On the number of spanning trees in random regular graphs

Catherine Greenhill[*]

School of Mathematics and Statistics
The University of New South Wales
Sydney NSW 2052, Australia
c.greenhill@unsw.edu.au

Matthew Kwan[†]

School of Mathematics and Statistics
The University of New South Wales
Sydney NSW 2052, Australia
matthew.a.kwan@gmail.com

David Wind

DTU Compute
Technical University of Denmark
DK-2800 Kongens Lyngby, Denmark
utdiscant@gmail.com

18 February 2014

Abstract

Let $d \geq 3$ be a fixed integer. We give an asympotic formula for the expected number of spanning trees in a uniformly random d-regular graph with n vertices. (The asymptotics are as $n \to \infty$, restricted to even n if d is odd.) We also obtain the asymptotic distribution of the number of spanning trees in a uniformly random cubic graph, and conjecture that the corresponding result holds for arbitrary (fixed) d. Numerical evidence is presented which supports our conjecture.

1 Introduction

In this paper, d denotes a fixed integer which is at least 2 (and usually at least 3). All asymptotics are taken as $n \to \infty$, with n restricted to even integers when d is odd.

The number of spanning trees in a graph, also called the *complexity* of the graph, is of interest for a number of reasons. The complexity of a graph is an evaluation of the Tutte polynomial (see for example [14]). The Merino-Welsh conjecture [10] relates the complexity of a graph with two other graph parameters, namely, the number of acyclic orientations and the number of totally cyclic orientations of a graph. (Noble and

---

Royle [13] recently proved that the Merino-Welsh conjecture is true for series-parallel graphs.) The complexity of a graph also plays a role in the theory of electrical networks (see for example [12]).

We are interested in the number of spanning trees in random regular graphs. The first significant result in this area is due to McKay [7], who proved that for $d \geq 3$, the nth root of the number of spanning trees of a random d-regular graph with n vertices converges to

$$\frac{(d-1)^{d-1}}{(d^2-2d)^{d/2-1}} \tag{1}$$

as $n \to \infty$, with probability one. An alternative proof of this was later given by Lyons [6, Example 3.16].

McKay [8, Theorem 4.2] gave an asymptotic expression for the expected number of spanning trees in a random graph with specified degrees, up to some unknown constant. His result holds when the maximum degree is bounded and the average degree is bounded away from 2 (independently of n). When specialised to regular degree sequences, [8, Theorem 4.2] states that the expected number of spanning trees in $\mathcal{G}_{n,d}$ is asymptotic to

$$\frac{c_d}{n} \left(\frac{(d-1)^{d-1}}{(d^2-2d)^{d/2-1}} \right)^n, \tag{2}$$

for some unknown constant c_d.

Other work on asymptotics for the number of spanning trees has focussed on circulant graphs, grid graphs and tori (see for example [2] and the references therein).

Our first result, Theorem 1.1, provides the value of the constant c_d from (2), proving that

$$c_d = \exp\left(\frac{6d^2 - 14d + 7}{4(d-1)^2} \right) \frac{(d-1)^{1/2}}{(d-2)^{3/2}}.$$

For our second result we investigate the distribution of the number of spanning trees in random d-regular graphs using the small subgraph conditioning method, and obtain the asymptotic distribution in the case of cubic graphs, presented in Theorem 1.2. We provide partial calculations for arbitrary fixed degrees, which lead us to conjecture that the corresponding result holds in general (see Conjecture 1.3).

In order to precisely state our main results we must introduce some notation and terminology.

1.1   Notation and our main results

Let $\mathbb{N}$ denote the natural numbers (which includes 0). For integers n, k let $(n)_k$ denote the falling factorial $n(n-1)\cdots(n-k+1)$. Square brackets without subscripts denote extraction of coefficients of a generating function. We use $\mathbf{1}(\cdot)$ to denote both the indicator variable of an event and the characteristic function of a set (the particular set will appear as a subscript). We use standard asymptotic notation throughout, with the exception that $\rightsquigarrow$ indicates convergence in distribution of a sequence of random

variables. (We use this notation rather than $\xrightarrow{d}$ to avoid overloading the symbol "d", which we use for the degree of the graph.)

Let $\mathcal{G}_{n,d}$ denote the uniform model of d-regular simple graphs on the vertex set $\{1, \ldots, n\}$. Define the random variable $Y_{\mathcal{G}}$ to be the number of spanning trees in a random $G \in \mathcal{G}_{n,d}$.

Clearly $Y_{\mathcal{G}}$ is identically zero if $n \geq 3$ and $d < 2$. A 2-regular graph has a spanning tree if and only if it is connected (that is, forms a Hamilton cycle), in which case it has exactly n spanning trees. Hence the distribution of $Y_{\mathcal{G}}$ can be inferred from [15, Equation (11)]. For the remainder of the paper we assume that $d \geq 3$.

Our first result gives an asymptotic expression for the expectation of $Y_{\mathcal{G}}$.

Theorem 1.1. *Let $d \geq 3$ be a fixed integer. Then*

$$\mathbb{E}Y_{\mathcal{G}} \sim \exp\left(\frac{6d^2 - 14d + 7}{4(d-1)^2}\right) \frac{(d-1)^{1/2}}{n(d-2)^{3/2}} \left(\frac{(d-1)^{d-1}}{(d^2-2d)^{d/2-1}}\right)^n.$$

This theorem is proved at the end of Section 3.

Next, for fixed $d \geq 3$ and for each positive integer j, define

$$\lambda_j(d) = \frac{(d-1)^j}{2j}, \qquad \zeta_j(d) = -\frac{2(d-1)^j - 1}{(d-1)^{2j}} \tag{3}$$

Our second theorem gives the asymptotic distribution of the number of spanning trees in the case of cubic graphs.

Theorem 1.2. *Let $Z_j \sim \text{Poisson}(\lambda_j(3))$, with each Z_j independent. Consider the number of spanning trees in a random cubic graph, normalized by the expectation given in Theorem 1.1 for $d = 3$. The asymptotic distribution of this quantity is given by*

$$\frac{Y_{\mathcal{G}}}{\mathbb{E}Y_{\mathcal{G}}} \rightsquigarrow \prod_{j=3}^{\infty} (1 + \zeta_j(3))^{Z_j} \, e^{-\lambda_j(3)\zeta_j(3)}.$$

This theorem is proved in Section 4. We conjecture that an analogous result holds for arbitrary (fixed) $d \geq 3$.

Conjecture 1.3. *Let $d \geq 3$ be fixed. Then*

$$\frac{Y_{\mathcal{G}}}{\mathbb{E}Y_{\mathcal{G}}} \rightsquigarrow \prod_{j=3}^{\infty} (1 + \zeta_j(d))^{Z_j} \, e^{-\lambda_j(d)\zeta_j(d)},$$

where $Z_j \sim \text{Poisson}(\lambda_j(d))$ and each Z_j is independent.

We present numerical evidence which supports this conjecture in Section 4.1.

1.2 Plan of attack

From now on, we omit explicit mention of d in the constants $\zeta_j = \zeta_j(d)$ and $\lambda_j = \lambda_j(d)$ from (3).

As is standard in this area, most of our calculations will be performed in the uniform probability space $\mathcal{P}_{n,d}$ of *pairings* (also called the *configuration model* [1, 9, 15]. Let d and n be positive integers such that dn is even. Consider a set of dn *prevertices* distributed evenly into n sets, called *buckets*. (We prefer the terminology "prevertices" to "points".) A *pairing* is a partition of the prevertices into $dn/2$ sets of size 2, called *pairs*. Then

$$|\mathcal{P}_{n,d}| = \#P(dn) = \frac{(dn)!}{(dn/2)!\, 2^{dn/2}} \sim \sqrt{2}\left(\frac{dn}{e}\right)^{dn/2}, \tag{4}$$

using Stirling's formula.

By contracting the prevertices in each bucket to a vertex, each pairing projects to a labelled d-regular multigraph, with loops permitted. Let $\Omega_{n,d}$ denote the set of such multigraphs, and denote the projection of a pairing P by $G(P)$. (We will occasionally informally refer to "partial" pairings, where only a subset of the prevertices are paired. The projection of a partial pairing is defined in the same way.)

Each $G \in \mathcal{G}_{n,d}$ is the projection of $(d!)^n$ different pairings (permuting the prevertices in each bucket), so we can recover the uniform model $\mathcal{G}_{n,d}$ from $\mathcal{P}_{n,d}$ by conditioning on the event that the projected multigraph of a random pairing is simple.

We will apply the small subgraph conditioning method in the form given by Janson [4, Theorem 1].

Theorem 1.4. *Let $\lambda_j > 0$ and $\zeta_j \geq -1$, $j = 1, 2, \ldots$, be constants. Suppose that for each n we have a sequence $\boldsymbol{X} = (X_1, X_2, X_3, \ldots)$ of non-negative integer valued random variables and a random variable Y with $\mathbb{E}Y \neq 0$ (at least for large n). Further suppose the following conditions are satisfied:*

(A1) *For $m \geq 1$, $(X_1, \ldots, X_m) \rightsquigarrow (Z_1, \ldots, Z_m)$, where Z_j are independent Poisson random variables with means λ_j;*

(A2) *For any $m \geq 0$, $\rho \in \mathbb{N}^m$,*

$$\frac{\mathbb{E}[Y|X_1 = \rho_1, \ldots, X_m = \rho_m]}{\mathbb{E}Y} \longrightarrow \prod_{j=1}^{m}(1+\zeta_j)^{\rho_j} e^{-\lambda_j \zeta_j};$$

(A3) $\displaystyle\sum_{j=1}^{\infty} \lambda_j \zeta_j^2 < \infty;$

(A4) $\displaystyle\frac{\mathbb{E}Y^2}{(\mathbb{E}Y)^2} \to \exp\left(\sum_{j=1}^{\infty} \lambda_j \zeta_j^2\right).$

Then

$$\frac{Y}{\mathbb{E}Y} \rightsquigarrow W := \prod_{j=1}^{\infty} (1+\zeta_j)^{Z_j} e^{-\lambda_j \zeta_j}. \tag{5}$$

Moreover, this and the convergence in (A1) *hold jointly.*

We also need a related lemma:

Lemma 1.5. ([4, Lemma 1]) *Let* $\lambda_j' > 0$, $j = 1, 2, \ldots$, *be constants. Suppose that* (A1) *holds and that* $Y \geq 0$. *Suppose:*

(A2') *For any* $m \geq 0$, $\rho \in \mathbb{N}^m$,

$$\frac{\mathbb{E}\left[Y \prod_{j=1}^{m} (X_{j,n})_{\rho_j}\right]}{\mathbb{E}Y} \longrightarrow \prod_{j=1}^{m} (\lambda_j')^{\rho_j}.$$

Then (A2) *holds with* $\lambda_j(1+\zeta_j) = \lambda_j'$ *for all positive integers* j.

We now define the random variables X_j and Y to which these results will be applied.

For each $j \geq 1$, let $\gamma_j : \Omega_{n,d} \to \mathbb{N}$ give the number of cycles of length j in a multigraph. (A loop is a 1-cycle, and a pair of edges on the same two vertices is a 2-cycle.) Then the random variable $X_j = \gamma_j \circ G$ is the number of j-cycles in the projection of a random pairing $P \in \mathcal{P}_{n,d}$. Write $\boldsymbol{X} = (X_j)_{j \geq 1}$ for the sequence of all cycle counts. It is well known [1] that for any positive integer m, the random variables $X_1, \ldots, X_m$ are asymptotically independent Poisson random variables, and that the mean of X_j tends to the quantity $\lambda_j = \lambda_j(d)$ given in (3). Hence Condition (A1) of Theorem 1.4 holds.

Let $\tau : \Omega_{n,d} \to \mathbb{N}$ be the function which counts spanning trees in d-regular multigraphs. Define $Y_{\mathcal{G}}$ as the restriction of τ to $\mathcal{G}_{n,d}$, and define $Y = \tau \circ G$. Then, $Y_{\mathcal{G}}$ is the number of spanning trees in a random $G \in \mathcal{G}_{n,d}$, as in Section 1.1. Y is accordingly the number of spanning trees in the projection of a random pairing $P \in \mathcal{P}_{n,d}$. We will investigate the asymptotic distribution of $Y_{\mathcal{G}}$ through analysis of Y.

In Section 2 we obtain an asymptotic formula for the expected value of Y. In Section 3 we analyse the interaction of the number of spanning trees with short cycles, establishing that (A2) holds for λ_j and ζ_j as given in (3). This enables us to prove Theorem 1.1 and to prove that (A3) holds. Then in Section 4 we investigate the second moment of Y. We can prove that Condition (A4) holds when $d = 3$, leading to a proof of Theorem 1.2. Using our partial calculations for general degrees, we provide numerical evidence that strongly supports Conjecture 1.3.

2 Expected number of spanning trees

In this section we compute $\mathbb{E}Y$. Let $\mathcal{T}_n$ denote the set of labelled trees on n vertices, so that $|\mathcal{T}_n| = n^{n-2}$ by Cayley's formula (see for example, [11]).

Recalling the definition of Y, we have

$$\mathbb{E}Y = \sum_{P \in \mathcal{P}_{n,d}} \frac{1}{|\mathcal{P}_{n,d}|} \tau(G(P)),$$

and hence

$$|\mathcal{P}_{n,d}| \, \mathbb{E}Y = \sum_{P \in \mathcal{P}_{n,d}} \sum_{T \in \mathcal{T}_n} M_{T,P}, \tag{6}$$

where $M_{T,P}$ is the number of ways to embed T into the multigraph $G(P)$. (When $G(P)$ is simple, $M_{T,P}$ is zero or one).

Now, we want to condition on the degree of each of the n vertices in T. Define the set of possible degree sequences

$$\mathcal{D}_n = \left\{ \delta \in \mathbb{N}^n : \quad \sum_{j=1}^n \delta_j = 2(n-1) \right\}.$$

We can decompose $|\mathcal{P}_{n,d}| \, \mathbb{E}Y$ as

$$\sum_{\delta \in \mathcal{D}_n} \sum_{T \in \mathcal{T}_n} \mathbf{1}(T \sim \delta) \sum_{P \in \mathcal{P}_{n,d}} M_{T,P}, \tag{7}$$

where $T \sim \delta$ denotes the event that vertex j has degree δ_j in T, for all $j = 1, \ldots n$.

To evaluate the innermost sum in (7), fix some $\delta \in \mathcal{D}_n$ and some $T \in \mathcal{T}_n$ with $T \sim \delta$. We need to count the number of pairings that include T, with the embedding of T identified. That is, if for some pairing P, the tree T can be embedded in $G(P)$ in multiple ways, then we count each different way separately.

Now, exactly δ_j of the prevertices in bucket j must contribute to T, and there are $(d)_{\delta_j}$ ways to choose and order these prevertices. So, there are

$$\prod_{j=1}^n (d)_{\delta_j}$$

ways to pair up the $n-1$ edges corresponding to a copy of T. Then, there are

$$dn - 2(n-1) = (d-2)n + 2$$

prevertices remaining, which can be paired in $\#P((d-2)n + 2)$ ways. This yields

$$|\mathcal{P}_{n,d}| \, \mathbb{E}Y = \#P((d-2)n + 2) \sum_{\delta \in \mathcal{D}_n} \prod_{j=1}^n (d)_{\delta_j} \sum_{T \in \mathcal{T}_n} \mathbf{1}(T \sim \delta). \tag{8}$$

The inner sum in (8) is the number of trees with degree sequence δ, which is the multinomial

$$\binom{n-2}{\delta_1 - 1, \ldots, \delta_n - 1} = \frac{(n-2)!}{\prod_{j=1}^n (\delta_j - 1)!}. \tag{9}$$

6

(See for example Moon [11, Theorem 3.1].)

Hence

$$|\mathcal{P}_{n,d}|\,\mathbb{E}Y = (n-2)!\,\#P(nd-2(n-1))\sum_{\delta\in\mathcal{D}_n}\prod_{j=1}^{n}\frac{(d)_{\delta_j}}{(\delta_j-1)!}.$$

It follows that the total number of ways to choose a spanning tree on n vertices and choose a partial pairing that projects to that tree is

$$
\begin{aligned}
(n-2)!\sum_{\delta\in\mathcal{D}_n}\prod_{j=1}^{n}\frac{(d)_{\delta_j}}{(\delta_j-1)!} &= (n-2)!\left[x^{2(n-1)}\right]\left(\sum_{j=1}^{\infty}\frac{(d)_j}{(j-1)!}\,x^j\right)^n \\
&= (n-2)!\left[x^{2(n-1)}\right]\left(dx(1+x)^{d-1}\right)^n \\
&= (n-2)!\,d^n\binom{(d-1)n}{n-2}.
\end{aligned}
\tag{10}
$$

Hence, by Stirling's approximation and (4) we conclude that

$$
\begin{aligned}
|\mathcal{P}_{n,d}|\,\mathbb{E}Y &= (n-2)!\,\#P(nd-2(n-1))\,d^n\binom{(d-1)n}{n-2} \\
&\sim \frac{\sqrt{2}(d-1)^{1/2}}{n(d-2)^{3/2}}\left(d(d-2)(d-1)^{d-1}\left(\frac{n}{(d-2)e}\right)^{d/2}\right)^n.
\end{aligned}
\tag{11}
$$

It follows that

$$
\begin{aligned}
\mathbb{E}Y &= \frac{(n-2)!\,\#P(nd-2(n-1))}{\#P(nd)}\,d^n\binom{(d-1)n}{n-2} \\
&\sim \frac{(d-1)^{1/2}}{n(d-2)^{3/2}}\left(\frac{(d-1)^{d-1}}{(d^2-2d)^{d/2-1}}\right)^n.
\end{aligned}
\tag{12}
$$

Hence for $d\ge 3$ and n sufficiently large, we have $\mathbb{E}Y\ne 0$.

3   Interaction with short cycles

Recall that $X_j = \gamma_j\circ G$ is the number of cycles of length j in the projection of a random pairing $P\in\mathcal{P}_{n,d}$. For some fixed $m\ge 0$, $\rho\in\mathbb{N}^m$, let $X_\rho = \prod_{j=1}^{m}(X_j)_{\rho_j}$. In this section we will compute an asymptotic formula for $\mathbb{E}[YX_\rho]/\mathbb{E}Y$, in the form required by Condition (A2′).

We have

$$\mathbb{E}[YX_\rho] = \sum_{P\in\mathcal{P}_{n,d}}\frac{1}{|\mathcal{P}_{n,d}|}\,\tau(G(P))\prod_{j=1}^{m}(\gamma_j(G(P)))_{\rho_j}.$$

Note that

$$\prod_{j=1}^{m}(\gamma_j(G(P)))_{\rho_j}$$

7

is the number of ways to choose, for each $j \in \{1, \ldots, m\}$, an ordered set of ρ_j cycles of length j. This will result in an ordered set of

$$|\rho| = \sum_{j=1}^{m} \rho_j$$

cycles.

We make the decomposition

$$\prod_{j=1}^{m} (\gamma_j(G(P)))_{\rho_j} = \gamma_\rho^{(0)} + \gamma_\rho',$$

where $\gamma_\rho^{(0)}$ is the number of ordered sets of cycles in which each cycle is disjoint, and γ_ρ' is the number of ordered sets in which some vertices are shared between multiple cycles. We can further decompose γ_ρ' by the structure of the interaction between the cycles. That is, according to the multigraph that is the union of the cycles, and the specification of which edges of this union belong to which cycle. This expresses γ_ρ' as a sum of terms $\gamma_\rho^{(j)}$. The number of terms J in this decomposition depends on ρ, but is $O(1)$ as $n \to \infty$.

Define

$$E^{(j)} = \sum_{P \in \mathcal{P}_{n,d}} \tau(G(P)) \, \gamma_\rho^{(j)},$$

so that we have $|\mathcal{P}_{n,d}| \, \mathbb{E}[Y X_\rho] = \sum_{j=1}^{J} E^{(j)}$.

We proceed to calculate $E^{(0)}$. As is standard when applying this method (see for example, [5, Theorem 9.6]), $\mathbb{E}[Y X_\rho]$ is asymptotically dominated by $E^{(0)}$ (the contribution due to disjoint cycles). See Lemma 3.1 for some justification for this fact.

Let $\mathcal{C}_{n,j}$ be the set of all j-cycles on the vertex set $\{1, \ldots, n\}$, and define the Cartesian product

$$\mathcal{C}_{n,\rho} = \prod_{j=1}^{m} \mathcal{C}_{n,j}^{\rho_j}.$$

Each $R \in \mathcal{C}_{n,\rho}$ is an ordered set of $|\rho|$ cycles. We use the notation $R_{j,k}$ for the kth cycle of length j in R. Next, define

$$\mathcal{R}_{n,\rho} = \{R \in \mathcal{C}_{n,\rho} : \text{ the cycles in } R \text{ are pairwise disjoint}\}.$$

Similarly to (6), we have

$$E^{(0)} = \sum_{R \in \mathcal{R}_{n,\rho}} \sum_{T \in \mathcal{T}_n} \sum_{P \in \mathcal{P}_{n,d}} M_{(T,R),P},$$

where $M_{(T,R),P}$ is the number of ways to embed the tree T and the cycles in R into the multigraph $G(P)$.

8

We further condition on the edge intersection between the embedding of T and the cycles in R. We use a binary sequence of length j to encode the intersection of an j-cycle with a spanning tree. Picking an arbitrary start vertex and direction for the cycle, if the kth edge of the cycle is to be included in the intersection, then the kth element of the corresponding sequence is one; otherwise it is zero. All sequences $q \in \{0,1\}^j$ represent possible intersections, except the sequence $(1, \ldots, 1)$, because a tree contains no cycles.

Define the set of all possible intersection sequences for a cycle of length j, by

$$\mathcal{I}_j = \{0,1\}^j \setminus (1, \ldots, 1).$$

Also, define the Cartesian product

$$\mathcal{Q}_\rho = \prod_{j=1}^{m} \mathcal{I}_j^{\rho_j}.$$

So, for each $R \in \mathcal{R}_{n,\rho}$, specifying some $Q \in \mathcal{Q}_\rho$ fully specifies the intersection between the cycles in R and a tree T.

We have

$$E^{(0)} = \sum_{Q \in \mathcal{Q}_\rho} \sum_{R \in \mathcal{R}_{n,\rho}} \sum_{T \in \mathcal{T}_n} \sum_{P \in \mathcal{P}_{n,d}} M_{(T,R,Q),P}, \tag{13}$$

where $M_{(T,R,Q),P}$ is the number of ways to embed T and the cycles in R in P, such that the intersection between the embedding of T and the cycle $R_{j,k}$ is consistent with $Q_{j,k}$, for $j = 1, \ldots, m$ and $k = 1, \ldots, \rho_j$.

Fixing $Q \in \mathcal{Q}_\rho$, we will now evaluate the innermost triple sum in (13). Consider the following process:

1. Choose some $R \in \mathcal{R}_{n,\rho}$.

2. Choose a partial pairing that projects to R.

3. Extend this to a pairing of a spanning tree consistent with Q.

4. Pair the remaining prevertices arbitrarily.

We will find that the number of ways to complete each step is independent of the other steps. Then, $E^{(0)}$ is a product of the number of ways to complete each step, summed over all $Q \in \mathcal{Q}_\rho$.

First let $n_\rho = \sum_{j=1}^{m} j\rho_j$ be the total number of vertices in each $R \in \mathcal{R}_{n,\rho}$. The number of ways to choose the vertices for some $R \in \mathcal{R}_{n,\rho}$ is

$$\binom{n}{n_\rho}$$

and the number of different arrangements of disjoint cycles on those vertices is

$$\frac{n_\rho!}{\prod_{j=1}^{m}(j\chi(j))^{\rho_j}},$$

9

where $j\chi(j)$ is the size of the automorphism group of a j-cycle:

$$\chi(j) = \begin{cases} 1 & \text{for } j \le 2 \\ 2 & \text{for } j > 2. \end{cases}$$

That is, the number of ways to complete Step 1 is

$$s_1^{(0)} = \frac{n!}{(n - n_\rho)! \prod_{j=1}^{m} (j\chi(j))^{\rho_j}}.$$

Next, the number of ways to complete Step 2 is

$$s_2^{(0)} = \prod_{j=1}^{m} \left(\frac{\chi(j)(d(d-1))^j}{2} \right)^{\rho_j}.$$

Note for future reference that we have

$$s_1^{(0)} s_2^{(0)} \sim n^{n_\rho} \prod_{j=1}^{m} \left(\frac{(d(d-1))^j}{2j} \right)^{\rho_j}. \tag{14}$$

Next, we count the number of ways to extend this pairing to a tree T consistent with Q. We do this by constructing a new irregular pairing model $\mathcal{P}_{n,\rho}$ from the prevertices still unpaired after Step 2. Recall that Q describes a union of disjoint paths; for each of these paths, combine the unpaired prevertices remaining in each constituent vertex of the path to form a *super-bucket*. If the path has k vertices then the resulting super-bucket has $k(d-2)$ prevertices. Let $|Q|$ be the number of super-buckets formed in this way, so the total number of buckets in $\mathcal{P}_{n,\rho}$ is $n' := n - n_\rho + |Q|$.

Now, consider an extension of a pairing of cycles from Step 2, to a (partial) pairing of the edges of a tree T consistent with Q, as per Step 3. The pairs from this extension correspond uniquely to a (partial) pairing P' in the pairing model $\mathcal{P}_{n,\rho}$. By the construction of $\mathcal{P}_{n,\rho}$, the projection $T' = G(P')$ of this pairing is simply T with some subpaths contracted to single vertices. Since contracting edges of a tree cannot create cycles, $G(P')$ is itself a (spanning) tree. Similarly, every pairing of a tree in $\mathcal{P}_{n,\rho}$ corresponds to an extension of a pairing of cycles to a pairing of a tree in $\mathcal{P}_{n,d}$ consistent with Q. So the number of ways $s_3^{(0)}$ to complete Step 3 equals the number of ways to choose and pair up a spanning tree in $\mathcal{P}_{n,\rho}$.

We will perform this count as in Section 2, by conditioning on the degree in T' of each bucket in $\mathcal{P}_{n,\rho}$. Put an arbitrary ordering on the $|Q|$ super-buckets, and let d_j be the number of prevertices in the jth super-bucket. For a degree sequence δ, let $|\delta|$ be its degree sum. Define the sets

$$\mathcal{D}_Q = \left\{ \delta' \in \mathbb{N}^{|Q|} : \quad \delta'_j \le d_j \text{ for all } j \right\},$$

$$\mathcal{D}_{\delta'} = \left\{ \delta \in \mathbb{N}^{n-n_\rho} : \quad \sum_{j=1}^{n-n_\rho} \delta_j = 2(n'-1) - |\delta'| \right\}.$$

10

The set $\mathcal{D}_Q$ contains all possible degree-in-T' sequences for the $|Q|$ super-buckets. For some $\delta' \in \mathcal{D}_Q$, the set $\mathcal{D}_{\delta'}$ contains all possible degree sequences for the $n - n_\rho$ remaining ordinary buckets. So, we have

$$s_3^{(0)} = \sum_{\delta' \in \mathcal{D}_Q} \sum_{\delta \in \mathcal{D}_{\delta'}} \sum_{T' \in \mathcal{T}_{n'}} \mathbf{1}(T' \sim (\delta', \delta)) \sum_{P' \in \mathcal{P}_{n,\rho}} M_{T',P},$$

where $T' \sim (\delta', \delta)$ denotes the event that the super-buckets have degree-in-T' sequence δ' and the remaining vertices have degree-in-T' sequence δ. Proceeding as before, after fixing some (T', δ', δ), there are

$$\left(\prod_{j=1}^{n-n_\rho} (d)_{\delta_j} \right) \left(\prod_{j=1}^{|Q|} (d_j)_{\delta'_j} \right)$$

ways to pair the edges of T'. There are

$$\frac{(n' - 2)!}{\left(\prod_{j=1}^{n-n_\rho} (\delta_j - 1)! \right) \left(\prod_{j=1}^{|Q|} (\delta'_j - 1)! \right)}$$

trees with $T' \sim (\delta', \delta)$, by (9). So, we have

$$s_3^{(0)} = \sum_{\delta' \in \mathcal{D}_Q} A_{\delta'} \prod_{j=1}^{|Q|} \frac{(d_j)_{\delta'_j}}{(\delta'_j - 1)!}, \tag{15}$$

where

$$
\begin{aligned}
A_{\delta'} &= (n' - 2)! \sum_{\delta \in \mathcal{D}_{\delta'}} \prod_{j=1}^{n-n_\rho} \frac{(d)_{\delta_j}}{(\delta_j - 1)!} \\
&= (n' - 2)! \left[x^{2(n'-1)-|\delta'|} \right] \left(dx(1 + x)^{d-1} \right)^{n-n_\rho} \\
&= (n' - 2)! \, d^{n-n_\rho} \binom{(d-1)(n - n_\rho)}{2(n' - 1) - |\delta'| - (n - n_\rho)} \\
&\sim \frac{(d-2)^{2|Q|-|\delta'|-5/2}(d-1)^{1/2}}{n^{n_\rho - |Q| + 2}} \left(\frac{(d-2)^{d-2}}{d(d-1)^{d-1}} \right)^{n_\rho} \left(\frac{d(d-1)^{d-1}n}{e(d-2)^{d-2}} \right)^n. \tag{16}
\end{aligned}
$$

Finally, for Step 4 there are $dn - 2n_\rho - 2(n' - 1) = (d-2)n - 2(|Q| - 1)$ prevertices remaining, which can be paired in

$$s_4^{(0)} = \#P((d-2)n - 2(|Q| - 1)) \sim \sqrt{2} \left(\frac{(d-2)n}{e} \right)^{(d/2-1)n} ((d-2)n)^{1-|Q|} \tag{17}$$

ways.

Combining (11), (14), (15), (16) and (17), we have

$$\frac{E^{(0)}}{|\mathcal{P}_{n,d}|\,\mathbb{E}Y} = \sum_{Q\in\mathcal{Q}_\rho} \frac{s_1^{(0)}s_2^{(0)}s_3^{(0)}s_4^{(0)}}{|\mathcal{P}_{n,d}|\,\mathbb{E}Y}$$

$$\to \left(\prod_{j=1}^{m}\left(\frac{(d(d-1))^j}{2j}\right)^{\rho_j}\right)$$

$$\times \sum_{Q\in\mathcal{Q}_\rho}\sum_{\delta'\in\mathcal{D}_Q}(d-2)^{|Q|-|\delta'|}\left(\frac{(d-2)^{d-2}}{d(d-1)^{d-1}}\right)^{n_\rho}\prod_{j=1}^{|Q|}\frac{(d_j)_{\delta'_j}}{(\delta'_j-1)!}. \qquad (18)$$

As is standard in these arguments, the only significant contribution to $\mathbb{E}[YX_\rho]$ comes from $E^{(0)}$, where the cycles do not overlap. For completeness we sketch a proof of this below.

Lemma 3.1. $\mathbb{E}[YX_\rho]$ *is dominated by the contribution from $E^{(0)}$. That is,*

$$\frac{\mathbb{E}[YX_\rho]}{\mathbb{E}Y} \sim \frac{E^{(0)}}{|\mathcal{P}_{n,d}|\,\mathbb{E}Y}.$$

Proof. We can estimate general $E^{(j)}$ with some slight modifications to the above calculations. We would need a different $\mathcal{R}'_{n,\rho} \subseteq \mathcal{C}_{n,\rho}$ that contains all possible ways to embed an ordered set of cycles with a particular union U into the vertex set $\{1,\ldots,n\}$. The intersection between a spanning tree and the cycles in some $R \in \mathcal{R}_{n,\rho}$ would then be a subforest of U, so it would be more complicated to explicitly define a set $\mathcal{Q}_\rho$ that encodes all possibilities for the intersection. However, the number of possible intersections is still independent of n.

Using the same 4 steps, the decomposition $E^{(j)} = s_1^{(j)}s_2^{(j)}s_3^{(j)}s_4^{(j)}$ is still valid. Let $|U|$ and $\|U\|$ be the number of vertices and edges in the multigraph U, respectively. Carefully adjusting the calculations for $E^{(0)}$, we have $s_1^{(j)}/s_1^{(0)} = O(n^{|U|-n_\rho})$, $s_2^{(j)}/s_2^{(0)} = O(1)$ and $s_3^{(j)}/s_3^{(0)} = O(1)$. For Step 4 there would be $dn - 2\|U\| - 2(n'-1)$ prevertices remaining, so $s_4^{(j)}/s_4^{(0)} = O(n^{n_\rho-\|U\|})$.

We conclude that $E^{(j)}/E^{(0)} = O(n^{|U|-\|U\|})$. Any non-disjoint union between distinct cycles has more edges than vertices, so the lemma is proved. $\qquad \square$

We now want to to express (18) in the form required by (A2'). So, we consider each cycle independently. For a sequence $q \in \mathcal{I}_j$, let $q[k]$ be the number of paths with k vertices in the intersection encoded by q. So, for $Q \in \mathcal{Q}_\rho$ we have

$$|Q| = \sum_{j=1}^{m}\sum_{\ell=1}^{\rho_j}\sum_{k=1}^{j}Q_{j,\ell}[k].$$

Also, let $|q| = \sum_{k=1}^{j}q[k]$ be the total number of paths in the intersection encoded by q.

Now, recall that if the jth super-bucket was collapsed from a path of length k, then $d_j = k(d - 2)$. Also recall that $n_\rho = \sum_{j=1}^{m} j\rho_j$ and note that

$$(d-2)^{|Q|-|\delta'|} = \prod_{j=1}^{|Q|} (d-2)^{1-\delta'_j}.$$

As a result, we have

$$\frac{\mathbb{E}[Y X_\rho]}{\mathbb{E}Y} \to \prod_{j=1}^{m} (\lambda'_j)^{\rho_j},$$

where

$$\lambda'_j = \frac{(d(d-1))^j}{2j} \left(\frac{(d-2)^{d-2}}{d(d-1)^{d-1}} \right)^j \sum_{q \in \mathcal{I}_j} \prod_{k=1}^{j} \left(\sum_{\ell=1}^{k(d-2)} \frac{(k(d-2))_\ell}{(\ell-1)!\,(d-2)^{\ell-1}} \right)^{q[k]}.$$

Note that ℓ takes the role of δ'_j for the jth super-bucket. We have proved that Condition (A2′) is satisfied. It remains to simplify our expression for λ'_j. We have

$$\lambda'_j = \frac{1}{2j} \left(\frac{d-2}{d-1} \right)^{j(d-2)} \sum_{q \in \mathcal{I}_j} \prod_{k=1}^{j} \left(k(d-2) \sum_{\ell=0}^{k(d-2)-1} \binom{k(d-2)-1}{\ell} (d-2)^{-\ell} \right)^{q[k]}$$

$$= \frac{1}{2j} \left(\frac{d-2}{d-1} \right)^{j(d-2)} \sum_{q \in \mathcal{I}_j} \prod_{k=1}^{j} \left(k(d-2) \left(\frac{1}{d-2} + 1 \right)^{k(d-2)-1} \right)^{q[k]}$$

$$= \frac{1}{2j} \sum_{q \in \mathcal{I}_j} \left(\frac{(d-2)^2}{d-1} \right)^{|q|} \prod_{k=1}^{j} k^{q[k]}.$$

Now, recall that $(1,\ldots,1) \notin \mathcal{I}_j$. So, to evaluate the sum over q, we may identify a particular element in the sequence to be zero. By symmetry, we arbitrarily choose the last. We also condition on $|q|$: define

$$\mu = \frac{(d-2)^2}{d-1}, \qquad \Lambda_{j,t} = \sum_{\substack{q \in \mathcal{I}_j \\ |q|=t \\ q_j = 0}} \mu^t \prod_{k=1}^{j} k^{q[k]}. \tag{19}$$

Note that $\Lambda_{j,1} = j\mu$ for all j, because the only sequence $q \in \mathcal{I}_j$ with $|q| = 1$ and $q_j = 0$ is $(1,\ldots,1,0)$. For $|q| > 1$, the first path in (the intersection encoded by) a sequence can contain anywhere between 1 and $j-1$ vertices. Ranging over the possibilities, we have

$$\Lambda_{j,t} = \sum_{k=1}^{j-1} k\mu \Lambda_{j-k,t-1}$$

13

for $t > 0$. To solve this recurrence, define the generating function

$$\Lambda(x, y) = \sum_{j=1}^{\infty} \sum_{t=1}^{\infty} \Lambda_{j,t} x^j y^t.$$

We have

$$\Lambda(x, y) - \sum_{j=1}^{\infty} \Lambda_{j,1} x^j y = \sum_{j=1}^{\infty} \sum_{t=2}^{\infty} \sum_{k=1}^{j-1} k\mu\Lambda_{j-k,t-1} x^j y^t,$$

so

$$\Lambda(x, y) = \sum_{k=1}^{\infty} kx^k y\mu \sum_{j=k+1}^{\infty} \sum_{t'=1}^{\infty} \Lambda_{j-k,t'} x^{j-k} y^{t'} + \sum_{j=1}^{\infty} jx^j y\mu$$

$$= (\Lambda(x, y) + 1) \sum_{k=1}^{\infty} kx^k y\mu.$$

Now, defining

$$g(x) = \sum_{k=1}^{\infty} kx^k \mu = \frac{x\mu}{(1-x)^2},$$

we have

$$\Lambda(x, y) = \frac{g(x)y}{1 - g(x)y}.$$

If $q \in \mathcal{I}_j$, then there are j positions to place a zero, and if $|q| = t$ then there are t zeros in q. Hence

$$\lambda'_j = \frac{1}{2j} \sum_{t=1}^{j} \frac{j\Lambda_{j,t}}{t}$$

$$= \frac{1}{2} [x^j] \sum_{t=1}^{j} \frac{1}{t} [y^{t-1}] \frac{g(x)}{1 - g(x)y}$$

$$= \frac{1}{2} [x^j] \int_0^1 \frac{g(x)}{1 - g(x)y} dy$$

$$= -\frac{1}{2} [x^j] \log(1 - g(x)).$$

Now, defining $\kappa = \sqrt{d-1}$ we have

$$1 - g(x) = \frac{1 + x^2 - (2 + \mu)x}{(1-x)^2} = \frac{(1 - \kappa^2 x)(1 - \kappa^{-2} x)}{(1-x)^2},$$

14

so

$$\lambda_j' = \frac{1}{2}[x^j]\left(2\log(1-x) - \log\left(1-\kappa^2 x\right) - \log\left(1-\kappa^{-2}x\right)\right)$$

$$= \frac{1}{2}[x^j]\sum_{k=1}^{\infty}\frac{-2x^k + \left(\kappa^2 x\right)^k + \left(\kappa^{-2}x\right)^k}{k}$$

$$= \frac{1}{2j}\left(\kappa^j - \kappa^{-j}\right)^2$$

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

for $j \geq 1$.

To complete this section we will establish that conditions (A2) and (A3) of Theorem 1.4 hold, and prove Theorem 1.1.

Lemma 3.2. *Let $d \geq 3$ be a fixed integer. Then Conditions (A2) and (A3) of Theorem 1.4 are satisfied, and*

$$\exp\left(\sum_{j=1}^{\infty}\lambda_j\zeta_j^2\right) = \frac{d^2}{\sqrt{(d-1)(d-2)(d^2-d+1)}}.$$

Proof. The calculations of this section show that Condition (A2$'$) of Lemma 1.5 is satisfied with λ_j' given by (20). Then Lemma 1.5 guarantees that (A2) is satisfied. Using the Taylor expansion of $\log(1-z)$, it follows from (3) that

$$\sum_{j=1}^{\infty}\lambda_j\zeta_j^2$$

$$= \sum_{j=1}^{\infty}\frac{1}{2j}\left(4(d-1)^{-j} - 4(d-1)^{-2j} + (d-1)^{-3j}\right)$$

$$= \frac{1}{2}\left(-4\log\left(1-(d-1)^{-1}\right) + 4\log\left(1-(d-1)^{-2}\right) - \log\left(1-(d-1)^{-3}\right)\right).$$

Taking the exponential of both sides and rearranging establishes the stated expression for $\exp\left(\sum_{j=1}^{\infty}\lambda_j\zeta_j^2\right)$, which is finite for $d \geq 3$. Hence Condition (A3) holds, as required. $\square$

Proof of Theorem 1.1. We know that condition (A2) holds (as proved above), and hence

$$\mathbb{E}Y_{\mathcal{G}} = \mathbb{E}[Y|X_1 = X_2 = 0] \to \mathbb{E}Y\,\exp(-\lambda_1\zeta_1 - \lambda_2\zeta_2).$$

Substituting using (3) and (12) completes the proof. $\square$

15

4 The second moment

We now want to calculate $\mathbb{E}Y^2$. As a first step we transform this problem into one of evaluating the coefficient of a certain generating function.

Lemma 4.1. *Let $d \geq 3$ be fixed and define*

$$N(n,d) = \begin{cases} n & \text{if } d \geq 4, \\ n/2 + 2 & \text{if } d = 3. \end{cases}$$

Then

$$
\begin{aligned}
&|\mathcal{P}_{n,d}| \, \mathbb{E}Y^2 \\
&= \frac{n! \, ((d-2)n)! \, d^n}{2^{(d/2-1)n+2}} \sum_{b=1}^{N(n,d)} \frac{2^b}{b! \, ((d/2-1)n - b + 2)!} \, [z^n] \left(\sum_{j=1}^{\infty} \binom{(d-1)j}{j} z^j \right)^b.
\end{aligned}
$$

Proof. We write

$$|\mathcal{P}_{n,d}| \, \mathbb{E}Y^2 = \sum_{P \in \mathcal{P}_{n,d}} \sum_{T_1 \in \mathcal{T}_n} \sum_{T_2 \in \mathcal{T}_n} M_{(T_1,T_2),P},$$

where $M_{(T_1,T_2),P}$ is the number of ways to embed the ordered pair of trees (T_1,T_2) into the multigraph $G(P)$. We will estimate this sum by choosing some $T_1, T_2 \in \mathcal{T}_n$ and counting the ways to pair up their edges, then counting the ways to complete the pairing. We break up this process in a similar way to Section 3:

1. Choose $b \in \{1, \ldots, n\}$, which will be the number of connected components in the intersection of the embeddings of T_1 and T_2. (As we will see later, when $d = 3$ we must restrict to $b \in \{1, \ldots, n/2 + 2\}$.)

2. Choose a partition $(\nu_1, \ldots, \nu_b)$ of n into positive parts. That is, ν_j is a positive integer for $j = 1, \ldots, b$ and $\sum_{j=1}^{b} \nu_j = n$. Here ν_j will be the number of vertices in the jth connected component of the interesection. (We should divide by $b!$ to account for our assumption that the connected components are labelled).

3. Choose a partition of the n vertices into b groups, where the size of the jth group is ν_j.

4. In each group, choose a spanning tree on that group and choose a partial pairing that projects to that tree. This specifies a component of the intersection.

   Now, collapse the buckets in each group into a single super-bucket, giving exactly b super-buckets. The jth super-bucket has $d\nu_j - 2(\nu_j - 1)$ unpaired prevertices. We now want to pair up two pair-disjoint spanning trees T_1', T_2' in the collapsed pairing model. These will extend to T_1 and T_2 using the intersection subtrees chosen in Step 4.

5. For $j = 1, \ldots, b$, choose δ_j^1 and δ_j^2, the degree of vertex j in T_1' and T_2' respectively, in such a way that $\delta_j^1 + \delta_j^2 \leq d\nu_j - 2(\nu_j - 1)$. These must also satisfy

$$\sum_{j=1}^{b} \delta_j^1 = \sum_{j=1}^{b} \delta_j^2 = 2(b-1),$$

as they are the degree sequence of a spanning tree on b vertices.

6. Choose two trees T_1', T_2' on the b vertices that are consistent with the degree sequences chosen in Step 6.

7. Pair up these two trees in a pair-disjoint way.

8. Pair all remaining prevertices to complete a d-regular pairing.

Given ν, the number of ways to complete Step 3 is

$$s_3 = \binom{n}{\nu_1, \ldots, \nu_b}.$$

By (10), the number of ways to complete Step 4 is

$$s_4 = d^n \prod_{j=1}^{b} (\nu_j - 2)! \binom{(d-1)\nu_j}{\nu_j - 2}.$$

The number of ways to complete Step 6 is

$$s_6 = \binom{b-2}{\delta_1^1 - 1, \ldots, \delta_b^1 - 1} \binom{b-2}{\delta_1^2 - 1, \ldots, \delta_b^2 - 1},$$

by (9), and the number of ways to complete Step 7 is

$$s_7 = \prod_{j=1}^{b} ((d-2)\nu_j + 2)_{\delta_j^1 + \delta_j^2}.$$

Finally, for Step 8 there are

$$\sum_{j=1}^{b} \left(d\nu_j - 2(\nu_j - 1) - \delta_j^1 - \delta_j^2 \right) = (d-2)n - 2(b-2)$$

prevertices remaining, so the number of ways to complete Step 8 is

$$s_8 = \#P((d-2)n - 2(b-2)).$$

For this construction to make sense, the quantity $(d-2)n - 2(b-2)$ must be nonnegative. This is certainly true when $d \geq 4$, but when $d = 3$ this imposes the constraint that $b \leq n/2 + 2$. (This explains the definition of $N(n, d)$ in the lemma statement.)

It will be convenient to work with the nonnegative variables

$$\eta_j^1 = \delta_j^1 - 1, \quad \eta_j^2 = \delta_j^2 - 1 \quad \text{and} \quad \eta_j^3 = (d-2)\nu_j - \eta_j^1 - \eta_j^2 \tag{21}$$

defined for $j = 1, \ldots, b$. Let

$$\mathcal{S}_2(b) = \{\nu \in \{1, \ldots, n\}^b : \sum_{j=1}^b \nu_j = n\}$$

be the set of possible sequences ν from Step 2 and let

$$\mathcal{S}_5(\nu) = \left\{ (\eta^1, \eta^2, \eta^3) \in (\mathbb{N}^b)^3 : \quad \eta_j^1 + \eta_j^2 + \eta_j^3 = (d-2)\nu_j \text{ for } j = 1, \ldots, b, \right.$$
$$\left. \sum_{j=1}^b \eta_j^1 = \sum_{j=1}^b \eta_j^2 = b - 2 \right\}$$

be the set of sequences arising from Step 5 using (21).

Combining all of the above gives

$$|\mathcal{P}_{n,d}| \, \mathbb{E} Y^2 = \sum_{b=1}^{N(n,d)} \frac{1}{b!} \sum_{\nu \in \mathcal{S}_2(b)} s_3 s_4 \sum_{(\eta^1, \eta^2, \eta^3) \in \mathcal{S}_5(\nu)} s_6 s_7 s_8$$

$$= \sum_{b=1}^{N(n,d)} \frac{n! \, d^n}{b! \, ((d/2 - 1)n - b + 2)! \, 2^{(d/2-1)n - b + 2}} \sum_{\nu \in \mathcal{S}_2(b)} \prod_{j=1}^b \frac{((d-1)\nu_j)!}{\nu_j!}$$

$$\times \sum_{(\eta^1, \eta^2, \eta^3) \in \mathcal{S}_5(\nu)} \binom{b-2}{\eta_1^1, \ldots, \eta_b^1} \binom{b-2}{\eta_1^2, \ldots, \eta_b^2} \binom{(d-2)n - 2(b-2)}{\eta_1^3, \ldots, \eta_b^3}.$$

Now

$$\sum_{(\eta^1, \eta^2, \eta^3) \in \mathcal{S}_5(\nu)} \binom{b-2}{\eta_1^1, \ldots, \eta_b^1} \binom{b-2}{\eta_1^2, \ldots, \eta_b^2} \binom{(d-2)n - 2(b-2)}{\eta_1^3, \ldots, \eta_b^3}$$

$$= \sum_{(\eta^1, \eta^2, \eta^3) \in \mathcal{S}_5(\nu)} \left[z_1^{\eta_1^1} \ldots z_b^{\eta_b^1} \right] \left(\sum_{j=1}^b z_j \right)^{b-2}$$

$$\times \left[z_1^{\eta_1^2} \ldots z_b^{\eta_b^2} \right] \left(\sum_{j=1}^b z_j \right)^{b-2} \left[z_1^{\eta_1^3} \ldots z_b^{\eta_b^3} \right] \left(\sum_{j=1}^b z_j \right)^{(d-2)n - 2(b-2)}$$

$$= \left[z_1^{(d-2)\nu_1} z_2^{(d-2)\nu_2} \cdots z_b^{(d-2)\nu_b} \right] \left(\sum_{j=1}^b z_j \right)^{(d-2)n}$$

$$= \binom{(d-2)n}{(d-2)\nu_1, \ldots, (d-2)\nu_b}.$$

It follows that

$$
\begin{aligned}
&|\mathcal{P}_{n,d}| \, \mathbb{E} Y^2 \\
&= \frac{n! \, ((d-2)n)! \, d^n}{2^{(d/2-1)n+2}} \sum_{b=1}^{N(n,d)} \frac{2^b}{b! \, ((d/2-1)n - b + 2)!} \sum_{\nu \in \mathcal{S}_2(b)} \prod_{j=1}^{b} \binom{(d-1)\nu_j}{\nu_j}.
\end{aligned}
$$

This is equal to the expression in the lemma statement, by definition of $\mathcal{S}_2(b)$. $\qquad\square$

We now seek to evaluate

$$
[z^n] \left(\sum_{j=1}^{\infty} \binom{(d-1)j}{j} z^j \right)^b.
$$

By Stirling's approximation and the ratio test, the radius of convergence of the series $\sum_{j=1}^{\infty} \binom{(d-1)j}{j} z^j$ equals $\frac{(d-2)^{d-2}}{(d-1)^{d-1}}$. Hence,

$$
f(z) := \sum_{j=1}^{\infty} \binom{(d-1)j}{j} \left(\frac{(d-2)^{d-2}}{(d-1)^{d-1}} \right)^j z^j
$$

is analytic in the disk $\{z : |z| < 1\}$. Define $\beta = b/n$ and let $r_\beta \in (0,1)$ be fixed for each β (we will determine this later). Then, with the contour $\Gamma : [-\pi, \pi] \to \mathbb{C}$ defined by $\theta \mapsto r_\beta e^{i\theta}$, we have

$$
\begin{aligned}
[z^n] f(z)^b &= \frac{1}{2\pi i} \int_\Gamma \frac{f(z)^b}{z^{n+1}} \, \mathrm{d}z \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{f(r_\beta e^{i\theta})^\beta}{r_\beta e^{i\theta}} \right)^n \mathrm{d}\theta,
\end{aligned}
\tag{22}
$$

by Cauchy's coefficient formula. Let

$$
\mathcal{X}_n = \left\{ \tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{N(n,d)}{n} \right\} \times [-\pi, \pi]
$$

be the sample space for pairs (β, θ), and define $a_n : \mathcal{X}_n \to \mathbb{C}$ by

$$
a_n(\beta, \theta) = \frac{2^{\beta n}}{(\beta n)! \, ((d/2 - 1)n - \beta n + 2)!} \left(\frac{f(r_\beta e^{i\theta})^\beta}{r_\beta e^{i\theta}} \right)^n.
\tag{23}
$$

Finally, let

$$
F_n = \sum_{b=1}^{N(n,d)} \int_{-\pi}^{\pi} a_n(b/n, \theta) \, \mathrm{d}\theta.
\tag{24}
$$

Then, by Lemma 4.1 and (22),

$$
|\mathcal{P}_{n,d}| \, \mathbb{E} Y^2 = \frac{(d-1)^{n(d-1)} n! \, ((d-2)n)! \, d^n}{2\pi (d-2)^{n(d-2)} 2^{(d/2-1)n+2}} F_n.
\tag{25}
$$

19

We now apply the saddle point method to estimate the sum in (24) in the case that $d = 3$. Our proof is adapted from that of [3, Theorem 2.3].

When $d = 3$ the function f satisfies

$$f(z) = \sum_{j=1}^{\infty} \binom{2j}{j} \left(\frac{z}{4}\right)^j = (1-z)^{-1/2} - 1.$$

We note for later that if $\theta \in [-\pi, \pi]$ is nonzero then

$$\left| f\left(r_\beta e^{i\theta}\right) \right| = \left| \sum_{j=1}^{\infty} \binom{2j}{j} \left(\frac{r_\beta}{4}\right)^j e^{ij\theta} \right| < \sum_{j=1}^{\infty} \binom{2j}{j} \left(\frac{r_\beta}{4}\right)^j = |f(r_\beta)|, \qquad (26)$$

using the triangle inequality. Hence for each β the function $\theta \mapsto |f(r_\beta e^{i\theta})|$ on $[-\pi, \pi]$ is uniquely maximised at $\theta = 0$.

Define $\mathcal{X} = (0,1] \times [-\pi, \pi]$ and let $\mathcal{X}^* \subset \mathcal{X}$ be a set (to be determined) such that for $(\beta, \theta) \in \mathcal{X}^*$, both β and $d/2 - 1 - \beta = 1/2 - \beta$ are bounded below by some positive constant. Then Stirling's approximation gives, for $\beta \in \mathcal{X}^* \cap \mathcal{X}_n$,

$$a_n(\beta, \theta)$$

$$\sim \frac{e^{n/2}}{2\pi \, n^{n/2+3} \sqrt{\beta(1/2 - \beta)} \, (1/2 - \beta)^2} \left(\frac{\left(2f\left(r_\beta e^{i\theta}\right)\right)^\beta}{r_\beta e^{i\theta} \beta^\beta (1/2 - \beta)^{(1/2-\beta)}} \right)^n. \qquad (27)$$

Next, define the half-spaces $\mathcal{X}^{1/2} = (0, 1/2] \times [-\pi, \pi]$ and $\bar{\mathcal{X}}^{1/2} = [0, 1/2] \times [-\pi, \pi]$. Define the real-valued sequence $(c_n)_{n \in \mathbb{N}}$ and the functions $\psi : \bar{\mathcal{X}}^{1/2} \to \mathbb{R}$ and $\phi : \mathcal{X}^{1/2} \to \mathbb{C}$ by

$$c_n = \frac{e^{n/2}}{2\pi n^{n/2+3}},$$
$$\psi(\beta, \theta) = \beta^{-1/2}(1/2 - \beta)^{-5/2},$$
$$\phi(\beta, \theta) = \beta \log\left(2f\left(r_\beta e^{i\theta}\right)\right) - \log r_\beta - i\theta - \beta \log \beta - (1/2 - \beta) \log(1/2 - \beta),$$

so that we have

$$a_n(\beta, \theta) \sim c_n \, \psi(\beta, \theta) \, e^{n\phi(\beta, \theta)} \qquad (28)$$

uniformly for $(\beta, \theta) \in \mathcal{X}_n \cap \mathcal{X}^*$.

Let D denote the differential operator

$$(D\phi(x))_j = \frac{\partial \phi(x)}{\partial x_j}.$$

We seek a stationary point of ϕ. The condition $\frac{\partial \phi(\beta, 0)}{\partial \theta} = 0$ is equivalent to the condition $\beta r_\beta f'(r_\beta) = f(r_\beta)$. Solving for r_β gives

$$r_\beta = \frac{1}{8}\left(8 - 4\beta - \beta^2 \pm \sqrt{\beta^3(8 + \beta)}\right).$$

20

We choose $r_\beta = \frac{1}{8}\left(8 - 4\beta - \beta^2 - \sqrt{\beta^3(8+\beta)}\right) \in (0,1)$, which ensures that $\frac{\partial \phi(\beta,0)}{\partial \theta} = 0$.
Next, we calculate that with this choice of r_β,

$$\frac{\partial \phi}{\partial \beta}(\beta,0) = \log\left(\frac{\left(4 - \beta - \sqrt{\beta(8+\beta)}\right)(1-2\beta)}{\beta\left(\beta + \sqrt{\beta(8+\beta)}\right)}\right).$$

Setting this equal to 0 and solving for β gives the equation $(3\beta - 1)(\beta^2 - 4\beta + 2) = 0$.
The only solution with $\beta \in \left(0, \frac{1}{2}\right]$ is $\beta = \frac{1}{3}$ so we choose $x^* = \left(\frac{1}{3}, 0\right)$ and check that $D\phi(x^*) = 0$.

Note that $\phi(x^*) = \log\left(4\sqrt{\frac{2}{3}}\right)$, and

$$H = -\begin{pmatrix} \frac{63}{5} & 0 \\ 0 & \frac{5}{2} \end{pmatrix}$$

is the Hessian matrix of ϕ at x^*. Define $C_1 = 5/8$, so that $-4C_1$ is the largest eigenvalue of H.

Now, define $\hat{\phi}$ by $\hat{\phi}(x) = \phi(x) - \phi(x^*)$, and define $\hat{a}_n : \mathcal{X}_n \to \mathbb{C}$ by

$$\hat{a}_n(x) = c_n^{-1} e^{-n\phi(x^*)} a_n(x).$$

With a Taylor expansion about x^*, for $x \in \mathcal{X}^{1/2}$ we have

$$\hat{\phi}(x) = \frac{1}{2}(x - x^*)^T H(x - x^*) + h(x)|x - x^*|^2, \tag{29}$$

where $h(x)$ is complex and $h(x) \to 0$ as $x \to x^*$. For all $v \in \mathbb{R}^2$ we have $v^T H v \leq -2C_1|v|^2$, so we can choose $\xi < \frac{1}{6}$ such that $\Re\hat{\phi}(x) \leq -C_1|x - x^*|^2$ for $|x - x^*| < \xi$. Define $\mathcal{X}^* = \left\{x \in \mathcal{X}^{1/2} : |x - x^*| < \xi\right\}$, satisfying the requirement for (27).

Next, define the sets

$$\mathcal{X}^{(1)} = \left\{x \in \mathcal{X}^* : |x - x^*| < n^{-1/3}\right\},$$
$$\mathcal{X}^{(2)} = \mathcal{X}^* \setminus \mathcal{X}^{(1)},$$
$$\mathcal{X}^{(3)} = \mathcal{X}^{1/2} \setminus \mathcal{X}^*,$$
$$\mathcal{X}^{(4)} = \mathcal{X} \setminus \mathcal{X}^{1/2},$$

so that with

$$F_n^{(j)} = \sum_{b=1}^{n/2+2} \int_{-\pi}^{\pi} \hat{a}_n(b/n, \theta) \, \mathbf{1}_{\mathcal{X}^{(j)}}(b/n, \theta) \, d\theta$$

we have

$$c_n^{-1} e^{-n\phi(x^*)} F_n = F_n^{(1)} + F_n^{(2)} + F_n^{(3)} + F_n^{(4)}. \tag{30}$$

Lemma 4.2. *With notation as above, we have*

$$F_n^{(1)} + F_n^{(2)} + F_n^{(3)} + F_n^{(4)} \sim F_n^{(1)} \sim \frac{144\pi}{\sqrt{7}}.$$

21

Proof. Note that ψ is a continuous function defined on a compact set. So, ψ is absolutely bounded on its domain, by C_2 say. By (28), it follows that $|\hat{a}_n(x)| = O\left(e^{n\Re\hat{\phi}(x)}\right)$ uniformly for $x \in \mathcal{X}^*$. For $x \in \mathcal{X}^{(2)}$ we have

$$n\Re\hat{\phi}(x) \leq -nC_1|x - x^*|^2 \leq -C_1 n^{1/3} \to -\infty$$

and consequently

$$\left|F_n^{(2)}\right| \leq \sum_{b=1}^{n} \int_{-\pi}^{\pi} |\hat{a}_n(b/n, \theta)|\, \mathbf{1}_{\mathcal{X}^{(2)}}(b/n, \theta)\, \mathrm{d}\theta = O\left(ne^{-C_1 n^{1/3}}\right) = o(1). \qquad (31)$$

Now (26) implies that for each β, $\Re\hat{\phi}(\beta, \theta)$ is uniquely maximized when $\theta = 0$. Also $\frac{\partial\Re\hat{\phi}}{\partial\beta}(\beta, 0) = \frac{\partial\hat{\phi}}{\partial\beta}(\beta, 0) = 0$ only for $(\beta, 0) = x^*$, since $\hat{\phi}$ is real along the line $\theta = 0$. Checking the values of $\Re\hat{\phi}(\beta, 0)$ in the limit as $\beta \to 0$ and $\beta \to \frac{1}{2}$, it follows that $\Re\hat{\phi}$ attains a unique maximum on $\mathcal{X}^{1/2}$ at x^*. Let $-C_3 < 0$ be the maximum value of $\Re\hat{\phi}$ on $\mathcal{X}^{(3)}$.

Let $u \vee w = \max\{u, w\}$ for real numbers u, w. We now redo the calculations of (27) using an alternate form of Stirling's inequality which holds for all $k \geq 0$, namely $\sqrt{k \vee 1}\left(\frac{k}{e}\right)^k \leq k!$. For $(\beta, \theta) \in \mathcal{X}^{1/2} \cap \mathcal{X}_n$,

$$|\hat{a}_n(\beta, \theta)| \leq \frac{ne^2}{\sqrt{(\beta n \vee 1)((n/2 - \beta n + 2) \vee 1)}(1/2 - \beta)^2} e^{n\Re\hat{\phi}(x)}$$
$$= e^{n\Re\hat{\phi}(x) + o(n)}.$$

It follows that

$$\left|F_n^{(3)}\right| = O\left(ne^{-C_3 n/2}\right) = o(1). \qquad (32)$$

Next suppose that $(\beta, \theta) \in F_n^{(4)} \cap \mathcal{X}_n$. Then we have $\frac{1}{2} < \beta \leq \frac{1}{2} + o(1)$ and $(n/2 - b + 2)! = 1$. By the alternate form of Stirling's inequality and (26),

$$|a_n(\beta, \theta)| \leq \frac{n^3 e^2}{\sqrt{(\beta n \vee 1)}} c_n \left|\frac{\left(2f\left(r_\beta e^{i\theta}\right)\right)^\beta}{r_\beta e^{i\theta}\beta^\beta}\right|^n$$
$$\leq e^{o(n)} c_n \left(\frac{\left(2f\left(r_{1/2}\right)\right)^{1/2}}{r_{1/2}\left(\frac{1}{2}\right)^{1/2}}\right)^n.$$

By direct computation,

$$\log \frac{\left(2f\left(r_{1/2}\right)\right)^{1/2}}{r_{1/2}\left(\frac{1}{2}\right)^{1/2}} = \phi(x^*) + C_4$$

for some $C_4 > 0$. It follows that

$$\left|F_n^{(4)}\right| = O\left(ne^{-C_4 n/2}\right) = o(1). \qquad (33)$$

22

It remains to consider $F_n^{(1)}$. Define $\lceil(\beta,\theta)\rceil = \left(\frac{\lceil\beta n\rceil}{n},\theta\right)$, so that $F_n = n\int_{\mathcal{X}}\hat{a}_n(\lceil x\rceil)\mathrm{d}x$. For any $y \in \mathbb{R}^2$, define $x_y = x^* + y/\sqrt{n}$ and $B_n = \{y : \lceil x_y\rceil \in \mathcal{X}^{(1)}\}$, so that we can make the change of variables

$$F_n^{(1)} = \int_{B_n}\hat{a}_n(\lceil x_y\rceil)\mathrm{d}y.$$

Note that

$$\left|y/\sqrt{n}\right| = |x_y - x^*| \le |x_y - \lceil x_y\rceil| + |\lceil x_y\rceil - x^*| = O\left(n^{-1/3} + n^{-1}\right) = O\left(n^{-1/3}\right)$$

for $y \in B_n$, so that B_n is approximately a ball of radius $O\left(n^{1/6}\right)$.

Next, a first-order Taylor expansion of $D\phi$ about x^* gives

$$|D\phi(x_y)| = O\left(|y/\sqrt{n}|\right) = O\left(n^{-1/3}\right).$$

Another first-order Taylor expansion of ϕ about x_y gives

$$\phi(\lceil x_y\rceil) - \phi(x_y) = O(|D\phi(x_y)||\lceil x_y\rceil - x_y|) = O\left(n^{-4/3}\right),$$

so that

$$e^{n\phi(x_y)} \sim e^{n\phi(\lceil x_y\rceil)} \tag{34}$$

uniformly. Now, for each $y \in \mathbb{R}^2$, we have $\lceil x_y\rceil \to x^*$. For n large enough so that $y \in B_n$, we have $\psi(\lceil x_y\rceil) \to \psi(x^*)$ by continuity and $e^{n\phi(x_y)} \to e^{\frac{1}{2}y^T H y}$ by (29). We therefore have $\mathbf{1}_{B_n}(y)\hat{a}_n(\lceil x_y\rceil) \to \psi(x^*)e^{\frac{1}{2}y^T H y}$ for all y.

Recalling that C_2 and $2C_1$ are bounds involving ψ and ϕ respectively, with (34) we have $|\mathbf{1}_{B_n}(y)\hat{a}_n(\lceil x_y\rceil)| \le 2C_2 e^{-C_1|y|^2}$ for sufficiently large n. Since

$$(\det(-H))^{-1/2} = \frac{1}{3}\sqrt{\frac{2}{7}}, \quad \psi(x^*) = 108\sqrt{2}$$

we obtain, by the dominated convergence theorem,

$$F_n^{(1)} \to \psi(x^*)\int_{\mathbb{R}^2}e^{\frac{1}{2}y^T H y}\,\mathrm{d}y = 2\pi\,\psi(x^*)\,(\det(-H))^{-1/2} = \frac{144\pi}{\sqrt{7}}.$$

Combining this with (31–33) completes the proof. □

We now pull these calculations together to prove the following.

Lemma 4.3. *Let $d = 3$. Then*

$$\mathbb{E}[Y^2] \sim \frac{18}{\sqrt{14}}\left(\frac{16}{3}\right)^n,$$

and hence

$$\frac{\mathbb{E}[Y^2]}{[\mathbb{E}Y]^2} \to \frac{9}{\sqrt{14}}.$$

It follows that Condition (A4) *holds when $d = 3$.*

23

Proof. Lemma 4.2 and (30) prove that

$$F_n \sim \frac{72}{n^3 \sqrt{7}} \left(4\sqrt{\frac{2e}{3n}} \right)^n.$$

Substituting $d = 3$ into (25) and applying (4) gives

$$\mathbb{E} Y^2 = \frac{(6\sqrt{2})^n (n!)^2}{4\pi \, \#P(3n)} F_n \sim \frac{18}{\sqrt{14}} \left(\frac{16}{3} \right)^n,$$

using Stirling's approximation. Then, with (12) and Lemma 3.2, we conclude that

$$\frac{\mathbb{E} Y^2}{(\mathbb{E} Y)^2} \to \frac{9}{\sqrt{14}} = \exp\left(\sum_{j=1}^{\infty} \lambda_j \gamma_j^2 \right).$$

This establishes Condition (A4), as required. $\qquad \square$

We can now complete the proof of Theorem 1.2.

Proof of Theorem 1.2. We will prove that for general $d \geq 3$, if Condition (A4) holds then Conjecture 1.3 is true. In particular, this will prove Theorem 1.2, using Lemma 4.3.

Suppose that Condition (A4) is satisfied for some fixed integer $d \geq 3$. Then by Lemma 3.2 we may apply Theorem 1.4 to conclude that (5) holds for Y. Therefore, for all real numbers y we have

$$\begin{aligned}
\mathbb{P}(Y_{\mathcal{G}}/\mathbb{E} Y_{\mathcal{G}} < y) &= \mathbb{P}(Y/\mathbb{E} Y_{\mathcal{G}} < y | X_1 = X_2 = 0) \\
&\to \mathbb{P}(W \exp(\lambda_1 \zeta_1 + \lambda_2 \zeta_2) < y | Z_1 = Z_2 = 0) \\
&= \mathbb{P}\left(\prod_{j=3}^{\infty} (1 + \zeta_j)^{Z_j} e^{-\lambda_j \zeta_j} < y \right).
\end{aligned}$$

Hence Conjecture 1.3 is a consequence of (A4). $\qquad \square$

4.1 Support for Conjecture 1.3

Let $p_d(n)$ denote the quotient

$$\frac{\mathbb{E} Y^2}{(\mathbb{E} Y)^2} \Big/ \exp\left(\sum_{j=1}^{\infty} \lambda_j \zeta_j^2 \right).$$

For any fixed integer $d \geq 4$, Conjecture 1.3 holds if and only Condition (A4) from Theorem 1.4 is satisfied; that is, if and only if $p_d(n) \sim 1$. Using Lemma 4.1 we can compute $p_d(100)$ for various values of d:

d	3	4	5	6	100
$p_d(100)$	0.9761	0.9881	0.9921	0.9942	0.9998

24

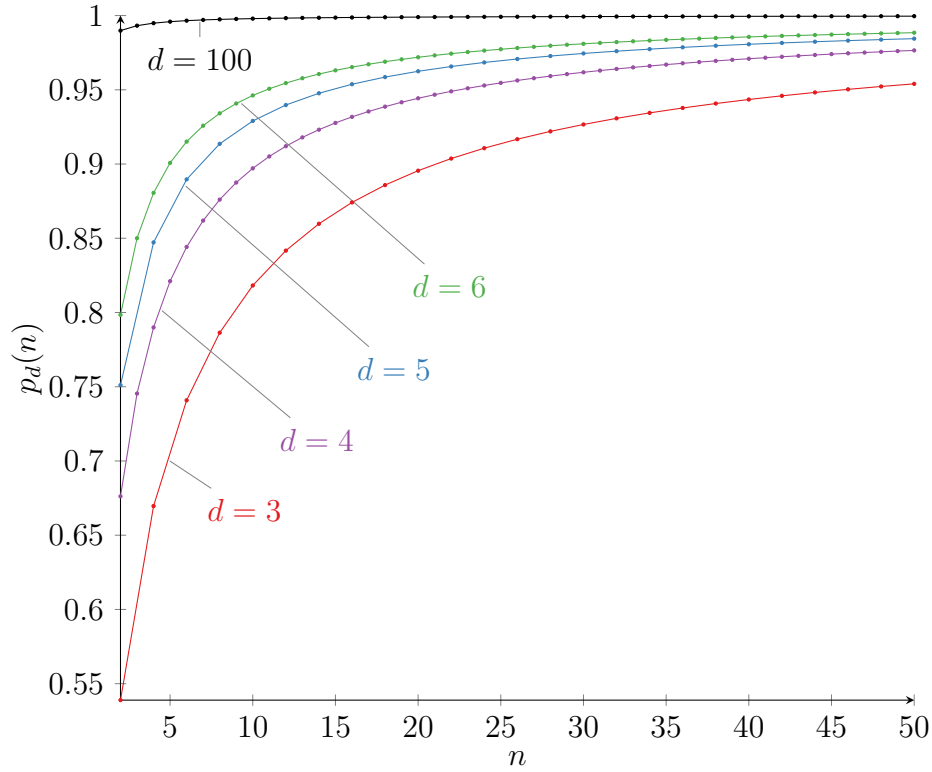Figure 1: A plot of $p_d(n)$ for $d \in \{3, 4, 5, 6, 100\}$

Figure 1 is a plot of $p_d(n)$ for $d \in \{3, 4, 5, 6, 100\}$ and $n \leq 50$.

This plot supports our conjecture that $p_d(n) \sim 1$ for all $d \geq 4$. Indeed, the rate of convergence to 1 appears to increase as d increases.

We now give an asymptotic result which is equivalent to Conjecture 1.3. Combining (4), (A4), (12), Lemma 3.2, Lemma 4.1 and applying Stirling's formula shows that for a fixed integer $d \geq 4$, Conjecture 1.3 holds if and only if

$$\sum_{b=1}^{n} \frac{2^b}{b! \left((d/2 - 1)n - b + 2\right)!} [z^n] \left(\sum_{j=1}^{\infty} \binom{(d-1)j}{j} z^j \right)^b$$

$$\sim \frac{2d^2}{\pi(d-2)^4 n^3} \sqrt{\frac{2d-2}{d^2 - d + 1}} \left(\frac{(d-1)^{2(d-1)}}{(d-2)^{2(d-2)}} \left(\frac{2e}{dn} \right)^{d/2-1} \right)^n . \tag{35}$$

Acknowledgements

25

References

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

[2] M.J. Golin, X. Yong and Y. Zhang, The asymptotic number of spanning trees in circulant graphs, *Discrete Mathematics* **310** (2010), 792–803.

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

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

[5] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley-Interscience, New York, 2000.

[6] R. Lyons, Asymptotic enumeration of spanning trees, *Combinatorics, Probability and Computing* **14** (2005), 491–522.

[7] B.D. McKay, Spanning trees in random regular graphs, in *Proceedings of the Third Caribbean Conference on Combinatorics and Computing*, 1981, 139–143.

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

[9] B.D. McKay, Asymptotics for 0-1 matrices with prescribed line sums, in *Enumeration and Design*, Academic Press, Toronto, 1984, pp. 225–238.

[10] C. Merino and D. Welsh, Forests, colorings, and acyclic orientations of the square lattice, *Annals of Combinatorics* **3** (1999), 417–429.

[11] J.W. Moon, *Counting Labelled Trees*, Canadian Mathematical Monographs, Vol. 1, Canadian Mathematical Congress, Montreal, 1970.

[12] B.R. Myers, On spanning trees, weighted compositions, Fibonacci numbers, and resistor networks, *SIAM Review* **17** (1975), 465–474.

[13] S.D. Noble and G.F. Royle, The Merino-Welsh conjecture holds for series-parallel graphs, *European Journal of Combinatorics* **38** (2014), 24–35.

[14] D. Welsh, The Tutte polynomial, *Random Structures and Algorithms* **15** (1999), 210–228.

[15] N.C. Wormald, (1999). Models of random regular graphs, in *Surveys in Combinatorics 1999* (J.D. Lamb and D.A. Preece, eds.), vol. 267 of LMS Lecture Note Series, Cambridge University Press, Cambridge, 1999, pp. 239–298.