The number of Euler tours of a random d-in/d-out graph

Páidí Creed[1†] and Mary Cryan[2‡]

[1] *Department of Computer Science, Royal Holloway University of London, Egham, Surrey, UK, TW20 0EX, UK.*
[2] *School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, UK.*

In this paper we obtain the expectation and variance of the number of Euler tours of a random d-in/d-out directed graph, for $d \geqslant 2$. We use this to obtain the asymptotic distribution and prove a concentration result. We are then able to show that a very simple approach for uniform sampling or approximately counting Euler tours yields algorithms running in expected polynomial time for almost every d-in/d-out graph. We make use of the BEST theorem of de Bruijn, van Aardenne-Ehrenfest, Smith and Tutte, which shows that the number of Euler tours of a d-in/d-out graph is the product of the number of arborescences and the term $[(d-1)!]^n/n$. Therefore most of our effort is towards estimating the asymptotic distribution of the number of arborescences of a random d-in/d-out graph.

Keywords: Euler tours, Random graphs, Directed graphs, Counting, Sampling

1 Introduction

Let $G = (V, E)$ be a directed graph. An *Euler tour* of G is any ordering $e_{\pi(1)}, \ldots, e_{\pi(|E|)}$ of the set of arcs E such that for every $1 \leqslant i < |E|$, the target vertex of arc $e_{\pi(i)}$ is the source vertex of $e_{\pi(i+1)}$, and such that the target vertex of $e_{\pi(|E|)}$ is the source of $e_{\pi(1)}$. We use $ET(G)$ to denote the set of Euler tours of G, where two Euler tours are considered to be equivalent if one is a cyclic permutation of the other. It is a well-known fact that a directed graph G has an Euler tour if and only if G is connected and if for each $v \in V$, the in-degree and out-degree of v are equal.

In this paper, we are interested in the number of Euler tours of a random d-in/d-out graph. Let $\mathcal{G}_{\mathrm{n}}^{\mathrm{d},\mathrm{d}}$ be the space of all d-in/d-out graphs on n vertices. We will consider the problem of determining the asymptotic distribution of the number of Euler tours of a random $G \in \mathcal{G}_{\mathrm{n}}^{\mathrm{d},\mathrm{d}}$. Similar results have been obtained for a variety of structures in uniformly random d-regular undirected graphs, e.g., Hamiltonian cycles [13, 14, 6], 1-factors [10], and 2-factors [12]. In each of these results, one of the goals was to prove that the structure of interest occurs in G with high probability when G is chosen uniformly at random from the set of all undirected d-regular graphs. Since every connected d-in/d-out graph has an Euler tour, the existence question is not of interest here. In the case of Hamiltonian cycles the asymptotic distribution was

†Email: paidi@cs.rhul.ac.uk. Supported by EPSRC grants EP/F01161X/1 and EP/D043905/1.
‡Supported by EPSRC grant EP/D043905/1

1365–8050 © 2010 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France

further used by Frieze et al.[6] to prove that very simple algorithms for random sampling and approximate counting of Hamiltonian cycles run in expected polynomial time for almost every d-regular graph. This paper contains analogous counting and sampling results for Euler tours of d-in/d-out graphs for $d \geqslant 2$.

Our result uses a well-known relationship between the Euler tours and *arborescences* of a graph. An *arborescence* of a directed graph $G = (V, E)$ is a rooted spanning tree of G in which all arcs are directed *towards* the root. We will use $ARBS(G)$ to denote the set of arborescences of G and, for any $v \in V$, use $ARBS(G, v)$ to denote the set of arborescences rooted at v. For any Eulerian directed graph G, the **BEST** Theorem (due to de **B**ruijn and van Aardenne-**E**hrenfest [18], extending a result of **S**mith and **T**utte [15]) reduces the problem of computing $|ET(G)|$ to the problem of computing $|ARBS(G, v)|$, for any vertex $v \in V$.

Theorem 1 ([15, 18]) *Let $G = (V, E)$ be an Eulerian directed graph. For any $v \in V$, we have*

$$|ET(G)| = \left[\prod_{u \in V} (outdeg(u) - 1)!\right] |ARBS(G, v)| . \tag{1}$$

We remark that the above theorem enables exact counting or sampling of Euler tours of any directed graph in polynomial time, as both these tasks can be carried out in polynomial time for arborescences of any directed graph [17, 4, 11]. The simple algorithms for uniform sampling and approximate counting we consider in Section 4, have so far only been analysed for Eulerian tournaments, in [9] (as part of their analysis of Euler tours on the undirected complete graph with an odd number of vertices).

We generate graphs using a directed version of the configuration model [2, 3]. We define the configuration space $\Phi_n^{d,d}$ as follows. Let S_v and T_v, $1 \leqslant v \leqslant n$, be disjoint d-sets and let $S = \cup S_v$ and $T = \cup T_v$. We say S_v is the set of *configuration points* available for arcs leaving v and T_v is the set of points available for arcs entering v. A *configuration* F is a perfect matching from S to T and $\Phi_n^{d,d}$ is the set of all configurations. Note that $|\,\Phi_n^{d,d}\,| = (dn)!$. Each configuration $F \in \Phi_n^{d,d}$ *projects* to a directed d-in/d-out pseudo-graph $\sigma(\mathrm{F})$ by identifying the elements of S_v and T_v. That is, $\sigma(\mathrm{F})$ has an arc (u, v) for each pair from $S_u \times T_v$ that is contained in F. This model has been analysed in the past for the case when $d = 2$ in [1, Section 7], to obtain an estimate of the expected number of Euler tours of a random $G \in \mathcal{G}_{\mathrm{n}}^{\mathrm{d,d}}$. One nice property of the model, and of the original configuration model, is that directed graphs (without loops or double arcs) are generated with equal probability. Hence, by studying properties of uniformly random configurations we can infer results about uniformly random elements of $\mathcal{G}_{\mathrm{n}}^{\mathrm{d,d}}$ by conditioning on there being no loops or double arcs.

The results in this paper are of an asymptotic nature. If a_n and b_n are sequences of numbers, we take $a_n \sim b_n$ to mean $\lim_{n\to\infty} a_n/b_n = 1$. Given a sequence of random variables X_n and random variable Z defined on the same probability space Ω, we say X_n *converges in distribution* to Z, or Z has the *asymptotic distribution* of X_n, if

$$\lim_{n\to\infty} \mathbb{P}[X_n = x] = \mathbb{P}[Z = x] , \quad \forall x \in \Omega .$$

The proof idea we use in this paper is that of *conditioning on short cycle counts*, pioneered by Robinson and Wormald in [13, 14]. Implicit in this pair of papers (and the subsequent work of Frieze et al. [6]) is a characterisation of the asymptotic distribution of the number of Hamiltonian cycles in a random d-regular graph in terms of random variables counting the number of i-cycles, for all fixed positive integers i. Janson [7] streamlined the technique of Robinson and Wormald and proved a general theorem (stated

here as Theorem 4). In Section 3, we use Theorem 4 to obtain an asymptotic distribution for the number of Euler tours of a random d-in/d-out graph.

In some places in the paper we only give proof sketches. For more detailed proofs, and some results for the more general model for in-degree/out-degree sequences $\vec{d} = (d_1, \ldots, d_n)$, see [5].

2 Expectation and Variance of Euler tours

In this section, we obtain the expectation and variance of the number of Euler tours of a random d-in/d-out graph. We will use two particular facts several times in the proofs of this section. Recall the definition of *falling factorial powers*: for every $n, k \in \mathbb{N}$,

$$(n)_k = n(n-1)(n-2)\cdots(n-k+1)\,.$$

Fact 1 *Falling factorial powers of sums obey the well known* multinomial theorem

$$(x_1 + x_2 + \cdots + x_l)_k = \sum_{\sum \delta_i = k} \binom{k}{\delta_1, \ldots, \delta_l} \prod_{i=1}^{l} (x_i)_{\delta_i}\,,$$

where the sum is taken over all partitions of k into l non-negative integer parts.

Fact 2 (see, e.g., [16]) *Let $V = \{1, 2, \ldots, n\}$. The number of k-forests on V in which v has δ_v children is*

$$\binom{n-1}{k-1}\binom{n-k}{\delta_v : v \in V}\,.$$

We use Fact 1 and Fact 2 to prove the following lemma. In this lemma, and in the proofs of subsequent results, we will speak of a configuration for an (in-directed) arborescence or forest. We take this to mean a partial matching from S to T (in the configuration model) that projects to a arborescence or forest.

Lemma 1 *Suppose we have a set of vertices $V = \{1, 2, \ldots, n\}$ for which there are x_v points for arcs entering $v \in V$ and y_v points for arcs leaving $v \in V$, with x_v not necessarily equal to y_v. Then, the number of ways to choose a configuration for an in-directed forest rooted at $R \subseteq V$ is*

$$\left(\prod_{v \notin R} y_v\right)\left(\sum_{v \in R} x_v\right)\left(\sum_{v \in V} x_v - 1\right)_{n-|R|-1}\,. \tag{2}$$

Proof: Let $\mathcal{F}$ be a forest on $\{1, 2, \ldots, n\}$ rooted at R and let δ_v be the number of children of v in $\mathcal{F}$, for each $v \in V$. The number of ways to choose points for the source and target vertex of each arc in $\mathcal{F}$ is

$$\left(\prod_{v \notin R} y_v\right)\left(\prod_{v \in V} (x_v)_{\delta_v}\right)\,, \tag{3}$$

since we must choose a point for the start of the arc directed away from each $v \notin R$ and choose one of the x_v points for the end of each of the δ_v arcs directed towards each $v \in V$.

Let $k = \sum_{v \in R} \delta_v$. We can construct a forest rooted at R by first choosing a k-forest on $V - R$, and then attaching each root of this forest as a child of some $v \in R$. By Fact 2, the number of k-forests on $V - R$ in which $v \in V - R$ has exactly δ_v children is

$$\binom{n - |R| - 1}{k - 1} \binom{n - |R| - k}{\delta_v : v \in V - R}, \tag{4}$$

and the number of ways to divide the roots of this forest amongst the members of R so that each $v \in R$ has δ_v children is

$$\binom{k}{\delta_v : v \in R}. \tag{5}$$

Combining (3), (4) and (5) and summing over all possible values for δ_v gives

$$\left(\prod_{v \notin R} y_v\right) \times \sum_{k=1}^{n-|R|} \binom{n - |R| - 1}{k - 1} \left(\sum_{\sum_{v \in R} \delta_v = k} \binom{k}{\delta_v : v \in R} \prod_{v \in R} (x_v)_{\delta_v}\right) \times \left(\sum_{\sum_{v \notin R} \delta_v = n - |R| - k} \binom{n - |R| - k}{\delta_v : v \notin R} \prod_{v \notin R} (x_v)_{\delta_v}\right). \tag{6}$$

By Fact 1, we see that the two sums over the different δ_v in (6) are expansions of the falling factorial powers $(\sum_{v \in R} x_v)_k$ and $(\sum_{v \notin R} x_v)_{n-|R|-k}$, respectively. Hence, (6) is equal to

$$\left(\prod_{v \notin R} y_v\right) \sum_{k=1}^{n-|R|} \binom{n - |R| - 1}{k - 1} (\sum_{v \in R} x_v)_k (\sum_{v \in \bar{R}} x_v)_{n-|R|-k}.$$

Applying Fact 1 again gives (2). □

We now use Lemma 1 to analyse the expectation and variance of the number of arborescences in $\sigma(F)$, when F is chosen uniformly at random from $\Phi_n^{d,d}$. We say $\mathcal{A} \subset \mathrm{F}$ is an arborescence of $\mathrm{F} \in \Phi_n^{d,d}$ if $\sigma(\mathcal{A})$ is an arborescence of $\sigma(F)$. In the following proofs, we will abuse terminology slightly and switch between speaking of arborescences of configurations and directed graphs arbitrarily.

Theorem 2 *Let d be some fixed integer, $d \geqslant 2$, and let $\mathcal{A}_n^\star$ denote the number of arborescences (rooted at any vertex) of a uniformly random $\mathrm{F} \in \Phi_n^{d,d}$. Then,*

$$\mathbb{E}[\mathcal{A}_n^\star] = d^{n-1};$$
$$\mathbb{E}[(\mathcal{A}_n^\star)^2] = \frac{n}{dn - n + 1} d^{2n-1}.$$

Proof: To calculate the first moment of $\mathcal{A}_n^\star$ we need to enumerate pairs $(\mathrm{F}, \mathcal{A})$, where $\mathrm{F} \in \Phi_n^{d,d}$ and $\mathcal{A}$ is an arborescence of F, and then divide by $|\Phi_n^{d,d}|$. Given $\mathcal{A}$, it is easy to count the number of configurations $\mathrm{F} \supset \mathcal{A}$: there are exactly $dn - n + 1$ arcs not contained in any particular arborescence of a d-in/d-out graph, and hence $(dn - n + 1)!$ ways to choose a configuration for those arcs. Applying Lemma 1 with $x_v = y_v = d$ for all $v \in V$ tells us that the number of arborescences rooted at any particular vertex is

$$d^n (dn - 1)_{n-2}.$$

Multiplying this by $n(dn-n+1)!$, and dividing by the total number of configurations, $(dn)!$, gives the claimed value for $\mathbb{E}[\mathcal{A}_n^\star]$.

To find the second moment, we need to count the number of ways to choose triples $(\mathrm{F}, \mathcal{A}, \mathcal{A}')$, where $\mathrm{F} \in \Phi_n^{d,d}$ and $\mathcal{A}$ and $\mathcal{A}'$ are both arborescences of F. We approach this task by first counting partial configurations for the intersection $\mathcal{A} \cap \mathcal{A}'$. We suppose $(\mathcal{A}, \mathcal{A}')$ is a pair of arborescences of some configuration F and that $\mathcal{F} = \mathcal{A} \cap \mathcal{A}'$ is a forest with l components. By Lemma 1 (with $x_v = y_v = d$), the number of possibilities for $\mathcal{F}$ is

$$\binom{n}{l} d^{n-l+1} l (dn-1)_{n-l-1} \,. \tag{7}$$

For each $t = 1 \ldots l$, let $\mathcal{F}_t$ denote the t-th component of $\mathcal{F}$, and let n_t be the number of vertices in $\mathcal{F}_t$. Then the total number of points available for arcs directed towards vertices of $\mathcal{F}_t$ is $x_t = (d-1)n_t + 1$. We choose the rest of $\mathcal{A}$ and $\mathcal{A}'$ as follows. For $t = 1 \ldots l$, we contract each $\mathcal{F}_t$ to a single vertex t. Now, choosing configurations for the rest of $\mathcal{A}$ and $\mathcal{A}'$ is equivalent to choosing configurations for a pair of disjoint trees on $\{1, 2, \ldots, l\}$, with d points available for arcs directed away from t and x_t points available for arcs directed towards t, for $t = 1 \ldots l$.

There are two cases. First suppose $\mathcal{F}_r$ is the component containing the root of $\mathcal{A}$, $\mathcal{F}_{r'}$ is the component containing the root of $\mathcal{A}'$, and $r \neq r'$. Assume we have already chosen $\mathcal{A}$ such that $\mathcal{A} \setminus \mathcal{A}'$ contains δ_t arcs directed towards vertices in $\mathcal{F}_t$ in $\mathcal{A}$. Then, by Lemma 1, the number of ways to choose $\mathcal{A}' \setminus \mathcal{A}$ is

$$d(x_{r'} - \delta_{r'})(d-1)^{l-2}(dn-n)_{l-2} \,.$$

The number of ways to choose $\mathcal{A} \setminus \mathcal{A}'$ is

$$d^{l-1} \sum_{|\boldsymbol{\delta}|=l-1, \delta_r \geqslant 1} \left[\binom{l-2}{\delta_r - 1; \delta_v : v \neq r} \prod_{t=1}^{l} (x_t)_{\delta_t} \right];$$

and therefore the number of ways to complete $\mathcal{F}$ to $\mathcal{A} \cup \mathcal{A}'$ is equal to $d^l(d-1)^{l-2}(dn-n)_{l-2}$ times

$$\begin{aligned} & x_{r'} \sum_{|\boldsymbol{\delta}|=l-1, \delta_r \geqslant 1} \binom{l-2}{\delta_r - 1; \delta_t : t \neq r} \prod_{t=1}^{l} (x_t)_{\delta_t} \\ & -(l-2) \sum_{|\boldsymbol{\delta}|=l-1, \delta_r, \delta_{r'} \geqslant 1} \binom{l-3}{\delta_r - 1, \delta_{r'} - 1; \delta_t : t \notin \{r, r'\}} \prod_{t=1}^{l} (x_t)_{\delta_t} \,. \end{aligned} \tag{8}$$

Using Fact 1, we can simplify the expression for the number of ways to complete $\mathcal{F}$ to $\mathcal{A} \cup \mathcal{A}'$ to be

$$\begin{aligned} & x_r x_{r'} d^l (d-1)^{l-2} (dn-n)_{l-2} \left\{ (dn-n+l-1)_{l-2} - (l-2)(dn-n+l-2)_{l-3} \right\} \\ = \;& x_r x_{r'} d^l (d-1)^{l-2} (dn-n+l-2)_{2l-4} \,. \end{aligned} \tag{9}$$

In the second case, suppose $r = r'$. Then a similar argument gives

$$(x_r)_2 d^{l-1} (d-1)^{l-1} (dn-n+l-2)_{2l-4} \,. \tag{10}$$

Multiplying (9) and (10) by $(d-1)^2$ and $d(d-1)$, respectively, and summing over r and r' gives

$$d^l (d-1)^l (dn-n+l)_{2l-2} \,.$$

Combining this with (7) we obtain the expression

$$\binom{n}{l} d^n (d-1)^l \frac{l}{n} (dn)! \tag{11}$$

which over-counts the number of triples $(\mathrm{F}, \mathcal{A}, \mathcal{A}')$ by weighting $(\mathrm{F}, \mathcal{A}, \mathcal{A}')$ by $d(d-1)$ if $\mathcal{A}$ and $\mathcal{A}'$ are rooted at the same vertex, and by $(d-1)^2$ if $\mathcal{A}$ and $\mathcal{A}'$ have different roots.

From the BEST Theorem (Theorem 1), we know that the number of triples $(\mathrm{F}, \mathcal{A}, \mathcal{A}')$ in which $\mathcal{A}$ is rooted at r and $\mathcal{A}'$ is rooted at r' does not depend on the choice of r and r'. Thus, it follows that the factor by which (11) over-counts the number of triples is

$$\frac{(d-1)^2 n(n-1) + d(d-1)n}{n^2} = \frac{(dn-n+1)_2}{n^2}\,. \tag{12}$$

Dividing (11) by (12) and $(dn)!$, and summing over l, yields

$$\mathbb{E}[(\mathcal{A}_n^\star)^2] = \frac{n}{(dn-n+1)_2} d^n \sum_{l=1}^{n} \binom{n}{l} l(d-1)^l = \frac{n}{dn-n+1} d^{2n-1}\,.$$

□

Recall that simple directed graphs are generated with equal probability in the configuration model. Thus, by conditioning on $\sigma(F)$ containing no loops or 2-cycles, we can infer the first two moments of the number of arborescences of a uniformly random $G \in \mathcal{G}_{\mathrm{n}}^{\mathrm{d},\mathrm{d}}$. We provide a sketch below in Theorem 3. For a detailed proof, see Section 3.3.1 of [5].

Theorem 3 *Let d be some fixed constant, $d \geqslant 2$, and let $\mathcal{A}_n$ denote the number of arborescences of a uniformly random $G \in \mathcal{G}_{\mathrm{n}}^{\mathrm{d},\mathrm{d}}$. Then,*

$$\mathbb{E}[\mathcal{A}_n] \sim e^1 d^{n-1};$$
$$\mathbb{E}[\mathcal{A}_n^2] \sim e^{2-1/d} \frac{1}{d-1} d^{2n-1}\,.$$

Proof: Let L_n and D_n denote random variables counting the number of loops and double edges in $\sigma(\mathrm{F})$ when F is chosen uniformly at random from $\Phi_n^{d,d}$. The directed graph $\sigma(\mathrm{F})$ is simple precisely when $L_n(\mathrm{F}) = D_n(\mathrm{F}) = 0$; that is,

$$\mathbb{P}(G \text{ is simple}) = \mathbb{P}(L_n = D_n = 0)\,,$$

for G obtained as the projection of uniformly random $\mathrm{F} \in \Phi_n^{d,d}$. Consider the term in $\mathbb{E}[(L_n)_k]$ corresponding to configurations in which j_i vertices have i loops:

$$(n)_{j_1+j_2+,\ldots+j_d} \prod_{i=1}^{d} ((d)_{2i})^{j_i} \frac{(dn-k)!}{(dn)!}\,.$$

Since $k = \sum_{i=1}^{d} i j_i$, this value is $O(n^{-j_2-2j_3-\cdots-(d-1)j_d})$. Hence, the only terms in $(L_n)_k$ which have non-negligible contribution, as $n \to \infty$, are those where each vertex has at most one loop. This means that, for every k, $\mathbb{E}[(L_n)_k] \sim d^k$, so L_n converges to a Poisson random variable with mean d. A similar

argument shows that D_n converges to a Poisson random variable with mean $(d-1)^2/2$, and that the asymptotic distributions of L_n and D_n are independent. Hence, we have $\mathbb{P}(L_n = D_n = 0) \sim \exp(-d - (d-1)^2/2)$.

Now, let $L_n^{(1)} = \mathcal{A}_n^\star L_n$ and $D_n^{(1)} = \mathcal{A}_n^\star D_n$. These random variables count the number of loops and double edges in $\sigma(F)$ weighted by the number of arborescences. This is equivalent to considering the number of loops or double edges in F when F is obtained by choosing a uniformly random element of

$$\bar{\Phi} = \{(F, \mathcal{A}) : F \in \Phi_n^{d,d}, \mathcal{A} \in ARBS(F)\}\,.$$

First, suppose we have a loop at v in F. A loop edge cannot be contained in any arborescence of $\sigma(F)$, so we can repeat the argument from Theorem 2, except with one vertex having in/out-degree $d-1$, giving $\mathbb{E}[L_n^{(1)}] = d-1$.

Now, suppose we have a pair of edges in F, $e, f \in S_u \times T_v$, for some $u \neq v$. Note that there are $2\binom{d}{2}^2 (n)_2$ ways to choose e and f. The number of arborescences rooted at u, which can be extended to $(F, \mathcal{A})$ pairs where F contains e and f, is equal to the number of arborescences rooted at u in a model where there are $d-2$ points available for arcs directed towards v, d points available for arcs directed towards each $w \neq v$, and d points available for arcs directed away from each $w \neq u$. By Lemma 1, this number is $d^n(dn-3)_{n-2}$. There are exactly $(dn-n-1)!$ ways to extend each such arborescence to a configuration containing e and f, and hence we can conclude that $\mathbb{E}[D_n^{(1)}] \sim (d-1)^2/2$. Moreover, $L_n^{(1)}$ and $D_n^{(1)}$ are in fact asymptotic to independent Poisson random variables. Hence, we can conclude that

$$\mathbb{P}[L_n^{(1)} = D_n^{(1)} = 0] \sim \exp(-d + 1 - (d-1)^2/2)\,,$$

and so $\mathbb{E}[\mathcal{A}_n] \sim e^1 \mathbb{E}[\mathcal{A}_n^\star]$.

Similarly, we define $L_n^{(2)} = (\mathcal{A}_n^\star)^2 L_n$ and $D_n^{(2)} = (\mathcal{A}_n^\star)^2 D_n$, which count the number of loops and double edges when F is obtained by choosing a random element of

$$\{(F, \mathcal{A}, \mathcal{A}') : F \in \Phi, \mathcal{A}, \mathcal{A}' \in ARBS(\sigma(F))\}\,.$$

As we did for $L_n^{(1)}$ and $D_n^{(1)}$, we can perform similar analysis to that used to compute the second moment of $\mathcal{A}_n^\star$, which we omit here due to lack of space (see [5] for details). We are able to show that $L_n^{(2)}$ and $D_n^{(2)}$ are asymptotic to independent Poisson random variables with means $(d-1)^2/d$ and $(d-1)^2/2$, respectively, from which we can conclude $\mathbb{E}[\mathcal{A}_n^2] \sim e^{2-1/d}\mathbb{E}[(\mathcal{A}_n^\star)^2]$. □

Given the expectation and variance of the number of arborescences of a random d-in/d-out graph we can, from the BEST Theorem (Theorem 1), deduce the expectation and variance of the number of Euler tours of a uniformly random $G \in \mathcal{G}_{\mathrm{n}}^{\mathrm{d,d}}$.

Corollary 1 *Let d be some fixed integer, $d \geqslant 2$, and let $\mathcal{T}_n$ denote the random variable counting the Euler tours of a uniformly random $G \in \mathcal{G}_{\mathrm{n}}^{\mathrm{d,d}}$. Then,*

$$\mathbb{E}[\mathcal{T}_n] \sim e^1 \frac{1}{dn}(d!)^n\,;$$
$$\mathbb{E}[\mathcal{T}_n^2] \sim e^{2-1/d} \frac{1}{d(d-1)n^2}(d!)^{2n}\,.$$

3 Asymptotic distribution of Euler tours

To compute the asymptotic distribution we will use the following general theorem of Janson [7] (see also [8, Chapter 9]).

Theorem 4 (Janson [7]) *Let $\lambda_i > 0$ and $\delta_i \geqslant -1$, $i = 1, 2, \ldots$, be constants and suppose that for each n there are random variables X_{in}, $i = 1, 2, \ldots$, and Y_n (defined on the same probability space) such that X_{in} is non-negative integer valued and $\mathbb{E}[Y_n] \neq 0$ (at least for large n) and furthermore the following conditions are satisfied:*

1. *$X_{in} \to X_{i\infty}$ (in distribution) as $n \to \infty$, jointly for all i, where $X_{i\infty}$ is a Poisson random variable with mean λ_i;*
2. *For any finite sequence $x_1, \ldots x_k$ of non-negative integers*

$$\frac{\mathbb{E}[Y_n | X_{1n} = x_1, \ldots X_{kn} = x_k]}{\mathbb{E}[Y_n]} \to \prod_{i=1}^{k} (1+\delta_i)^{x_i} e^{-\lambda_i \delta_i} \quad as\, n \to \infty\,;$$

3. *$\sum_i \lambda_i \delta_i^2 < \infty$;*
4. *$\frac{\mathbb{E}[Y_n^2]}{\mathbb{E}[Y_n]^2} \to \exp(\sum_i \lambda_i \delta_i^2)$;*

Then

$$\frac{Y_n}{\mathbb{E}[Y_n]} \to W = \prod_{i=1}^{\infty} (1+\delta_i)^{X_{i\infty}} e^{-\lambda_i \delta_i}\,.$$

Moreover, this and the convergence in 1 holds jointly. The infinite product defining W converges a.s. and in L_2, with $\mathbb{E}[W] = 1$ and $\mathbb{E}[W^2] = \exp(\sum_i \lambda_i \delta_i^2) = lim_{n\to\infty} \mathbb{E}[Y_n]^2 / \mathbb{E}[Y_n]^2$. Hence, the normalised variables are uniformly square integrable. Furthermore, the event $W > 0$ equals, up to a set of probability 0, the event that $X_{i\infty} > 0$ for some i with $\delta_i = -1$. In particular, $W > 0$ a.s. if and only if every $\delta_i > -1$.

In our application of Theorem 4 we will have $Y_n = \mathcal{T}_n$, the random variable counting Euler tours of d-in/d-out graphs, and X_{in} equal to the number of directed i-cycles in a random d-in/d-out graph. To apply Theorem 4 we need the following two lemmas.

Lemma 2 *For each positive integer i let X_{in} count the number of directed i-cycles in a directed graph obtained as the projection of a uniformly random $\mathrm{F} \in \Phi_n^{d,d}$. The variables X_{in} are asymptotically independent Poisson random variables with means $\mathbb{E}[X_{in}] = \lambda_i = \frac{d^i}{i}$.*

Proof: We say a set of i edges $e_1, e_2, \ldots, e_i$ in a configuration is an i-cycle if there is a sequence of distinct vertices $v_1, v_2, \ldots, v_i$ such that $e_j \in S_{v_j} \times T_{v_{j+1}}$ for $j < i$ and $e_i \in S_{v_i} \times T_{v_1}$. The probability of any particular i-cycle being contained in a random $\mathrm{F} \in \Phi_n^{d,d}$ is

$$\frac{(dn-i)!}{(dn)!} \sim \frac{1}{(dn)^i}\,.$$

So, to estimate $\mathbb{E}[X_{in}]$ all we need to do is count the number of different i-cycles that can occur in some $\mathrm{F} \in \Phi_n^{d,d}$ and then divide by $(dn)^i$. Let I be some i-subset of $\{1, 2, \ldots, n\}$. There are $(i-1)!$ different

ways to arrange I into an i-cycle $(v_1, v_2, \ldots, v_i)$ and then d^{2i} ways to choose edges $e_j \in S_{v_j} \times T_{v_{j+1}}$ for $1 \leqslant j < i$ and $e_i \in S_{v_i} \times T_{v_1}$. Hence,

$$\mathbb{E}[X_{in}] \sim \frac{(i-1)!}{(dn)^i}\binom{n}{i}d^{2i}\,,$$

and so $\mathbb{E}[X_{in}] \sim \lambda_i$. The contribution to $\mathbb{E}[(X_{in})_k]$ from configurations in which there are s intersections between i-cycles is $O(n^{-s})$. Hence, we can assume that cycles occur independently and $\mathbb{E}[(X_{in})_k] \sim (\lambda_i)^k$, for every k, so X_{in} converges to a Poisson random variable with mean λ_i. Moreover, by a similar argument, we can show that products of factorial moments converge,

$$\mathbb{E}\left[\prod_{i=1}^{\ell}(X_{in})_{k_i}\right] \sim \prod_{i=1}^{\ell}\lambda_i^{k_i}\,,$$

whence the asymptotic distributions of the X_{in} are independent Poisson variables. □

Lemma 3 *Let X_{in} be as in Lemma 3 and let $\mu_i = \frac{d^i-1}{i}$. Then, for any fixed set of integers $j_1, j_2, \ldots, j_k$ we have*

$$\frac{\mathbb{E}[\mathcal{A}_n^\star \prod_{i=1}^k (X_{in})_{j_i}]}{\mathbb{E}[\mathcal{A}_n^\star]} \sim \prod_{i=1}^k \mu_i^{j_i}\,.$$

Proof: We only verify

$$\frac{\mathbb{E}[\mathcal{A}_n^\star\, X_{in}]}{\mathbb{E}[\mathcal{A}_n^\star]} \sim \mu_i\,;$$

the factorial moments then hold for the same reasons as were given in Lemma 2.

Let $\bar{\Phi}$ be the set defined in the proof of Theorem 3 and let I be an i-subset of $\{1, \ldots, n\}$. As in the previous lemma, there are $(i-1)!d^{2i}$ ways to choose a configuration for an i-cycle on I. To estimate $\mathbb{E}[\mathcal{A}_n^\star\, X_{in}]$ we need to calculate the probability that a particular i-cycle C is contained in F when $(\mathrm{F}, \mathcal{A})$ is chosen uniformly at random from $\bar{\Phi}$. Suppose $C \cap \mathcal{A}$ has l components, $P_1, P_2, \ldots, P_l$, each of which is a directed path, and let v_k be the final vertex in the path P_k for $1 \leqslant k \leqslant l$. Choosing the remainder of $\mathcal{A}$ is then equivalent to choosing an arborescence on $(V \backslash I) \cup \{v_k : 1 \leqslant k \leqslant l\}$, where we have collapsed each path to a single vertex. Each $v \in V \backslash I$ has d points available for arcs directed towards or away from v. For each $k = 1 \ldots l$, there are $|P_k|(d_u - 1)$ points available for arcs directed towards v_k, and $d-1$ points available for arcs directed away from v_k. Once we have chosen $\mathcal{A}$, there are $(dn - n - l - 1)!$ ways to complete F. Hence, using Lemma 1, we can deduce that the number of ways to choose the remainder of $(\mathrm{F}, \mathcal{A})$ is

$$n(d-1)^l d^{n-i}(dn - i - 1)!\,.$$

Summing over all the possible choices for $P = \{v_1, v_2, \ldots, v_l\}$ and dividing by $|\bar{\Phi}| = d^{n-1}(dn)!$ and applying Stirling's formula yields

$$\mathbb{E}[\mathcal{A}_n^\star\, X_{in}] \sim \mu_i \mathbb{E}[\mathcal{A}_n^\star]\,.$$ □

Corollary 2 *Let $d \geqslant 2$ be some fixed integer, and let $\mathcal{T}_n$ denote the number of Euler tours in a directed graph G chosen uniformly at random from $\mathcal{G}_{\mathrm{n}}^{\mathrm{d,d}}$. For any fixed set of integers $j_2, \ldots, j_k$ we have*

$$\frac{\mathbb{E}[\mathcal{T}_n \prod_{i=1}^k (X_{in})_{j_i}]}{\mathbb{E}[\mathcal{T}_n]} \sim \prod_{i=2}^k \mu_i^{j_i}\,.$$

We now have sufficient ammunition to apply Janson's theorem and obtain an asymptotic distribution for the number of Euler tours of a uniformly random $G \in \mathcal{G}_{\mathrm{n}}^{\mathrm{d},\mathrm{d}}$.

Theorem 5 *Let $d \geqslant 2$ be some fixed integer, and let $\mathcal{T}_n$ denote the number of Euler tours in a directed graph G chosen uniformly at random from $\mathcal{G}_{\mathrm{n}}^{\mathrm{d},\mathrm{d}}$. Then,*

$$\frac{\mathcal{T}_n}{\mathbb{E}[\mathcal{T}_n]} \sim \prod_{i=2}^{\infty} \left(1 - \frac{1}{d^i}\right)^{Z_i} e^{1/i},$$

where the Z_i are independent Poisson random variables with means d^i/i.

Proof: It suffices to show that conditions (1) to (4) of Theorem 4 are satisfied by $\mathcal{T}_n$ and $\{X_{in} : i \geqslant 2\}$, where X_{in} is the random variable counting i-cycles. Lemma 2 and Corollary 2 provide conditions (1) and (2) with

$$\lambda_i = \frac{d^i}{i} \quad \text{and} \quad \delta_i = -\frac{1}{d^i}.$$

With these values, evaluating the sum in condition (3) gives

$$\sum_{i=2}^{\infty} \frac{1}{id^i} = -\frac{1}{d} + \log\left(\frac{d}{d-1}\right). \tag{13}$$

Finally, Corollary 1 provides condition (4). □

4 Generating and counting Euler tours

Consider the following process for generating a random Euler tour of a directed Eulerian graph G. Start at any vertex v and at each vertex choose a random arc leaving v, subject to the condition that no arc is chosen more than once. Eventually, we will reach v with no available outgoing arcs, at which point we will have constructed a circuit. If this is an Euler tour, accept; otherwise, repeat the process. The probability that the circuit generated is an Euler tour is

$$\frac{d_v |ET(G)|}{\prod_{u \in V} d_u!}. \tag{14}$$

This procedure was first considered in [9], where the authors showed it terminates in expected polynomial time when $G = K_n$ for odd n. If $G \in \mathcal{G}_{\mathrm{n}}^{\mathrm{d},\mathrm{d}}$ then (14) becomes $d|ET(G)|/(d!)^n$. The following theorem uses the results of the previous section (by a similar argument to that used in [6, Lemma 1]) to show that this value is $\Omega(n^{-2})$ with high probability when G is chosen uniformly at random from $\mathcal{G}_{\mathrm{n}}^{\mathrm{d},\mathrm{d}}$. When this is the case, we can generate uniformly random Euler tours of G in expected polynomial time. Moreover, by generating a number of samples and recording the proportion of executions which yielded Euler tours, we can approximate $|ET(G)|$.

Theorem 6 *Let d be some fixed integer, $d \geqslant 2$, and let G be chosen uniformly at random from $\mathcal{G}_{\mathrm{n}}^{\mathrm{d},\mathrm{d}}$. Then,*

$$\mathbb{P}\left[\frac{|ET(G)|}{(d!)^n} \in \Omega(n^{-2})\right] \to 1,$$

as $n \to \infty$.

Proof: We first note that by the estimate for $\mathbb{E}[\mathcal{T}_n]$ given in Corollary 1, the statement above is equivalent to showing that

$$\mathbb{P}\left[\frac{|\,\mathcal{T}_n\,|}{\mathbb{E}[\mathcal{T}_n]} \in \Omega(n^{-1})\right] \to 1\,.$$

We will combine our results to date to derive this statement. For $\mathbf{x} = (x_2, \ldots, x_k)$ we define $\mathcal{G}_{\mathbf{x}}$ to be the set of all d-in/d-out graphs containing exactly x_i directed cycles of length i for each $i = 2 \ldots k$, and

$$W(\mathbf{x}) = \prod_{i=2}^{k} \left(1 - \frac{1}{d^i}\right)^{x_i} e^{1/i}\,.$$

For each fixed $\gamma > 0$ we define

$$S(\gamma) = \{\mathbf{x} : x_i \leqslant \lambda_i + \gamma\lambda_i \text{ for } 2 \leqslant i \leqslant k\}\,.$$

From Lemma 2 (and Lemma 3 of [6]), we can deduce that the probability of a random d-in/d-out graph G not being contained in $\mathcal{G}_{\mathbf{x}}$ for some $\mathbf{x} \in S(\gamma)$ is $O(e^{-a\gamma})$, where a is an absolute constant independent of γ. Hence, to verify the theorem all we need do is show that

$$W(\mathbf{x}) \geqslant e^{-(b+c)\gamma} \quad \forall \mathbf{x} \in S(\gamma)\,, \tag{15}$$

where b and c are absolute constants independent of γ. For any particular b, c and γ, we can choose n sufficiently large so that $e^{-(b+c\gamma)} \geqslant n^{-1}$. Then, if (15) holds, we have

$$\mathbb{P}\left[\mathcal{T}_n \geqslant n^{-1}\mathbb{E}[\mathcal{T}_n]\right] \geqslant 1 - e^{-a\gamma}\,.$$

The above holds for any constant γ, and so can be taken as equal to 1 in the limiting case. Using the value of $\mathbb{E}[\mathcal{T}_n]$ provided by Corollary 1 will, as noted above, then give our proof.

So, it remains to prove (15). For $x \in S(\gamma)$ we have $W(\mathbf{x}) = AB^{\gamma}$, where

$$A = \prod_{i \geqslant 2} \left(1 - \frac{1}{d^i}\right)^{\lambda_i} e^{1/i} \tag{16}$$

$$B = \prod_{i \geqslant 2} \left(1 - \frac{1}{d^i}\right)^{\lambda_i^{2/3}}\,. \tag{17}$$

We can bound the right hand side of (16) as

$$A \geqslant \prod_{i=2}^{\infty} \exp\left(\frac{1}{i} - \frac{d^i}{i(d^i-1)}\right) = \exp\left(\sum_{i=2}^{\infty} -\frac{1}{i(d^i-1)}\right)\,.$$

The sum inside the exponential is clearly convergent, so we can conclude that $A \geqslant e^{-b}$ for some absolute constant b. Similarly, we can bound B by

$$B \geqslant \exp\left(-\sum_{i=2}^{\infty} \frac{1}{(i)^{2/3}d^{i/3}}\right)\,,$$

and, again, the sum in the exponential is convergent, so $B^{\gamma} \geqslant e^{-c\gamma}$ for some absolute constant c. □

Acknowledgements

This work has benefited from conversations with Kyriakos Kalorkoti. The original idea for applying Robinson and Wormald's technique to the number of Euler tours came from Alan Frieze.

References

[1] Richard Arratia, Béla Bollobás, Don Coppersmith, and Gregory B. Sorkin. Euler circuits and DNA sequencing by hybridization. *Discrete Applied Mathematics*, **104**(1-3):63–96, 2000.

[2] Edward A. Bender and E. Rodney Canfield. The Asymptotic Number of Labeled Graphs with Given Degree Sequences. *Journal of Combinatorial Theory, Series A*, **24**(3):296–307, 1978.

[3] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, **1**:311–316, 1980.

[4] Charles J. Colbourn, Wendy J. Myrvold, and Eugene Neufeld. Two Algorithms for Unranking Arborescences. *Journal of Algorithms*, **20**(2):268–281, 1996.

[5] Patrick J. Creed. *Counting and Sampling problems on Eulerian graphs*. PhD thesis, School of Informatics, University of Edinburgh, 2010. Submitted.

[6] Alan M. Frieze, Mark Jerrum, Michael Molloy, Robert W. Robinson, and Nicholas C. Wormald. Generating and Counting Hamilton Cycles in Random Regular Graphs. *Journal of Algorithms*, **21**(1):176–198, 1996.

[7] Svante Janson. Random Regular Graphs: Asymptotic Distributions and Contiguity. *Combinatorics, Probability & Computing*, 4:369–405, 1995.

[8] Svante Janson, Tomasz Łuczak, and Andrzej Riciński. *Random Graphs*. Wiley, 2000.

[9] Brendan D. McKay and Robert W. Robinson. Asymptotic Enumeration of Eulerian Circuits in the Complete Graph. *Combinatorics, Probability & Computing*, **7**(4):437–449, 1998.

[10] Michael S. O. Molloy, Hanna D. Robalewska, Robert W. Robinson, and Nicholas C. Wormald. 1-Factorizations of random regular graphs. *Random Structures and Algorithms*, **10**(3):305–321, 1997.

[11] James Gary Propp and David Bruce Wilson. How to Get a Perfectly Random Sample From a Generic Markov Chain and Generate a Random Spanning Tree of a Directed Graph. *Journal of Algorithms*, **27**:170–217, 1998.

[12] Hanna D. Robalewska. 2-factors in random regular graphs. *Journal of Graph Theory*, **23**(3):215–224, 1996.

[13] Robert W. Robinson and Nicholas C. Wormald. Almost All Cubic Graphs Are Hamiltonian. *Random Structures & Algorithms*, **3**(2):117–126, 1992.

[14] Robert W. Robinson and Nicholas C. Wormald. Almost All Regular Graphs Are Hamiltonian. *Random Structures & Algorithms*, **5**(2):363–374, 1994.

[15] C. A. B. Smith and W. T. Tutte. On unicursal paths in networks of degree 4. *American Mathematical Monthly*, **48**:233–237, 1941.

[16] Richard P. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, 2001.

[17] W.T. Tutte. *Graph Theory, Encyclopedia of Mathematics and it Applications*, volume 21. Addison-Wesley Publishing Company, 1984.

[18] Tanja van Aardenne-Ehrenfest and Nicholas G. de Bruijn. Circuits and trees in oriented linear graphs. *Simon Stevin*, **28**:203–217, 1951.