n^{(d-2)n/2}} \cdot \sqrt{\frac{\Big(\sum(k-i)B_i\Big)\Big(\frac{dn}{2}-\sum(k-i)B_i\Big)}{\Big(\frac{dn}{2k}-\sum B_i\Big)^2\Big(n-\frac{dn}{k}+\sum B_i\Big)\prod_{i=0}^{d-k}B_i}} \\
& \times \frac{\Big(\sum(k-i)B_i\Big)^{\sum(k-i)B_i}\Big(\frac{dn}{2}-\sum(k-i)B_i\Big)^{dn/2-\sum(k-i)B_i}}{\binom{d}{k}^{2\sum B_i}\Big(\frac{dn}{2k}-\sum B_i\Big)^{dn/k-2\sum B_i}\Big(n-\frac{dn}{k}+\sum B_i\Big)^{n-dn/k+\sum B_i}} \cdot \prod_{i=0}^{d-k}\left(\frac{\tau_i}{B_i}\right)^{B_i} \\
& \times \left(1 + O\left(\frac{1}{\frac{dn}{2}+1-\sum(k-i)B_i} + \frac{1}{\frac{dn}{2k}+1-\sum B_i}\right.\right. \\
& \qquad\qquad \left.\left. + \frac{1}{n-\frac{dn}{k}+1+\sum B_i} + \sum_{i=0}^{d-k}\frac{1}{B_i+1}\right)\right),
\end{aligned} \tag{23}$$

where any factor in the denominator which equals zero should be replaced by 1.

5.1 Laplace summation

We now wish to apply Laplace's method to compute the asymptotic value of the summation (22). The following lemma, from Greenhill, Janson and Ruciński [10], will help us perform our calculations. (Some notation has been changed slightly to avoid clashes with notation used here.)

Lemma 5.1 ([10, Lemma 6.3])**.** *Suppose the following:*

(i) $\mathcal{L} \subset \mathbb{R}^m$ *is a lattice with full rank* m.

(ii) $K \subset \mathbb{R}^m$ *is a compact convex set with non-empty interior.*

(iii) $\varphi : K \to \mathbb{R}$ *is a continuous function with a unique maximum at some interior point* $z_0 \in K^\circ$.

(iv) φ *is a twice continuously differentiable in a neighbourhood of* z_0 *and the Hessian* $H_0 := D^2\varphi(z_0)$ *is strictly negative definite.*

(v) $\psi : K_1 \to \mathbb{R}$ *is a continuous function on some neighbourhood* $K_1 \subset K$ *of* z_0 *with* $\psi(z_0) > 0$.

(vi) *For each positive integer* n *there is a vector* $\ell_n \in \mathbb{R}^m$.

(vii) *For each positive integer* n *there is a positive real number* A_n *and a function* $J_n : (\mathcal{L} + \ell_n) \cap nK \to \mathbb{R}$ *such that, as* $n \to \infty$,

$$J_n(\ell) = O(A_n e^{n\varphi(\ell/n)+o(n)}), \qquad \ell \in (\mathcal{L} + \ell_n) \cap nK,$$

and

$$J_n(\ell) = A_n(\psi(\ell/n) + o(1))e^{n\varphi(\ell/n)}, \qquad \ell \in (\mathcal{L} + \ell_n) \cap nK_1,$$

uniformly for ℓ *in the indicated sets.*

Then, as $n \to \infty$,

$$\sum_{(\mathcal{L}+\ell_n)\cap nK} J_n(\ell) \sim \frac{(2\pi)^{m/2}\,\psi(z_0)}{\det(\mathcal{L})\,\det(-H_0)^{1/2}}\, A_n\, n^{m/2}\, e^{n\varphi(z_0)}.$$

To apply this lemma to (22), we first introduce the shorthand notation

$$\beta_{\boldsymbol{b}} = \sum_{i=0}^{d-k} b_i, \qquad \gamma_{\boldsymbol{b}} = \sum_{i=0}^{d-k} (k-i)b_i.$$

Then define the scaled domain

$$K = \Big\{ \boldsymbol{b} = (b_0, \ldots, b_{d-k}) \in \mathbb{R}^{d-k+1} \;\Big|\; \tfrac{d-k}{k} \le b_0 + \cdots + b_{d-k} \le \tfrac{d}{2k}, \quad b_0, \ldots, b_{d-k} \ge 0 \Big\}$$

and let

$$A_n = \frac{d^{-dn/2}}{\sqrt{2(2\pi n)^{d-k+1}}} \binom{d}{k}^{dn/k}, \qquad \psi(\boldsymbol{b}) = \sqrt{\frac{\gamma_{\boldsymbol{b}}\,(d/2 - \gamma_{\boldsymbol{b}})}{(d/(2k) - \beta_{\boldsymbol{b}})^2\,(1 - d/k + \beta_{\boldsymbol{b}})\,\prod_{i=0}^{d-k} b_i}},$$

$$\varphi(\boldsymbol{b}) = g(\gamma_{\boldsymbol{b}}) + g(d/2 - \gamma_{\boldsymbol{b}}) - 2g(d/(2k) - \beta_{\boldsymbol{b}}) - g(1 - d/k + \beta_{\boldsymbol{b}}) - 2\beta_{\boldsymbol{b}} \log \binom{d}{k} + \sum_{i=0}^{d-k} b_i \log(\tau_i) - \sum_{i=0}^{d-k} g(b_i),$$

where $g(x) = x \log x$ for $x > 0$ and $g(0) = 0$.

The following crucial result, which establishes property (iii) of Lemma 5.1, will be proved in Section 6.

Lemma 5.2. *Assume that* d, k *are integers such that* $d \ge 5$ *and at least one of the following two conditions hold:*

(I) $d/2+1 < k \le k_{\mathrm{exist}}(d)$;

(II) $d/2 < k \le d/2 + \max\left\{1, \frac{\log d}{6}\right\}$.

Then the unique global maximum of φ *over* K *occurs at the point* $\boldsymbol{b}^*$ *where*

$$b_i^* = \left(\frac{d}{2k}\right)^2 \binom{d}{k}^{-2} \tau_i \quad \textit{for } i = 0, \ldots, d-k. \tag{24}$$

Next we prove some useful facts about b^*.

Lemma 5.3. *The point* $\boldsymbol{b}^*$ *satisfies*

$$\beta_{\boldsymbol{b}^*} = \sum_{i=0}^{d-k} b_i^* = \left(\frac{d}{2k}\right)^2, \qquad \gamma_{\boldsymbol{b}^*} = \sum_{i=0}^{d-k} (k-i)\, b_i^* = \frac{d}{4},$$

$$\begin{aligned}
\varphi(\boldsymbol{b}^*) &= \frac{d(k-2)}{2k} \log d + 2\log k - (d-2)\log 2 - \frac{2k-d}{k}\log(2k-d),\\
\psi(\boldsymbol{b}^*) &= \frac{2k^3}{(2k-d)^2} \left(\prod_{i=0}^{d-k} b_i^*\right)^{-1/2}.
\end{aligned}$$

Furthermore,

$$\sum_{i=0}^{d-k} (k-i)^2\, b_i^* = \frac{d}{4} + \frac{d(k-1)^2}{4(d-1)}.$$

Proof. Recall that $\tau_i = \binom{d}{k-i,i,i,d-k-i}$ is the number of ways to choose an ordered pair of subsets A_1, A_2, both of size k, from a set of size d, so that $A_1 \cap A_2 = k-i$. Hence $\sum_{i=0}^{d-k} \tau_i = \binom{d}{k}^2$. The stated value of $\beta_{\boldsymbol{b}^*}$ is obtained after multiplication by $\left(\frac{d}{2k}\right)^2 \binom{d}{k}^{-2}$, by definition of $\boldsymbol{b}^*$. Next,

$$\begin{aligned}
\sum_{i=0}^{d-k} (k-i)\tau_i &= \sum_{i=0}^{d-k} \frac{d!}{(k-i-1)!\, i!^2\, (d-k+i)!}\\
&= d \sum_{i=0}^{d-k} \frac{(d-1)!}{(k-1-i)!\, i!^2\, (d-k+i)!} = d\binom{d-1}{k-1}^2,
\end{aligned}$$

arguing as above. The stated value of $\gamma_{\boldsymbol{b}^*}$ follows, and now the expressions for $\varphi(\boldsymbol{b}^*)$ and $\psi(\boldsymbol{b}^*)$ can be obtained by direct substitution. Finally

$$\begin{aligned}
\sum_{i=0}^{d-k} (k-i)^2 \tau_i &= \left(\sum_{i=0}^{d-k} (k-i)\tau_i\right) + \left(\sum_{i=0}^{d-k} (k-i)_2 \tau_i\right)\\
&= d\binom{d-1}{k-1}^2 + \sum_{i=0}^{d-k} \frac{d!}{(k-i-2)!\, i!^2\, (d-k+i)!}
\end{aligned}$$

$$= d\binom{d-1}{k-1}^2 + d(d-1)\binom{d-2}{k-2}^2,$$

and multiplying by $\left(\frac{d}{2k}\right)^2 \binom{d}{k}^{-2}$ completes the proof of the final identity. □

To calculate $\det(-H_*)$ we prove the following lemma. The main tools in the proof are the Matrix Determinant Lemma [18, equation (6.2.3)], which says

$$\det(A + \boldsymbol{u}\boldsymbol{v}^T) = (1 + \boldsymbol{v}^T A^{-1} \boldsymbol{u}) \det A, \tag{25}$$

and the Sherman–Morrison Theorem [18, equation (3.8.2)]), which says

$$(A + \boldsymbol{u}\boldsymbol{v})^{-1} = A^{-1} - \frac{A^{-1} \boldsymbol{u}\boldsymbol{v}^T A^{-1}}{1 + \boldsymbol{v}^T A^{-1}\boldsymbol{u}}, \tag{26}$$

for any square matrix A and vectors $\boldsymbol{u}$, $\boldsymbol{v}$ of the same dimension.

Lemma 5.4. *Let D be a diagonal $m \times m$ real matrix with entries $(d_1, \dots, d_m)$, and let $\boldsymbol{w} = (w_1, \dots, w_m)^T$, $\boldsymbol{v} = (v_1, \dots, v_m)^T \in \mathbb{R}^m$, for some positive integer m. Then*

$$\det(D + \boldsymbol{v}\boldsymbol{v}^T - \boldsymbol{w}\boldsymbol{w}^T) = \left(\Big(1 + \sum_{i=1}^m v_i^2/d_i\Big)\Big(1 - \sum_{i=1}^m w_i^2/d_i\Big) + \Big(\sum_{i=1}^m v_i w_i/d_i\Big)^2\right) \prod_{j=1}^n d_j.$$

Proof. Applying (25) twice, we obtain

$$\begin{aligned}\det(D + \boldsymbol{v}\boldsymbol{v}^T - \boldsymbol{w}\boldsymbol{w}^T) &= \det(D - \boldsymbol{w}\boldsymbol{w}^T)\ \big(1 + \boldsymbol{v}^T(D - \boldsymbol{w}\boldsymbol{w}^T)^{-1} \boldsymbol{v}\big) \\ &= \det(D) \left(1 - \sum_{i=1}^m w_i^2/d_i\right) \Big(1 + \boldsymbol{v}^T(D - \boldsymbol{w}\boldsymbol{w}^T)^{-1} \boldsymbol{v}\Big).\end{aligned}$$

Then (26) implies that

$$\begin{aligned}1 + \boldsymbol{v}^T(D - \boldsymbol{w}\boldsymbol{w}^T)^{-1} \boldsymbol{v} &= 1 + \boldsymbol{v}^T \left(D^{-1} + \frac{D^{-1}\boldsymbol{w}\boldsymbol{w}^T D^{-1}}{1 - \sum_{i=1}^m w_i^2/d_i}\right) \boldsymbol{v} \\ &= 1 + \sum_{i=1}^m v_i^2/d_i + \frac{\big(\sum_{i=1}^m v_i w_i/d_i\big)^2}{1 - \sum_{i=1}^m w_i^2/d_i},\end{aligned}$$

and combining these expressions completes the proof. □

Since we will need them to calculate the Hessian, we note here that the partial derivatives of φ are given by

$$\begin{aligned}\frac{\partial\varphi}{\partial b_i} = -2\log\binom{d}{k} &+ \log\tau_i + (k-i)\log(\gamma_{\boldsymbol{b}}) - (k-i)\log(d/2 - \gamma_{\boldsymbol{b}}) - \log b_i \\ &+ 2\log(d/(2k) - \beta_{\boldsymbol{b}}) - \log(1 - d/k + \beta_{\boldsymbol{b}})\end{aligned} \tag{27}$$

for $i = 0, \dots, d-k$. Direct substitution shows that the point $\boldsymbol{b}^*$, given in (24), is a stationary point.

Lemma 5.5. *Suppose that $4k-d-2 > (2k-d)^2$, and recall that H_* denotes the Hessian of φ at the point $\boldsymbol{b}^*$. Then*

$$\det(-H_*) = \frac{2k^2}{(d-1)(2k-d)^2}\left(4k-d-2-(2k-d)^2\right)\prod_{i=0}^{d-k}\frac{1}{b_i^*} > 0.$$

Furthermore, if $\boldsymbol{b}^*$ is the unique global maximum of φ on K then H_* is negative definite.

Proof. Recall that $\beta_{\boldsymbol{b}}$, $\gamma_{\boldsymbol{b}}$ are functions of $\boldsymbol{b}$. Differentiating (27) again gives

$$\frac{\partial^2\varphi}{\partial b_i\,\partial b_j} = \frac{(k-i)(k-j)}{\gamma_{\boldsymbol{b}}} + \frac{(k-i)(k-j)}{d/2-\gamma_{\boldsymbol{b}}} - \frac{\mathbb{1}(i=j)}{b_i} - \frac{2}{d/(2k)-\beta_{\boldsymbol{b}}} - \frac{1}{1-d/k+\beta_{\boldsymbol{b}}}$$

where the indicator function $\mathbb{1}(i=j)$ equals 1 if $i=j$, and equals zero otherwise. Substituting $\boldsymbol{b}=\boldsymbol{b}^*$ and using Lemma 5.3, it follows that

$$-H_* = D + \boldsymbol{v}\boldsymbol{v}^T - \boldsymbol{w}\boldsymbol{w}^T$$

where D is a $(d-k+1)\times(d-k+1)$ diagonal matrix with diagonal entries $d_i = 1/b_i^*$ for $i=0,\ldots,d-k$, and the entries of $\boldsymbol{v}$ and $\boldsymbol{w}$ are given by

$$v_i = \frac{2k}{2k-d}\sqrt{\frac{4k-d}{d}},\qquad w_i = (k-i)\sqrt{8/d}\qquad \text{for } i=0,\ldots,d-k.$$

Using Lemma 5.3 we calculate

$$\sum_{i=0}^{d-k} v_i^2/d_i = \frac{d(4k-d)}{(2k-d)^2},\qquad \sum_{i=0}^{d-k} v_iw_i/d_i = \frac{k\sqrt{2(4k-d)}}{2k-d},\qquad \sum_{i=0}^{d-k} w_i^2/d_i = 2+\frac{2(k-1)^2}{d-1}.$$

The claimed value for $\det(-H_*)$ is obtained by applying Lemma 5.4 and simplifying, and the lemma assumption implies that $\det(-H_*) > 0$. The fact that H_* is negative definite follows from this as $\boldsymbol{b}^* \in K^\circ$ is the unique global maximum of φ on K, with no local maximum of φ on the boundary of K. □

5.2 Proof of Theorem 1.4 and Theorem 1.5

The result for $(d,k)=(4,3)$ is established in [7]. In the following we assume that $d\geq 5$ and $k > d/2$, and show that the conditions of Theorem 3.1 hold under the assumptions of Theorem 1.4 and Theorem 1.5, where the random variables Y, X_j are defined in Section 1.3.

Condition (A1) follows from Bollobás [2] with $\lambda_j = (d-1)^j/(2j)$ for $j\geq 1$, while condition (A2) is established in Lemma 4.3 with δ_j defined in (20). Next, (5) holds by definition of $k_{\text{exist}}(d)$ if $k\leq k_{\text{exist}}(d)$, and by direct computation if $k \leq d/2+\max\left\{1, \frac{\log d}{6}\right\}$. Hence the assumptions of Lemma 4.4 are satisfied, which implies that condition (A3) holds.

To show that condition (A4) holds we will apply Lemma 5.1. First we check the assumptions of this lemma:

(i) Let $\mathcal{L} = \mathbb{Z}^{d-k+1}$, which is a lattice with rank $m = d - k + 1$ and $\det(\mathcal{L}) = 1$.

(ii) The domain K is compact and convex with a non-empty interior.

(iii) The function $\varphi : K \to \mathbb{R}$ is continuous, and has a unique global maximum $\boldsymbol{b}^*$, by Lemma 5.2. Note that each value of k in the range $d/2 < k \le k_{\text{exist}}(d)$ is covered by assumption (I) or assumption (II).

(iv) The function $\varphi : K \to \mathbb{R}$ is twice differentiable in the interior of K, with a strictly negative definite Hessian, by Lemma 5.5. This uses the fact that (5) holds, as proved above.

(v) Let K_1 be an open ball around $\boldsymbol{b}^*$ with radius small enough to guarantee that $K_1 \subset K$. Then the function $\psi : K_1 \to \mathbb{R}$ is a continuous function and $\psi(\boldsymbol{b}^*) > 0$, by Lemma 5.3.

(vi) Let ℓ_n be the zero vector in $\mathbb{Z}^{d-k+1}$, for each n.

(vii) This condition follows from (23).

Thus, we can apply Lemma 5.1 to see that

$$\begin{aligned}\mathbf{E}Y^2 &\sim \frac{(2\pi)^{(d-k+1)/2}\ \psi(\boldsymbol{b}^*)}{\det(\mathcal{L})\ \det(-H_*)^{1/2}}\ A_n\, n^{(d-k+1)/2}\, e^{n\,\varphi(\boldsymbol{b}^*)} \\ &= \frac{k^2\sqrt{d-1}}{(2k-d)\sqrt{4k-d-2-(2k-d)^2}}\ \left(\frac{\binom{d}{k}^d\, k^{2k}}{2^{k(d-2)}\, d^d\, (2k-d)^{2k-d}}\right)^{n/k}.\end{aligned}$$

Dividing by $(\mathbf{E}Y)^2$, using (16), proves that

$$\frac{\mathbf{E}Y^2}{(\mathbf{E}Y)^2} \sim \sqrt{\frac{d-1}{4k-d-2-(2k-d)^2}}.$$

This shows that condition (A4) of Theorem 3.1 holds, by Lemma 4.4.

Since $\delta_j > -1$ for all $j \ge 1$, we conclude from Theorem 3.1 that a.a.s. $Y > 0$. By (14), this completes the proof of Theorem 1.5 and Theorem 1.4.

6 Unique global maximum

In this section we prove Lemma 5.2, showing that $\boldsymbol{b}^*$ is the unique global maximum of φ in K.

6.1 No global maxima on the boundary

First we consider the boundary of the domain K.

Lemma 6.1. *Let the assumptions of Lemma 5.2 hold. Suppose that we know that $\boldsymbol{x}^*$ is the unique maximum of φ in the interior of K. Then $\varphi(\boldsymbol{b}^*) > \varphi(\boldsymbol{b})$ for all points $\boldsymbol{b}$ on the boundary of K.*

Proof. First we consider the point $\boldsymbol{a} = (\frac{d}{2k}, 0, \ldots, 0)$; that is, $a_0 = \frac{d}{2k}$ and all other a_i are zero. Then $\beta_{\boldsymbol{a}} = \frac{d}{2k}$ and $\gamma_{\boldsymbol{a}} = \frac{d}{2}$. Plugging these values into φ, we find that

$$\varphi(\boldsymbol{a}) = \frac{d(k-1)}{k}\log(d) + \log(k) - \frac{2k-d}{2k}\log(2k-d) - \frac{(d-2)}{2}\log 2 - \frac{d}{2k}\log\binom{d}{k},$$

using the fact that $\tau_0 = \binom{d}{k}$. Then $\varphi(\boldsymbol{b}^*) > \varphi(\boldsymbol{a})$ if and only if

$$d\,\log\binom{d}{k} + 2k\log k > d\log d + k(d-2)\log 2 + (2k-d)\log(2k-d).$$

But this inequality is equivalent to (17). Note that $k_{\text{exist}}(d) \le k^+_{\text{ind}}(d)$ by definition, and $\lceil\frac{d+1}{2}\rceil \le k^+_{\text{ind}}(d)$ since (2) holds for this value of k, by direct computation. Furthermore, if $d \ge 404 > e^6$ then

$$d/2 + \tfrac{\log d}{6} \le k^+_{\text{ind}}(d).$$

(To see this, note that (2) holds when $k = 1.01d/2$ and $\frac{\log d}{6} < d/200$ when $d \ge 404$.) Hence the conditions of Lemma 4.1 hold if either (I) or (II) hold, so (17) holds and hence $\boldsymbol{a}$ is not a global maximum of φ.

Next, suppose that $\boldsymbol{b} \ne \boldsymbol{a}$ satisfies $\sum_{i=0}^{d-k} b_i = \frac{d}{2k}$. Then each of $\gamma_{\boldsymbol{b}}$, $\frac{d}{2} - \gamma_{\boldsymbol{b}}$ and $1 - \frac{d}{k} + \beta_{\boldsymbol{b}}$ are positive, noting that $\gamma_{\boldsymbol{y}} = \frac{d}{2}$ if and only if $\boldsymbol{y} = \boldsymbol{a}$. Choose i such that $b_i > 0$. In the expression for $\partial\varphi/(\partial b_i)$, every term is bounded except for the term corresponding to $\frac{d}{2k} - \beta_{\boldsymbol{b}}$, and this term contributes $-\infty$. Hence reducing b_i slightly while holding all other entries steady will lead to an increase in φ, showing that no point on the boundary $\sum_{i=0}^{d-k} b_i = \frac{d}{2k}$ can be a local maximum of φ.

Now, suppose that $\frac{d-k}{k} < \sum_{i=0}^{d-k} b_i < \frac{d}{2k}$ and $b_\ell = 0$ for some $\ell \in \{0, \ldots, d-k\}$. Take some $j \in \{0, \ldots, d-k\}$ such that $b_j > 0$, and replace (b_ℓ, b_j) by $(b_\ell+\varepsilon, b_j-\varepsilon) = (\varepsilon, b_j-\varepsilon)$. This leaves $\beta_{\boldsymbol{b}}$ unchanged. Let $\widehat{\varphi}(\varepsilon) = \varphi(\boldsymbol{b}+\varepsilon\,\boldsymbol{e}_\ell-\varepsilon\,\boldsymbol{e}_j)$, where $\boldsymbol{e}_i$ denotes the ith standard basis vector for $i = \ell, j$. By the chain rule,

$$\widehat{\varphi}'(0) = \log\tau_\ell - \log\tau_j + (\ell-j)\log(\gamma_{\boldsymbol{b}}) - (\ell-j)\log(d/2-\gamma_{\boldsymbol{b}}) - \log(b_\ell) + \log(b_j).$$

The only unbounded term is $-\log b_\ell$, which equals $+\infty$. Therefore, replacing $\boldsymbol{b}$ by $\boldsymbol{b} + \varepsilon\boldsymbol{e}_\ell - \varepsilon\boldsymbol{e}_j$ gives a new vector in K with a higher value of φ, for some sufficiently small $\varepsilon > 0$. This proves that no vector in K with a zero entry can be a local maximum of φ.

Finally, suppose that $\boldsymbol{b}$ satisfies $\sum_{i=0}^{d-k} b_i = \frac{d-k}{k}$ and that all entries of $\boldsymbol{b}$ are positive. Then the only unbounded term in $\partial\varphi/(\partial b_1)$ is $-\log(1 - \frac{d}{k} + \beta_{\boldsymbol{b}})$, which equals $+\infty$.

Hence increasing b_1 slightly leads to a new vector which lies inside K and has a higher value of φ. Therefore the boundary given by $\sum_{i=0}^{d-k} b_i = \frac{d-k}{k}$ does not contain any local maximum of φ. This completes the proof. $\square$

6.2 Stationary points

In this section we write $\mu = \binom{d}{k}$. Setting all partial derivatives equal to zero using (27), we find that at any stationary point of φ in the interior of K,

$$b_i = \frac{\tau_i \left(\frac{d}{2k} - \beta\right)^2}{\mu^2 \left(1 - \frac{d}{k} + \beta\right)} \left(\frac{\gamma}{d/2 - \gamma}\right)^{k-i} \tag{28}$$

for $i = 0, \ldots, d-k$. Here $(\beta, \gamma) = (\beta_{\boldsymbol{b}}, \gamma_{\boldsymbol{b}})$. Alternatively, we may treat β, γ as the variables and use (28) to define the values of b_i, whenever $\boldsymbol{b}$ is a stationary point of φ in the interior of K.

Next we write β, γ in terms of a single variable x. Define $x = \frac{\gamma}{d/2-\gamma}$. Rearranging for γ gives

$$\gamma = \gamma(x) = \frac{xd}{2(x+1)}. \tag{29}$$

Observe that $\gamma^* = \gamma(1)$. The polynomial f defined in (3) can also be written as

$$f(x) = \sum_{i=0}^{d-k} \frac{\tau_i}{\mu^2}\, x^{k-i}.$$

Summing (28), it follows that at any stationary point,

$$\beta = \beta(x) = \frac{\left(\frac{d}{2k} - \beta\right)^2}{\left(1 - \frac{d}{k} + \beta\right)} f(x). \tag{30}$$

Note that $\beta^* = \beta(1)$. Next, using $\gamma = \sum_{i=0}^{d-k} (k-i) b_i$ and substituting (28) and (29) into (30) gives

$$\begin{aligned} \frac{xd}{2(x+1)} &= \frac{\left(\frac{d}{2k} - \beta\right)^2}{\left(1 - \frac{d}{k} + \beta\right)} \sum_{i=0}^{d-k} \frac{(k-i)\,\tau_i}{\mu^2}\, x^{k-i} \\ &= \frac{\left(\frac{d}{2k} - \beta\right)^2}{\left(1 - \frac{d}{k} + \beta\right)}\, x f'(x) \\ &= \frac{\beta x f'(x)}{f(x)}, \end{aligned}$$

using (30) for the final equality. Dividing through by x, we obtain the identity

$$\frac{d}{2(x+1)} = \frac{\beta f'(x)}{f(x)}. \tag{31}$$

We want to show that this identity has a unique solution on $(0,\infty)$ at $x=1$.

Let

$$y = y(x) = \frac{(x+1)f'(x)}{k}.$$

Then (31) says that $\beta = \frac{d}{2k}\cdot\frac{f}{y}$, and we need $\frac{d-k}{k} < \beta < \frac{d}{2k}$ for the corresponding vector $\boldsymbol{b}$ to belong to K°. Therefore we are interested in values of x for which

$$\frac{2(d-k)}{d}\,y < f < y.$$

We can rewrite (30) as $(y-f)^2 = f - \frac{2(d-k)}{d}\,y$, or in terms of x:

$$\left(\frac{(x+1)f'}{k} - f\right)^2 = f - \frac{2(d-k)\,(x+1)\,f'}{dk}. \tag{32}$$

We need to show that $x=1$ is the unique solution of (32) on the interval $(0,1]$ under the constraints $\frac{2(d-k)}{d}\,y < f < y$.

Next, we solve (32) as a quadratic equation in y. We can rule out one of the branches using the constraint $y > f$, leading to

$$y = f - \frac{d-k}{d} + \frac{\sqrt{(d-k)^2 + d(2k-d)\,f}}{d} = (1+\eta)f$$

where $\eta = \eta(x)$ is defined in (4). It suffices to prove that the equation

$$\frac{(x+1)\,f'(x)}{k} = (1+\eta(x))\,f(x) \tag{33}$$

has a unique solution on $(0,\infty)$ at $x=1$. In the following subsections, we give two separate arguments under assumptions (I) and (II). First we collect some useful identities and facts about f.

Note that f is an increasing function of x when $x>0$, and hence $0 < f(x) < 1$ when $x \in (0,1)$. We will also use the expansion of f around 1, which can be obtained using inclusion-exclusion (see Wilf [22, Section 4.2]):

$$f(x) = \mu^{-1}\sum_{j=0}^{k}\binom{k}{j}\binom{d-j}{k-j}(x-1)^j. \tag{34}$$

For later use, observe that (34) gives

$$f(1) = 1, \qquad f'(1) = \frac{k^2}{d}, \qquad f''(1) = \frac{k^2(k-1)^2}{d(d-1)}, \tag{35}$$

and

$$\begin{aligned} y(x) &= \frac{(x-1)f'(x)}{k} + \frac{2f'(x)}{k} \\ &= \mu^{-1}\sum_{j=0}^{k}\binom{k}{j}\binom{d-j}{k-j}\left(\frac{j}{k} + \frac{2(k-j)^2}{k(d-j)}\right)(x-1)^j. \end{aligned} \tag{36}$$

6.3 Proof of Lemma 5.2. Part (I)

Here, we prove Lemma 5.2 under assumption (I), which is $d/2+1 < k \le k_{\text{exist}}(d)$. From Table 1, we find that $d \ge 9$. Then inequality (5) implies that $k \le (d+1+\sqrt{d-1})/2$, and hence $3k \le 2d$. Using condition (6), to complete the proof of Lemma 5.2 we must show that the equation (33) has no solution on $\left(0, \left(1+\frac{(2k-d)^2 d}{k(d-k)(4k-d-2-(2k-d)^2)}\right)^{-1}\right) \cup \left(\frac{5k-2d}{d-k}, \infty\right)$. First we consider the large values of x.

Lemma 6.2. *Suppose that $d \ge 9$ and $\frac{d}{2}+1 < k \le \frac{2d}{3}$. If $x \ge \frac{5k-2d}{d-k}$ then $(x+1)f' < kf$.*

Proof. Using (34) and (36), we can write

$$f - \frac{(x+1)f'}{k} = \mu^{-1} \sum_{j=0}^{k} \binom{k}{j}\binom{d-j}{k-j}\left(1 - \frac{j}{k} - \frac{2(k-j)^2}{k(d-j)}\right)(x-1)^j \tag{37}$$

$$= \frac{1}{k\mu}\sum_{j=0}^{k}\binom{k}{j}\binom{d-j-1}{k-j-1}(d-2k+j)(x-1)^j$$

$$\ge \sum_{\ell=1}^{2k-d} \ell (x-1)^{2k-d}\left(c_{2k-d+\ell}(x-1)^{\ell} - c_{2k-d-\ell}(x-1)^{-\ell}\right) \tag{38}$$

where $c_j = \frac{1}{\mu k}\binom{k}{j}\binom{d-j-1}{k-j-1}$ for $j = 0, \ldots, k$. For the inequality, note that our assumptions imply that $4k-2d \le k$, so the right hand side of (38) is a sum of the $j = 0, \ldots, 4k-2d$ terms of (37). The omitted terms, with $4k-2d < j \le k$, all make a positive contribution to (37).

We claim that

$$\frac{c_{2k-d-\ell} \cdot c_{2k-d+\ell-1}}{c_{2k-d-\ell+1} \cdot c_{2k-d+\ell}} \le \frac{8}{9}(x-1)^2 \quad \text{for } \ell = 1, \ldots, 2k-d-1, \tag{39}$$

and note that this implies that for all $\ell = 1, \ldots, 2k-d-1$,

$$\frac{c_{2k-d+\ell}}{c_{2k-d-\ell}}(x-1)^{2\ell} \ge \frac{9}{8} \cdot \frac{c_{2k-d+\ell-1}}{c_{2k-d-\ell+1}}(x-1)^{2(\ell-1)} \ge \cdots \ge \left(\frac{9}{8}\right)^{\ell}. \tag{40}$$

Hence if the claim holds then all terms in (38) with $\ell \le 2k-d-1$ are positive. To bound the remaining term with $\ell = 2k-d$, observe that

$$\frac{(2k-d)c_0}{(2k-d-1)c_1(x-1)} = \frac{(2k-d)(d-1)}{(2k-d-1)k(k-1)(x-1)} \le \frac{(d-1)(d-k)}{3k(k-1)(2k-d-1)},$$
$$\frac{(2k-d)c_{4k-2d}(x-1)}{(2k-d-1)c_{4k-2d-1}} = \frac{(2d-3k+1)(2d-3k)(x-1)}{2(2k-d-1)(3d-4k)} \ge \frac{3(2d-3k)(2d-3k+1)}{2(d-k)(3d-4k)}.$$

Under our assumptions, we have that

$$\left(1 + \frac{3(2d-3k)(2d-3k+1)}{2(d-k)(3d-4k)}\right)\left(\frac{9}{8}\right)^2 > 1 + \frac{(d-1)(d-k)}{3k(k-1)(2k-d-1)}. \tag{41}$$

To see this, suppose first that $2k-d \geq 4$. Then bounding $d-1 \leq 2(k-1)$ and $d-k \leq k$, the right-hand side of (41) is at most $1+\frac{2}{9} < \left(\frac{9}{8}\right)^2$. Similarly, if $2k-d=3$ and $k \geq 7$ we bound the right-hand side of (41) by $\frac{4}{3}$ and estimate

$$1+\frac{3(2d-3k)(2d-3k+1)}{2(d-k)(3d-4k)} = 1+\frac{3(k-6)(k-5)}{2(k-3)(2k-9)} \geq \frac{4}{3} > \frac{4}{3}\left(\frac{8}{9}\right)^2.$$

Here we use the fact that $3(k-6)(k-5)/\big(2(k-3)(2k-9)\big)$ is monotonically increasing with k. Finally, if $2k-d=3$ and $k=6$ then $d=9$ and direct substitution shows that the right-hand side of (41) equals $1+\frac{2}{15} < \left(\frac{9}{8}\right)^2$. Hence (41) holds in all cases.

Combining the terms in (38) for $\ell \in \{2k-d, 2k-d-1\}$ and using (40), (41), we see that

$$\begin{aligned}
&\frac{1}{2k-d-1}\sum_{\ell\in\{2k-d,2k-d-1\}} \ell(x-1)^{2k-d}\left(c_{2k-d+\ell}(x-1)^{\ell}-c_{2k-d-\ell}(x-1)^{-\ell}\right)\\
&\geq \left(1+\frac{3(2d-3k)(2d-3k+1)}{2(d-k)(3d-4k)}\right)c_{4k-2d-1}(x-1)^{4k-2d-1}\\
&\qquad\qquad\qquad -\left(1+\frac{(d-1)(d-k)}{3k(k-1)(2k-d-1)}\right)c_1(x-1)\\
&> \left(1+\frac{(d-1)(d-k)}{3k(k-1)(2k-d-1)}\right)\left(\left(\frac{8}{9}\right)^2 c_{4k-2d-1}(x-1)^{4k-2d-1}-c_1(x-1)\right)\\
&\geq 0.
\end{aligned}$$

Thus, the sum in (38) is positive and the lemma is proved.

It remains to establish claim (39). Using $\frac{c_j}{c_{j+1}} = \frac{(j+1)(d-j-1)}{(k-j)(k-j-1)}$, we find that

$$\begin{aligned}
&\frac{c_{2k-d-\ell}}{c_{2k-d-\ell+1}}\cdot\frac{c_{2k-d+\ell-1}}{c_{2k-d+\ell}}\\
&\qquad = \frac{(2(d-k)-\ell)(2k-d+\ell)}{(d-k)^2-\ell^2}\cdot\frac{(2(d-k)+\ell-1)(2k-d-\ell+1)}{(d-k)^2-(\ell-1)^2}.
\end{aligned} \tag{42}$$

Since $3k \leq 2d$, it follows that the denominators on the right hand side of (42) are always positive for $\ell \leq 2k-d-1$. Now

$$\begin{aligned}
\frac{(2(d-k)+\ell-1)(2k-d-\ell+1)}{(d-k)^2-(\ell-1)^2} &\leq \frac{(2(d-k)+\ell-1)(2k-d-\ell+1)}{(d-k)^2-\ell^2}\\
&\leq \frac{2(2(d-k)+\ell)(2k-d-\ell)}{(d-k)^2-\ell^2},
\end{aligned}$$

and hence by (42), we obtain

$$\frac{c_{2k-d-\ell}}{c_{2k-d-\ell+1}}\cdot\frac{c_{2k-d+\ell-1}}{c_{2k-d+\ell}} \leq \frac{2\big(4(d-k)^2-\ell^2\big)\big((2k-d)^2-\ell^2\big)}{\big((d-k)^2-\ell^2\big)^2}.$$

Next, using $3k \le 2d$, we observe that

$$\left(1-\frac{\ell^2}{4(d-k)^2}\right)\left(1-\frac{\ell^2}{(2k-d)^2}\right) \le 1-\frac{\ell^2}{(d-k)^2}.$$

Hence for $\ell = 1, \dots, 2k-d-1$, using the lower bound on x, we have

$$\frac{2\big(4(d-k)^2-\ell^2\big)\big((2k-d)^2-\ell^2\big)}{\big((d-k)^2-\ell^2\big)^2} \le \frac{8(2k-d)^2}{(d-k)^2} \le \frac{8}{9}(x-1)^2.$$

This completes the proof. □

To prove Lemma 5.2 under assumption (I), it only remains to show that equation (33) has no solution when $0 < x < \left(1+\frac{d(2k-d)^2}{k(d-k)(4k-d-2-(2k-d)^2)}\right)^{-1}$. Since η is monotonically decreasing and $\eta(0) = \frac{2k-d}{2(d-k)}$, it is sufficient to show that, for such x,

$$\frac{(x+1)f'}{k} \ge \frac{d}{2(d-k)}f. \tag{43}$$

Let $t = 1/x$ and

$$g(t) = f(t^{-1})t^k = \mu^{-1}\sum_{j=0}^{d-k}\binom{k}{j}\binom{d-k}{j}t^j. \tag{44}$$

Note that $g(t)$ is the PGF of the hypergeometric distribution with parameters $(d, k, d-k)$. We will also use the expansion of g around 1, given by

$$g(t) = \mu^{-1}\sum_{j=0}^{d-k}\binom{k}{j}\binom{d-j}{k}(t-1)^j. \tag{45}$$

Now

$$f'(t^{-1}) = -\frac{g'(t)}{t^{k-2}} + \frac{kg(t)}{t^{k-1}}. \tag{46}$$

Rewriting equation (43) in terms of $g(t)$ and t, we get

$$\frac{1+t^{-1}}{k}\left(-\frac{g'(t)}{t^{k-2}}+\frac{kg(t)}{t^{k-1}}\right) \ge \frac{d}{2(d-k)}\cdot\frac{g(t)}{t^k},$$

which is equivalent to the inequality

$$A(t) = (t+1)g - \frac{d}{2(d-k)}g - \frac{t(t+1)g'}{k} \ge 0 \tag{47}$$

for $t \ge 1 + \frac{d(2k-d)^2}{k(d-k)(4k-d-2-(2k-d)^2)}$.

Lemma 6.3. *If* (5) *holds then* $A''(t) > 0$ *for all* $t \ge 1$ *and*

$$A'(1) = \frac{k(4k-d-2-(2k-d)^2)}{2d(d-1)}.$$

Proof. It is sufficient to show that $A = A(1) + \sum_{j=1}^{d-k+1} a_j(t-1)^j$ where $a_j > 0$ for all j. Using the expansion from (45), we have

$$\begin{aligned}(t+1)g &= 2g + (t-1)g \\ &= 2g(1) + \mu^{-1} \sum_{j=1}^{d-k+1} \left(2\binom{k}{j}\binom{d-j}{k} + \binom{k}{j-1}\binom{d-j+1}{k} \right)(t-1)^j\end{aligned}$$

and

$$\begin{aligned}t(t+1)g' &= 2g' + 3(t-1)g' + (t-1)^2 g' \\ &= 2g'(1) + \mu^{-1} \sum_{j=1}^{d-k+1} \left(2(j+1)\binom{k}{j+1}\binom{d-j-1}{k} + 3j\binom{k}{j}\binom{d-j}{k} \right. \\ &\qquad \left. + (j-1)\binom{k}{j-1}\binom{d-j+1}{k} \right)(t-1)^j.\end{aligned}$$

Thus, we find that

$$a_{d-k+1} = \mu^{-1}\binom{k}{d-k}\left(1 - \tfrac{d-k}{k}\right) > 0$$

and, for $j = 1 \ldots, d-k$, we have $a_j = \mu^{-1}\binom{k}{j}\binom{d-j}{k} h_j$, where

$$\begin{aligned}h_j &= 2 + \tfrac{j(d-j+1)}{(k-j+1)(d-k-j+1)} - \tfrac{d}{2(d-k)} - \tfrac{2(k-j)(d-k-j)}{k(d-j)} - \tfrac{3j}{k} - \tfrac{j(j-1)(d-j+1)}{k(k-j+1)(d-k-j+1)} \\ &= 2 + \tfrac{j(d-j+1)}{k(d-k-j+1)} - \tfrac{d}{2(d-k)} - \tfrac{3j}{k} - \tfrac{2(k-j)(d-k-j)}{k(d-j)} \\ &= 2 + \tfrac{j}{d-k-j+1} - \tfrac{d}{2(d-k)} - \tfrac{2(d-k)}{d-j}.\end{aligned}$$

Direct calculations show that

$$h_1 = \frac{4k - d - 2 - (2k-d)^2}{2(d-1)(d-k)}. \tag{48}$$

Since (5) holds we see that $h_1 > 0$. For any positive $1 \le z \le d-k$, let

$$h(z) = 2 + \tfrac{z}{d-k-z+1} - \tfrac{d}{2(d-k)} - \tfrac{2(d-k)}{d-z}.$$

Note that $d - z \ge 2(d-k-z+1)$ for all $z \ge 1$. Therefore,

$$h'(z) = \tfrac{d-k+1}{(d-k-z+1)^2} - \tfrac{2(d-k)}{(d-z)^2} \ge \frac{4(d-k+1) - 2(d-k)}{(d-z)^2} > 0,$$

which implies that $h_j = h(j) \ge h(1) > 0$ for all $j = 2, \ldots, d-k+1$. This completes the proof of the first statement, and the expression for $A'(1)$ follows as $A'(1) = a_1$. $\square$

We find from (45) that

$$g(1) = 1 \quad \text{and} \quad g'(1) = k(d-k)/d. \tag{49}$$

Thus, $A(1) = -\frac{(2k-d)^2}{2d(d-k)}$. Using Lemma 6.3, we find that, for $t \geq 1 + \frac{(2k-d)^2 d}{k(d-k)(4k-d-2-(2k-d)^2)}$,

$$A(t) \geq A(1) + (t-1)A'(1) \geq -\frac{(2k-d)^2}{2d(d-k)} + \frac{(2k-d)^2 d}{k(4k-d-2-(2k-d)^2)} A'(1) \geq 0.$$

This establishes (47) and completes the proof of Lemma 5.2 under assumption (I).

6.4 Proof of Lemma 5.2. Part (II) for $k = \lceil (d+1)/2 \rceil$.

In this section we prove Lemma 5.2 for the case when $k = k_0(d)$ defined by

$$k_0(d) = \begin{cases} (d+1)/2 & \text{when } d \text{ is odd,} \\ d/2 + 1 & \text{when } d \text{ is even.} \end{cases}$$

However, some of the lemmas in this section will be proved under weaker assumptions on k, so that we can reuse them in Section 6.5.

Lemma 6.4. *Suppose that $d \geq 5$ and $k = k_0(d)$. If $x > 1$ then*

$$\frac{(x+1)\, f'(x)}{k} < \big(1 + \eta(x)\big)\, f(x).$$

Proof. Since equality holds in (33) when $x = 1$, it suffices to prove that the derivative of the LHS is strictly smaller than the derivative of the RHS whenever $x > 1$. The left hand side of (33) equals $y(x)$. Define

$$R_1(x) = f'(x), \qquad R_2(x) = \frac{(2k-d)\, f'(x)}{2\sqrt{(d-k)^2 + d(2k-d)\, f(x)}}.$$

Then the derivative of the left hand side of (33) is given by

$$y'(x) = \frac{1}{k}\big((x+1)\, f''(x) + f'(x)\big),$$

while the derivative of the right hand side equals $R_1(x) + R_2(x)$. We claim that

(a) $y'(1) < R_1(1) + R_2(1)$;

(b) $(R_1 - y')^{(\ell)}(1) \geq 0$ for all $\ell \geq 1$. That is, the ℓ-th order derivatives of $R_1(x) - y'(x)$ are all nonnegative at $x = 1$;

(c) $R_2'(x) > 0$ for all $x > 1$.

Before establishing these facts, we show why they complete the proof. From (b), we conclude that $R_1(x) - y'(x)$ is an increasing function of x on $(1, \infty)$, as all non-constant coefficients of the Taylor expansion of $R_1(x) - y'(x)$ around $x = 1$ are nonnegative. From (c) we conclude that $R_2(x)$ is a strictly increasing function of x on $(1, \infty)$. Then (a) implies that $y'(x) < R_1(x) + R_2(x)$ for all $x > 1$, as required. Thus it remains to establish (a), (b) and (c).

Using (35), we calculate that

$$R_1(1) = f'(1) = \frac{k^2}{d}, \quad R_2(1) = \frac{(2k-d)k}{2d}, \quad y'(1) = \frac{k\big(d - 1 + 2(k-1)^2\big)}{d(d-1)}.$$

Direct substitution shows that

$$(4k - d - 2)(d - 1) > 4(k - 1)^2$$

when $k = k_0(d)$ and $d \geq 5$. Hence (a) holds.

Since $(f - y)' = R_1 - y'$ and

$$1 - \frac{j}{k} - \frac{2(k-j)^2}{k(d-j)} = \frac{(k-j)(d-2k+j)}{k(d-j)},$$

it follows from (34) and (45) that all non-constant coefficients of $R_1(x) - y'(x)$ are nonnegative, noting that the non-constant coefficients of $R_1 - y'$ corresponds to terms with $j \geq 2$. By Taylor's Theorem, we conclude that (b) holds.

Finally we consider (c). Note that

$$\left(R_2^2\right)' = \frac{(2k-d)^2\, f'}{d^2\, \left((d-k)^2 + d(2k-d)\, f\right)^2} \left(f'' \left(2(d-k)^2 + 2d(2k-d)\, f\right) - d(2k-d)\, (f')^2\right).$$

Furthermore, $R_2(x) > 0$ for all $x \geq 1$, by definition. Hence $R_2'(x) > 0$ if and only if

$$f''(x) \left(2(d-k)^2 + 2d(2k-d)\, f(x)\right) - d(2k-d) \left(f'(x)\right)^2 > 0. \tag{50}$$

Using (3) we can write $f(x) = \sum_{j=0}^{k} \alpha_j x^j$, where the coefficients α_j are positive. Then

$$f''(x) + x^{-1} f'(x) = \alpha_1 x^{-1} + \sum_{j=2}^{k} j^2 \alpha_j\, x^{j-2} = \sum_{j=1}^{k} \left(j\sqrt{\alpha_j x^{j-2}}\right)^2,$$

and $f(x) \geq \sum_{j=1}^{k} \left(\sqrt{\alpha_j\, x^j}\right)^2$. Then the Cauchy–Schwarz inequality implies that

$$\left(f''(x) + x^{-1}\, f'(x)\right) f(x) \geq \left(\sum_{j=1}^{k} j\, \alpha_j x^{j-1}\right)^2 = \left(f'(x)\right)^2.$$

Hence, to establish (50) it suffices to prove that

$$d(2k-d) \left(f''(x) + x^{-1}\, f'(x)\right) f(x) < f''(x) \left(2(d-k)^2 + 2d(2k-d)\, f(x)\right),$$

which we rearrange as

$$d(2k-d)\,x^{-1}\,f'(x)\,f(x) < f''(x)\,\big(2(d-k)^2 + d(2k-d)\,f(x)\big).$$

But

$$f''(x)\,\big(2(d-k)^2 + d(2k-d)\,f(x)\big) \geq d(2k-d)\,f''(x)\,f(x)$$

so the desired inequality holds if $xf''(x) > f'(x)$ for all $x \geq 1$. Since f'' is an increasing function on $x \geq 1$, by the Mean Value Theorem we have

$$f'(x) \leq f'(1) + (x-1)f''(x),$$

so it suffices to check that $f'(1) \leq f''(1)$, again using the fact that f'' is monotonic increasing for $x \geq 1$. By (35),

$$\frac{d(d-1)}{k^2}\big(f''(1) - f'(1)\big) = (k-1)^2 - (d-1) \geq \left(\frac{d+1}{2}\right)^2 - d + 1 = \frac{(d-1)(d-5)}{4}.$$

Hence the result holds. □

It remains to prove that (31) has no solution with $x \in (0,1)$. This will follow from the next three lemmas.

Lemma 6.5. *If $x \in (0,1)$ then*

$$(1+\eta)f < \frac{d}{2(d-k)}\,f - \frac{(2k-d)^2}{2d(d-k)}\,f^2.$$

Proof. Since $f < 1$ for $x \in (0,1)$, we get that

$$\begin{aligned}\frac{d}{2(d-k)} &- \frac{d}{d-k+\sqrt{(d-k)^2 + d(2k-d)\,f}} \\ &= \frac{d^2(2k-d)\,f}{2(d-k)\big(d-k+\sqrt{(d-k)^2+d(2k-d)\,f}\,\big)^2} \\ &> \frac{2k-d}{2(d-k)}\,f.\end{aligned}$$

Therefore

$$\eta = \frac{2k-d}{d-k+\sqrt{(d-k)^2+d(2k-d)\,f}} < \frac{2k-d}{2(d-k)} - \frac{(2k-d)^2}{2d(d-k)}\,f.$$

The required bound follows after multiplying by $(2k-d)f/d$ and rearranging. □

Recall the function g from (44).

Lemma 6.6. *For all $t > 0$, we have that*

$$g(tg'' + g') > t(g')^2$$

and hence the rational function tg'/g is monotonically strictly increasing on $t > 0$.

Proof. Using the Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
g(tg''+g') &> t\mu^{-2}\sum_{j=1}^{d-k}\binom{k}{j}\binom{d-k}{j}t^{j-1}\cdot\sum_{j=1}^{d-k}j^2\binom{k}{j}\binom{d-k}{j}t^{j-1}\\
&\geq t\mu^{-2}\left(\sum_{j=1}^{d-k}j\binom{k}{j}\binom{d-k}{j}t^{j-1}\right)^2 = t(g')^2
\end{aligned}
$$

The first inequality is strict because we dropped the constant term in g. □

Now we are ready to prove the following bound.

Lemma 6.7. *Assume that* $2d \geq 3k$ *and* (5) *holds. Then, for any* $x \in (0,1)$*,*

$$\frac{(1+x)f'}{k} > (1+\eta)f.$$

Proof. By Lemma 6.5, it suffices to prove that for all $x \in (0,1)$,

$$\frac{(1+x)f'}{k} \geq \frac{d}{2(d-k)}\,f - \frac{(2k-d)^2}{2d(d-k)}\,f^2.$$

Using (46) we rewrite this inequality in terms of g: it suffices to prove that for all $t > 1$,

$$\frac{1+t^{-1}}{k}\left(-\frac{g'}{t^{k-2}}+\frac{kg}{t^{k-1}}\right) \geq \frac{d}{2(d-k)}\,\frac{g}{t^k} - \frac{(2k-d)^2}{2d(d-k)}\,\frac{g^2}{t^{2k}}.$$

This is equivalent to the inequality $A+B \geq 0$, where A is defined in (47) and

$$B = B(t) = \frac{(2k-d)^2\,g^2}{2d(d-k)\,t^k}.$$

Next we state a sequence of claims which together will imply the result.

Claim 1. $A'(t) \geq 0$ and $A''(t) \geq 0$ for $t \geq 1$.

Claim 2. Either $B''(t) \geq 0$ for all $t \geq 1$, or there exists some $t_0 > 1$ such that $B'(t) \geq 0$ for $t \geq t_0$ and $B''(t) \geq 0$ for $1 \leq t \leq t_0$.

Claim 3. $A(1)+B(1) = 0$ and $A'(1)+B'(1) \geq 0$.

First, we complete the proof based on these claims. By a slight abuse of notation, set $t_0 = \infty$ in the first case of Claim 2. Note that

$$(A+B)' \geq A'(1)+B'(1) \geq 0$$

for $1 \leq t \leq t_0$, since both A and B are convex on this interval. For $t \geq t_0$, we have $(A+B)' \geq 0$ since both A and B are increasing. Therefore, $A+B$ is an increasing function on $[1, +\infty)$ and thus $A + B \geq A(1) + B(1) = 0$, as required. Next, observe that Claim 1 follows from Lemma 6.3. Hence it remains to establish Claims 2 and 3.

For Claim 2, direct calculations show that

$$\frac{t\,g'(t)}{g(t)} \to \begin{cases} k(d-k)/d & \text{as } t \to 1, \\ d-k & \text{as } t \to \infty, \end{cases}$$

By our assumptions, we have $k(d-k)/d < k/2 \leq d-k$. Hence, using Lemma 6.6, if $3k < 2d$ then there is a unique point $t_0 \in (1, \infty)$ such that

$$\frac{t_0\,g'(t_0)}{g(t_0)} = \frac{k}{2},$$

while if $3k = 2d$ then tg'/g increases strictly with limit $k/2$ as $t \to \infty$, and we set $t_0 = \infty$. We calculate directly that

$$\left(\frac{g(t)^2}{t^k}\right)' = \frac{g}{t^{k+1}}\,(2t\,g' - k\,g)\,.$$

This expression is strictly positive when $t > t_0$, by definition of t_0. This proves that $B'(t) \geq 0$ for $t \geq t_0$, as B is a positive multiple of g^2/t^k. Differentiating again gives

$$\begin{aligned}\left(\frac{g(t)^2}{t^k}\right)'' &= \frac{2\,(g')^2 + 2g\,g''}{t^k} - \frac{4kg\,g'}{t^{k+1}} + \frac{k(k+1)\,g^2}{t^{k+2}} \\ &\geq \frac{4t^2\,(g')^2 - 2(2k+1)t\,g\,g' + k(k+1)\,g^2}{t^{k+2}} + \frac{2g\,g' - 2t\,(g')^2 + 2t\,g\,g''}{t^{k+1}} \\ &\geq \frac{1}{t^{k+2}}\,(k\,g - 2t\,g')\,((k+1)g - 2t\,g')\,.\end{aligned}$$

The final inequality follows since the last term of the penultimate line (with denominator t^{k+1}) is nonnegative, by the first statement of Lemma 6.6. Therefore if $1 \leq t \leq t_0$ then, again by Lemma 6.6,

$$\frac{t\,g'(t)}{g(t)} \leq \frac{k}{2}$$

and hence $B''(t) \geq 0$. This completes the proof of Claim 2.

For Claim 3, recalling (49), we have

$$A(1) = -\frac{(2k-d)^2}{2d(d-1)} = -B(1).$$

Using (48), we obtain that

$$A'(1) = a_1 = \mu^{-1}k\binom{d-1}{k}c_1 \geq \frac{k(4k-d-2-(2k-d)^2)}{2d^2}.$$

Observe also that

$$B'(1) = \frac{(2k-d)^2}{2d(d-k)}\left(\frac{2k(d-k)}{d} - k\right) = -\frac{k(2k-d)^3}{2d^2(d-k)}.$$

Therefore, by assumptions, we find that

$$A'(1) + B'(1) \geq \frac{k}{2d^2(d-k)}\left((d-k)(4k-d-2-(2k-d)^2) - (2k-d)^3\right).$$

Now using (5), the right hand side is nonnegative if $d-k > (2k-d)^3$, and this inequality holds when $3k \leq 2d$. This completes the proof of Claim 3 and the lemma. □

If $d \geq 5$ and $k = k_0(d)$ then $3k \leq 2d$ and (5) holds. Hence, by Lemma 6.7 we conclude that (31) has no solution in $(0,1)$. Recalling Lemma 6.4, this completes the proof of Lemma 5.2 when $k = k_0(d)$ and $d \geq 5$.

6.5 Proof of Lemma 5.2. Part (II) for large d.

In this subsection, we prove Lemma 5.2 for values of k in the range $k_0(d) < k \leq d/2 + \frac{\log d}{6}$ that is not covered in the previous subsection. Thus, we can assume that $d \geq 404 = \lceil e^6 \rceil$.

Lemma 6.8. *Assume that $d \geq 404$ and $d/2 + 1 < k \leq 2d/3$. Then*

$$\frac{(x+1)f'(x)}{k} < (1+\eta(x))f(x))$$

for all $x \in (1,2]$ such that $f(x) \leq k/(2k-d)^2$.

Proof. Recall that $f(1) = 1$ and hence $\eta(1) = (2k-d)/d$. Since both f and f' are strictly increasing functions of x, applying the Mean Value Theorem to η on the interval $[1,x]$ gives

$$\begin{aligned}\frac{\eta(1)-\eta(x)}{x-1} &= \frac{d(2k-d)^2\, f'(\xi)/\sqrt{(d-k)^2 + d(2k-d)\, f(\xi)}}{\left(d-k+\sqrt{(d-k)^2 + d(2k-d)\, f(\xi)}\right)^2} \\ &\leq \frac{(2k-d)^2\, f'(x)}{2kd} \leq \frac{f'(x)}{2d\, f(x)}\end{aligned}$$

for some $\xi \in (1,x)$. This uses the assumed upper bound on $f(x)$ for the final inequality. Hence to prove the lemma, it suffices to prove that

$$\frac{(x+1)f'}{k} < \frac{2k\, f}{d} - \frac{(x-1)f'}{2d}. \tag{51}$$

Similarly to the proof of Lemma 6.2, using (34), we expand as follows:

$$\begin{aligned}&\frac{2k\, f}{d} - \frac{(x+1)f'}{k} - \frac{(x-1)f'}{2d} \\ &\quad = \mu^{-1}\sum_{j=0}^{k}\binom{k}{j}\binom{d-j}{k-j}\left(\frac{2k}{d} - \frac{j}{k} - \frac{2(k-j)^2}{k(d-j)} - \frac{j}{2d}\right)(x-1)^j.\end{aligned} \tag{52}$$

Observe that, for all $0 \le j \le k$,

$$\frac{2k}{d} - \frac{j}{k} - \frac{2(k-j)^2}{k(d-j)} - \frac{j}{2d} = \frac{j}{d}\left(\frac{(2k-d-j)(d-k)}{k(d-j)} + \frac{1}{2}\right). \tag{53}$$

Summing the terms of (52) corresponding to $j \in \{0, k-1, k\}$ together and using (53) gives

$$\begin{aligned}
\sum_{j\in\{1,k-1,k\}} & \binom{k}{j}\binom{d-j}{k-j}\frac{j}{d}\left(\frac{(2k-d-j)(d-k)}{k(d-j)} + \frac{1}{2}\right)(x-1)^j \\
&> \frac{k}{2d}\binom{d-1}{k-1}(x-1) - \frac{(k-1)(2d-3k)}{2d}(x-1)^{k-1} - \frac{2d-3k}{2d}(x-1)^k \\
&\ge \frac{k(x-1)}{2d}\left(\binom{d-1}{k-1} - (d-k+2)(2d-3k)\right).
\end{aligned} \tag{54}$$

Under our assumptions, we have

$$\binom{d-1}{k-1} - (d-k+2)(2d-3k) > \frac{(d-2)}{2}\big((d-1) - (d-k+2)\big) \ge 0,$$

since $\binom{d-1}{k-1} > (d-1)(d-2)/2$ and $2d-3k \le (d-2)/2$. Hence the sum over $j \in \{1, k-1, k\}$ in (54) is strictly positive.

Now, for $2 \le j \le k/2$, we will sum the terms corresponding to $(x-1)^j$ and $(x-1)^{k-j}$ in (52). First observe that when $2 \le j \le k/2$,

$$\frac{2k-d-j}{d-j} > -\frac{1}{2}$$

and hence

$$\frac{(2k-d-j)(d-k)}{k(d-j)} + \frac{1}{2} \ge -\frac{(d-k)}{2k} + \frac{1}{2} = \frac{2k-d}{2k} > 0.$$

Next,

$$j\binom{k}{j}\binom{d-j}{k-j} \ge (k-j)\binom{k}{k-j}\binom{d-k+j}{j},$$

so we can bound the term corresponding to $(x-1)^j$ from below as

$$\begin{aligned}
&\binom{k}{j}\binom{d-j}{k-j}\frac{j}{d}(x-1)^j\left(\frac{(2k-d-j)(d-k)}{k(d-j)} + \frac{1}{2}\right) \\
&\qquad \ge \frac{(k-j)}{d}\binom{k}{k-j}\binom{d-k+j}{j}(x-1)^{k-j}\left(\frac{(2k-d-j)(d-k)}{k(d-j)} + \frac{1}{2}\right).
\end{aligned}$$

Finally, observe that

$$\frac{(2k-d-j)(d-k)}{k(d-j)} + \frac{(2k-d-(k-j))(d-k)}{k\big(d-(k-j)\big)} + 1$$

$$\geq \frac{(2k-d)(d-k)}{kd} + \frac{(k-d)(d-k)}{k(d-k)} + 1$$
$$= 1 - \frac{2(d-k)^2}{kd} \geq 0.$$

Combining these facts shows that the sum of the $(x-1)^j$ and $(x-1)^{k-j}$ terms is nonnegative for $2 \leq j \leq k/2$, completing the proof. □

To complete this section we prove the following.

Lemma 6.9. *Suppose that $d \geq 404$ and $d/2 + 1 < k \leq d/2 + \frac{\log d}{6}$. If $x > 1$ then*

$$\frac{(x+1)\, f'(x)}{k} < \big(1 + \eta(x)\big)\, f(x).$$

Proof. By our assumptions, $k \leq 4d/7$ and hence the assumptions of Lemma 6.2 and Lemma 6.8 hold and $\frac{5k-2d}{d-k} \leq 2$. By Lemma 6.2 and Lemma 6.8, if

$$f(x) \leq \frac{k}{(2k-d)^2} \qquad \text{for all } 1 < x \leq \frac{5k-2d}{d-k} \tag{55}$$

then the lemma holds. For the remainder of the proof, suppose that $1 < x \leq \frac{5k-2d}{d-k}$. It follows from (51) that

$$\frac{(x+1)f'}{k} \leq \frac{2k\, f}{d},$$

and hence $(\log f)' < \frac{2k^2}{d\,(x+1)}$. Integration gives

$$f(x) \leq \left(1 + \frac{x-1}{2}\right)^{2k^2/d} \leq \exp\left(\frac{k^2(x-1)}{d}\right)$$
$$\leq \exp\left(\frac{3k^2(2k-d)}{d(d-k)}\right).$$

Hence

$$\log(f(x)) \leq \frac{(d + \frac{1}{3}\log d)^2}{2d(d - \frac{1}{3}\log d)}\, \log d \leq \frac{3}{5}\log d,$$

while

$$\frac{k}{(2k-d)^2} \geq \frac{9d}{2(\log d)^2} = \frac{9d^{3/5}}{50}\left(\frac{d^{1/5}}{\log(d^{1/5})}\right)^2 \geq \frac{9e^2}{50}\, d^{3/5} \geq d^{3/5}.$$

This shows that (55) holds, completing the proof. □

As observed earlier, if $d \geq 404$ and $d/2 < k < d/2 + \frac{\log d}{6}$ then $3k \leq 2d$. Furthermore, (5) holds as $\frac{\log d}{6} < \frac{1+\sqrt{d-1}}{2}$ for $d \geq 404$. Hence the proof of Lemma 5.2 is completed in this case by combining Lemma 6.7 and Lemma 6.9, for $k > k_0(d)$, and using the results of Section 6.4 when $k = k_0(d)$.

Acknowledgements

Michelle Delcourt's research was supported by NSERC under Discovery Grant No. 2019-04269. Mikhail Isaev's research was supported by the Australian Research Council DECRA Fellowship DE200101045. Bernard Lidický's research was supported in part by NSF grant DMS-2152490 and the Scott Hanna Professorship. Luke Postle's research was partially supported by NSERC under Discovery Grant No. 2019-04304.

We thank the mathematical research institute MATRIX in Australia where part of this research was performed.

References

[1] E. A. Bender and E. R. Canfield, The asymptotic number of labeled graphs with given degree sequences, *Journal of Combinatorial Theory, Series A*, **24(3)** (1978), 296–307.

[2] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European Journal of Combinatorics*, **1(4)** (1980), 311–316.

[3] B. Bollobás, The isoperimetric number of random regular graphs, *European Journal of Combinatorics* **9** (1988), 241–244.

[4] R. A. Cameron and D. Horsley, Decompositions of complete multigraphs into stars of varying sizes, *Journal of Combinatorial Theory (Series B)* **145** (2020), 32–64.

[5] Y. Caro and J. Schönheim, Decomposition of trees into isomorphic subtrees, *Ars Combinatoria* **9** (1980), 119–130.

[6] J. Ding, A. Sly and N. Sun, Maximum independent sets on random regular graph, *Acta Mathematica* **217** (2016), 263–340.

[7] M. Delcourt and L. Postle, Random 4-regular graphs have 3-star-decompositions asymptotically almost surely, *European Journal of Combinatorics* **72** (2018), 97–111.

[8] W. Duckworth and M. Zito, Large independent sets in random regular graphs, *Theoretical Computer Science* **410** (2009), 5236–5243.

[9] A. Frieze and T. Łuczak, On the independence and chromatic numbers of random regular graphs, *Journal of Combinatorial Series B* **54** (1992), 123–132.

[10] C. Greenhill, S. Janson and A. Ruciński, On the number of perfect matchings in random lifts, *Combinatorics, Probability and Computing* **19** (2010), 791 – 817.

[11] V. Harangi, Improved replica bounds for the independence ratio of random regular graphs, *Journal of Statistical Physics* **190** (2023), Article number:60.

[12] D. Hoffman, The real truth about star designs, *Discrete Mathematics* **284** (2004), 177–180.

[13] S. Janson, Random regular graphs: asymptotic distributions and contiguity, *Combinatorics, Probability and Computing* **4** (1995), 369–405.

[14] A. Kotzig, From the theory of finite regular graphs of degree three and four, *Casopis Pést. Mat.* **82** (1957), 76–92 (in Slovak).

[15] L. M. Lovász, C. Thomassen, Y. Wu and C.-Q. Zhang, Nowhere-zero 3-flows and modulo k-orientations, *Journal of Combinatorial Theory (Series B)* **103** (2013), 587–598.

[16] R. Marino and S. Kirkpatrick, Large independent sets on random d-regular graphs with fixed degree d, Preprint, 2020. arXiv:2003.12293.

[17] B. D. McKay, Independent sets in regular graphs of high girth, *Ars Combinatorica* **23A** (1987), 179–185.

[18] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[19] R. W. Robinson and N. C. Wormald, Almost all cubic graphs are Hamiltonian, *Random Structures & Algorithms* **3** (1992), 117–125.

[20] M. Tarsi, Decomposition of complete multigraphs into stars, *Discrete Mathematics* **26** (1979), 273–278.

[21] C. Thomassen, The weak 3-flow conjecture and the weak circular flow conjecture, *Journal of Combinatorial Theory (Series B)* **102** (2012), 521–529.

[22] H.S. Wilf, *generatingfunctionology*, Academic Press, New York, 1994.

[23] N. C. Wormald, The asymptotic connectivity of labelled regular graphs, *Journal of Combinatorial Theory (Series B)* **31** (1981), 156–167.

[24] N. C. Wormald, Differential equations for random processes and random graphs, *Annals of Applied Probability* **5** (1995), 1217–1235.

[25] N. C. Wormald, Models of random regular graphs, in *Surveys in Combinatorics, 1999* (J. D. Lamb and D. A. Preece, eds.), *London Mathematical Society Lecture Note Series, vol. 267*, Cambridge University Press, Cambridge, 1999.

[26] S. Yamamoto, H. Ikeda, S. Shige-eda, K. Ushio and N. Hamada, On claw-decompositions of complete graphs and complete bigraphs, *Hiroshima Math. J.* **5** (1975), 33–42.