Tournaments with many Hamilton cycles

N. C. Wormald *
Department of Combinatorics and Optimization
University of Waterloo
Waterloo ON
Canada N2L 3G1
nwormald@uwaterloo.ca

Abstract

The object of interest is the maximum number, $h(n)$, of Hamilton cycles in an n-tournament. By considering the expected number of Hamilton cycles in various classes of random tournaments, we obtain new asymptotic lower bounds on $h(n)$. The best result so far is approximately 2.85584... times the expected number $g(n)$ of Hamilton cycles in a random n-tournament, and it is conjectured that $h(n) \sim cg(n)$ where $c \approx 2.855958$. The same statements hold for Hamilton paths.

AMS Classifications: 05A16, 05C30, 05C80, 05C38, 05D40

1 Results

A tournament is an orientation of a complete graph. A *Hamilton* cycle or *path* is a tournament T is a directed cycle or path which contains all vertices in T. Let $H(T)$ $(H'(T))$ denote the number of Hamilton cycles (paths) in

*Research supported by the Canadian Research Chairs Program and NSERC

a tournament T, and let S_n denote the set of all tournaments on n vertices. Define

$$h(n) = \max_{T \in S_n} H(T), \quad h'(n) = \max_{T \in S_n} H'(T).$$

Let $g(n) = 2^{-n}(n-1)!$ and $g'(n) = 2^{-n+1}n!$. If T is chosen uniformly at random from S_n, it is easy to calculate that the expected value $\mathbf{E}[H]$ of H is $g(n)$ and similarly that $\mathbf{E}[H'] = g'(n)$, by which method Szele [10] originally showed that $h'(n) \geq g'(n)$. Adler et al. [1] showed that $h'(n) \geq (e - o(1))g'(n)$ as $n \to \infty$, the first improvement over Szele's original result. The proof examined the expected value of $f(T)$ when T is randomly drawn from a certain subset of the regular tournaments (i.e. tournaments with equal indegree and outdegree at every vertex), obtained by partitioning the arc set into triangles and randomly orienting each as a directed 3-cycle. The best known upper bound is of the order of $n^c g'(n)$ for $c = 3/2$ by Alon [2] using Brégman's theorem (formerly Minc's conjecture) on permanents. This was recently improved to c slightly less than $5/4$ by Friedgut and Kahn [3] based on Radhakrishnan's proof of Brégman's theorem. The Hamilton cycle version of both of these results also holds, with h and g in place of h' and g'.

The object of the present paper is to investigate various classes of tournaments to find clues on what affects the number of Hamilton cycles in a tournament. As it turns out, the expected number of Hamilton cycles in a random regular tournament, which can be calculated using the methods of McKay [6] combined with those in [8], coincides with the lower bound $(e + o(1))g(n)$ on $h(n)$ obtained in [1]. We find several classes of random tournaments which improve this result. Most of these, but not all, are regular.

The discussion in [3] also asked for the minimum value of $H(T)$ when $T \in S_n$ is regular, a question which we make little progress on here. However, we do find some classes of random regular tournaments with significantly fewer Hamilton cycles than the expected number.

We obtain analogous results for Hamilton paths, but prefer to present the cycle results first. In [1] it is noted as interesting to determine if $h'(n) = \Theta(n!/2^n)$. We make no conclusive progress on this, but the results here are suggestive enough for us to conjecture the affirmative and to estimate $h'(n)$.

Let $d \geq 1$ be odd and consider a fixed tournament T^* with d vertices.

Let $\mathcal{T}(T^*) = \mathcal{T}(T^*, n)$ denote the set of random tournaments on dn vertices constructed by choosing d tournaments $T_1, \ldots, T_d$ independently uniformly at random from S_n, expanding each vertex i of T^* to the tournament T_i, and expanding each arc ij of T^* to all arcs from vertices of T_i to vertices of T_j. (To keep vertex labels distinct, renumber the vertices of the resulting tournament in a canonical way.)

Also define $\mathcal{R}(T^*) = \mathcal{R}(T^*, n)$ to be the restricted probability space formed from $\mathcal{T}(T^*)$, by conditioning on the event that each T_i is a regular tournament. (Assume that n is odd for this definition.) Define $\mathcal{R}_0(T^*)$ similarly, with the restriction to the set of tournaments which can be obtained by orienting the triangles in a decomposition of the edges of the complete graph K_n into triangles. These were the tournaments considered in the main argument in [1]. For all our asymptotic statements, $n \to \infty$.

Theorem 1 *Let T^* be any fixed regular tournament with $d \geq 1$ vertices. Then*

$$\mathbf{E}[H(\mathcal{R}(T^*))] \sim \mathbf{E}[H(\mathcal{R}_0(T^*))] \sim e^{1/d}\mathbf{E}[H(\mathcal{T}(T^*))].$$

The proof of this theorem, along with those of the other results stated here, are in the next section.

The rest of the results concern $\mathbf{E}[H(\mathcal{T}(T^*))]$, for use in conjunction with Theorem 1. The following case is given separately because it does not require much computation. If C_3 denotes the directed cycle on three vertices, then $\mathcal{T}(C_3)$ consists of three random n-tournaments arranged in a 3-cycle with all arcs between the three tournaments having consistent cyclic orientation.

Theorem 2 *We have*

$$\mathbf{E}[H(\mathcal{T}(C_3))] \sim 2g(n).$$

For the next results, which concern $\mathbf{E}[H(\mathcal{T}(T^*))]$ for arbitrary regular tournaments T^*, we need some definitions. Given a regular tournament T^* on $d \geq 3$ vertices, let A be the adjacency matrix of T^*, i.e. $A = (a_{ij})$ where $a_{ij} = 1$ if ij is a directed edge of T and $a_{ij} = 0$ otherwise. Put $D =$

$\operatorname{diag}(x_1, \ldots, x_d)$, and define $f(x_1, \ldots, x_n)$ to be the $(1,1)$ entry of $(I_d - AD)^{-1}$ and $\hat{f}(x_1, \ldots, x_n) = \frac{1}{d}\operatorname{trace}((I_d - AD)^{-1})$. Put

$$\begin{aligned} f_n = f_n(T^*) &= [y_1^n \cdots y_d^n] f\left(2y_1/(2-y_1), \ldots, 2y_d/(2-y_d)\right), \qquad (1.1) \\ \hat{f}_n &= [y_1^n \cdots y_d^n] \hat{f}\left(2y_1/(2-y_1), \ldots, 2y_d/(2-y_d)\right), \end{aligned}$$

where square brackets denote the extraction of coefficients.

Theorem 3 *Let T^* be any fixed regular tournament with at least three vertices. Then $f_n \sim \hat{f}_n$ and*

$$\mathbf{E}[H(\mathcal{T}(T^*))] \sim f_n g(dn) \left(\frac{2}{d}\right)^{dn} (2\pi n)^{(d-1)/2} \frac{d^{3/2}}{d-1}.$$

The next result is an asymptotic evaluation of f_n, for which we need a bit more notation. Define B to be the matrix whose (i,j) entry B_{ij} is $-y_j$ if $A_{ij} = 1$ and 0 otherwise, except for diagonal entries $B_{ii} = 1 - \frac{1}{2}y_i$. That is, $B = I_d - (\frac{1}{2}I_d + A)D_y$ where $D_y = \operatorname{diag}(y_1, \ldots, y_d)$. Let $K = K(\mathbf{y}) = \det B$. Set

$$z = 2/d \quad \text{and} \quad \mathbf{z} = (z, z, \ldots, z). \qquad (1.2)$$

Writing K_j for the partial derivative of K with respect to y_j, it will be shown that $K_d(\mathbf{z}) \neq 0$ and $K(\mathbf{z}) = 0$, so (e.g. by observing that K is linear in y_d) one may solve $K(\mathbf{y}) = 0$ in the neighbourhood of $\mathbf{z}$ and obtain $y_d = h(y_1, \ldots, y_{d-1})$ for an analytic function h. Define a function $\tilde{f}$ in the neighbourhood of $(0, 0, \ldots, 0)$ (the $(d-1)$-dimensional vector) by

$$\tilde{f}(\theta_1, \ldots, \theta_{d-1}) = \log h(ze^{i\theta_1}, \ldots, ze^{i\theta_{d-1}}).$$

Let $\mathcal{H}$ denote the determinant of the Hessian (matrix of second order partials) of $\tilde{f}$ at $\theta_j = 0$ for $j = 1, \ldots, d-1$.

Theorem 4 *Provided $\mathcal{H} \neq 0$,*

$$\mathbf{E}[H(\mathcal{T}(T^*))] \sim g(dn)\sqrt{\frac{d}{\mathcal{H}}}.$$

Notes It will be shown in the proof that the same result holds if we make definitions based on K_j for any $1 \le j \le d$. Also, as will be seen in the proof, $K_j(z)$ is equal to $-d/2$ times the (j,j) cofactor in $\det B(\mathbf{z})$. Perhaps the condition $\mathcal{H} \neq 0$ can be shown to be true for all regular tournaments T^*, but its value needs to be computed in any case.

The required calculations can be done individually for small tournaments T^*. Let R_d denote the cyclic regular tournament of order d, with arcs from i to $i+r \pmod d$ for $1 \le r < d/2$. For $d = 9$ only, let R_9' be the regular 9-vertex tournament in $\mathcal{R}(R_3, 3)$ where each of T_i is a copy of R_3.

The values of $\sqrt{d/\mathcal{H}}$ in the following table were obtained by the algebraic manipulation package Maple. (Presumably there is a more direct way to calculate this. In particular, it is unexplained why $d/\mathcal{H}$ turns out to be a perfect square in all the examples computed.) For large d, the exact ratio is omitted to save space. The next two rows are just to make comparisons easy. The second line is $\rho_{\mathcal{T}}(T^*) := \lim_{n\to\infty} \mathbf{E}[H(\mathcal{T}(T^*))]/g(dn)$, which (by Theorem 4) is the same quantity as in the first line, rounded to four decimal places for $d \le 13$. Notice that these results refer to the non-regular tournaments in $\mathcal{T}(T^*)$. For regular tournaments, the bottom row gives $\rho_{\mathcal{R}}(T^*) := \lim_{n\to\infty} \mathbf{E}[H(\mathcal{R}(T^*))]/g(dn)$, which comes by Theorem 1 on multiplying by $e^{1/d}$.

T^*	R_3	R_5	R_7	R_9	R_9'	R_{11}	R_{13}
$\sqrt{d/\mathcal{H}}$	2	$\frac{44}{19}$	$\frac{3032}{1231}$	$\frac{437584}{171769}$	$\frac{5488}{2197}$	$\frac{108859552}{41833969}$	$\frac{41637757888}{15769379963}$
$\rho_{\mathcal{T}}(T^*)$	2.0000	2.3158	2.4630	2.5475	2.4980	2.6022	2.6404
$\rho_{\mathcal{R}}(T^*)$	2.7912	2.8285	2.8413	2.8469	2.7915	2.8498	2.8515

T^*	R_{21}	R_{41}	R_{81}	R_{101}
$\rho_{\mathcal{T}}(T^*)$	2.7215188...	2.7867038...	2.820801899...	2.827747088 ...
$\rho_{\mathcal{R}}(T^*)$	2.8542501...	2.8555079...	2.855842467...	2.855883644...

(To avoid using prohibitively large amounts of memory when calculating with Maple for large d, the (j,k) entry of the Hessian was computed separately for each j and k, first setting $y_i = 2/d$ in the matrix B if $i \notin \{j,k,d\}$.)

All other regular tournaments T^* which were examined in this way produced lower numerical values. The 7-vertex tournament with cyclic triples abc and $a'b'c'$, with arcs aa', bb', cc', ad, bd and cd, and other arcs making it regular, gave $\sqrt{d/\mathcal{H}} = \frac{3088}{1287} \approx 2.3994$. For 9-vertex tournaments, the circulant tournament with arcs ij iff $j - i (\mathrm{mod}\ d) \in D$ yields values for $\mathbf{E}[H(\mathcal{T}(T^*))]/g(dn)$ of 1.4582 ($D = \{1,3,4,5\}$), 1.6811 ($D = \{1,2,4,5\}$), 2.4557 ($D = \{1,2,3,5\}$), 1.5618 ($D = \{2,3,4,5\}$), 1.5529 ($D = \{1,3,5,6\}$), 1.9325 ($D = \{1,4,5,6\}$), 1.8117 ($D = \{2,3,5,6\}$). However, this examination was not exhaustive in any way.

We conclude from the R_{101} result that $h(n) > 2.85588g(n)$ for infinitely many n. Incidentally we obtained the asymptotic upper bound $1.4582e^{1/9}g(n) \approx 1.63g(n)$ on the minimum number of Hamilton cycles in a regular n-tournament, but make no further efforts in this direction. We turn instead to Hamilton paths.

Theorem 5 *Theorems 1 and 4 are both valid with H and g replaced by H' and g'.*

Thus the table above is valid when interpreted for Hamilton paths rather than Hamilton cycles. It follows that $h'(n) > 2.85588g'(n)$ for infinitely many n.

Concluding Remarks and Conjectures

In this paper we do not consider the variance of H or H'. Routine calculations show that if T is chosen uniformly at random from S_n, then the variance of H is $o(g(n)^2)$. Thus, by Chebyshev's inequality, a random tournament with probability close to 1 contains approximately $g(n)$ Hamilton cycles. A significantly larger variance would have implied the existence of tournaments with significantly more than $g(n)$ Hamilton cycles. Presumably a similar concentration holds for the other models of random tournament considered in this paper.

In [2, Proposition 2.5] it was shown that for every tournament Q there exists a tournament with exactly one more vertex than Q (and containing Q) and with at least $H'(Q)/4$ Hamilton cycles. Thus $h'(n) > cg'(n)$ implies that $h(n) > cg(n)$, and therefore the conclusions of this type in this paper can be obtained by looking at H' alone. However, this argument does not quite permit us to derive $\mathbf{E}[H(\mathcal{T}(T^*))]$ given $\mathbf{E}[H'(\mathcal{T}(T^*))]$, and it is of interest to know about both of these quantities. In any case, for the results in this paper the cycle case is a little simpler to consider than the path case.

It is interesting to speculate why R_9 has more Hamilton cycles than R_9': perhaps something to do with the decreased number of 3-cycles expected to be spanned by the out-set of each vertex. This suggests that perhaps $T^* = R_d$ maximises $\mathbf{E}[H(\mathcal{R}(T^*))]$ for $|V(T^*)| = d$. Alon [2] felt it plausible that the cyclic tournament R_n maximises the number of Hamilton cycles of any n-tournament, that is, $H(R_n) = h(n)$. The results reported in and after the table suggest this quite strongly, so we have turned this into a conjecture (see below). The tournaments in $H(\mathcal{T}(R_d))$ possess a large proportion (approximately $(d-1)/d$) of the arcs of R_{dn}, with the other arcs random. $H(\mathcal{T}(R_d))$ is a regular version of this.

Both [1] and [3] consider that $h(n) = \Theta(g(n))$ is possible, which would imply the existence of $\liminf_{d\to\infty} \max_{|V(T^*)|=d} \mathbf{E}[H(\mathcal{R}(T^*))]/g(dn)$. We can estimate this lim inf to be approximately 2.8559579 in the following way, assuming that R_d achieves the maximum for each d. Firstly, fitting the polynomial $c_0 + c_1/d + c_2/d^2 + c_3/d^3 + c_4/d^4$ in $1/d$ to $\rho_{\mathcal{R}}(R_d)$ for $d = 11$, 21, 41, 81 and 101 gives $c_0 = 2.85595787\ldots$. Secondly, the value of c_1 obtained is less than 10^{-4}. This suggests that perhaps $c_1 = 0$. (Also c_3 appears to be small by the same criterion, but not c_5.) This is supported by the fact that fitting $c_0 + c_2/d^2 + c_3/d^3 + c_4/d^4$ to $\rho_{\mathcal{R}}(R_d)$ for $d = 11$, 21, 41 and 81 only, gives a polynomial $P(1/d)$ for which $\rho_{\mathcal{R}}(R_{101}) - P(1/101) < 2 \times 10^{-9}$. Fitting $c_0 + c_2/d^2 + c_3/d^3 + c_4/d^4 + c_5/d^5$ to $\rho_{\mathcal{R}}(R_d)$ for $d = 11$, 21, 41, 81 and 101 gives $c_0 = 2.85595789\ldots$, in agreement with the estimate above.

One might also attempt to extrapolate the exact values of $\sqrt{d/\mathcal{H}} = \rho_{\mathcal{T}}(R_d)$ from the values for small d. However, this may not be easy, as indicated by the prime factorisation of the denominator in the case $d = 11$, as 53×297535471.

The results above lead to the following conjectures, some of which are obviously stronger than the assertion that $h(n) = \Theta(g(n))$.

Conjecture 1 *For all $d \geq 3$, the maximum of $\mathbf{E}[H(\mathcal{T}(T^*))]$, over all tournaments T^* on d vertices (d odd, and n sufficiently large), is achieved for*

$T^* = R_d$.

Conjecture 2 *The limit* $\lim_{d\to\infty} \rho_{\mathcal{R}}(R_d)$ *(d odd) exists.*

Conjecture 3 $H(R_n) = h(n)$ *for* n *sufficiently large (and odd).*

Conjecture 4 *The limit* $\lim_{n\to\infty} h(n)/g(n)$ *exists.*

Assuming that Conjectures 2 and 4 are true, we believe that they refer to the same constant as estimated above, as follows.

Conjecture 5 *The limits* $\lim_{d\to\infty} \rho_{\mathcal{R}}(R_d)$ *and* $\lim_{n\to\infty} h(n)/g(n)$ *are equal and have the approximate value 2.855958.*

In the absence of verification of Conjectures 1 and 3, the limits $\lim_{d\to\infty} \max_{|V(T^*)|=d} \rho_{\mathcal{R}}(T^*)$ and $\lim_{n\to\infty} H(R_n)/g(n)$ (n odd) are also of interest, and we conjecture that these limits exist and have the value in Conjecture 5.

We also make the analogous conjectures for H', h' and g'. Of independent interest we have the following.

Conjecture 6 *For every regular tournament* T^* *on at least three vertices,* $\mathcal{H} \neq 0$ *and* $d/\mathcal{H}$ *is the square of a rational.*

Finally, we note that the two parts of Theorem 1, referring $\mathcal{R}_0$ to and to $\mathcal{R}$, have the same implications for $h(n)$, but we include both cases since it may be significant that they do have the same implications.

Acknowledgements Encouraging comments and suggestions on this problem were made by Jeff Kahn, Tomasz Łuczak and Svante Janson.

2 Proofs

We first derive two formulae for $\mathbf{E}[H(\mathcal{T}(T^*))]$.

Lemma 1 *Let* T^* *be any fixed tournament with at least three vertices. Then*

$$\mathbf{E}[H(\mathcal{T}(T^*))] = \left(\frac{n!}{2^n}\right)^d \sum_{j_1,\ldots,j_d\geq 1} j_1^{-1}[x_1^{j_1}\cdots x_d^{j_d}]f(x_1,\ldots,x_d)\prod_{i=1}^{d}\binom{n-1}{j_i-1}2^{j_i}. \quad (2.1)$$

Proof A Hamilton cycle in Q projects onto a closed walk W in T^* (by mapping T_i to the vertex i and ignoring loops). Let j_i denote the number of arcs of W whose head is i (which is of course equal to the number with tail i). Distinguish a special "initial" vertex v in T_1 such that the arc wv of Q has $w \notin V(T_1)$. This introduces a multiplicity of j_1 into the counting. For each i, the arcs of Q induce an ordered spanning forest F_i of T_i, i.e. a set of j_i disjoint directed paths (perhaps some containing only one vertex) which cover the vertices of T_i, the ordering determined by the initial vertex v. The number of such spanning forests which can ever appear is $n!\binom{n-1}{j_i-1}$, since they can be formed from a Hamilton path on $V(T_i)$ by deleting $j_i - 1$ edges. Any such spanning forest F_i contains $n - j_i$ arcs, so the probability that T_i contains F_i is 2^{-n+j_i}. Moreover since the T_i are independent, the probability that they contain a given set of forests $F_1, \ldots, F_d$ is $\prod_i 2^{-n+j_i}$. Given these forests and, in addition, the walk W, we can reconstruct the original Hamilton cycle in Q. The number of closed walks W starting at vertex 1, and with parameters j_i as above, is $[x_1^{j_1} \cdots x_d^{j_d}]f(x_1, \ldots, x_d)$, with $f(x_1, \ldots, x_d)$ defined as before. The lemma now follows by linearity of expectation (and dividing by j_1 to ignore the choice of v). ∎

Lemma 2 *If T^* is regular then, in the summation in (2.1), $(1 - o(1))$ of the contribution comes from terms with $j_i = (1 + o(1))(d - 1)n/d$ for all i. In particular,*

$$\mathbf{E}[H(\mathcal{T}(T^*))] \sim \frac{d(n!2^{-n})^d}{(d-1)n} \sum_{j_1,\ldots,j_d \geq 1} [x_1^{j_1} \cdots x_d^{j_d}]f(x_1, \ldots, x_d) \prod_{i=1}^{d} \binom{n-1}{j_i - 1} 2^{j_i}.$$

Proof Consider restricting the summation in (2.1) to those parameters such that $\sum_i j_i = dj$ for some fixed j. Since the out-degree of each vertex in T^* is $(d-1)/2$, there are $((d-1)/2)^{dj}$ walks of length dj, and moreover the position of a random walker will converge to the uniform distribution. Thus, the sum of all the coefficients of f in (2.1) with $\sum_i j_i = dj$ is asymptotic to $((d-1)/2)^{dj}/d$, and almost all of the value is contributed by the terms with $|j_i - j| < \epsilon j/2d$ $(1 \leq j \leq d)$ for any fixed $\epsilon > 0$. By making use of the log-concavity of the factors in the product in (2.1), it is straightforward to show that for the above-mentioned terms, this product is at least as large as it is if $|j_i - j| > \epsilon j$ for any i. Hence terms of the latter type contribute negligibly to the summation. For the other (non-negligible) terms, Stirling's formula shows that the product has value

$$\frac{n^{dn}2^{dj}\exp(\Theta(\epsilon^2 n))}{j^{dj}(n-j)^{d(n-j)}}. \tag{2.2}$$

On the other hand, the product of this with the sum of coefficients, $((d-1)/2)^{dj}/d$, is easily seen to be maximised for $j \sim (d-1)n/d$ (apart from the error term), and it is easy to see that unless $|j-(d-1)n/d| = O(\epsilon)$ the contribution will be negligible. The lemma follows (on taking $\epsilon \to 0$ slowly). ∎

Also for the proof of Theorem 1 we need to know about the expected number of spanning forests of paths in random regular tournaments. For a digraph D, let $E(D)$ be its arc set, $a(D) = |E(D)|$ its number of arcs, and $s(D)$ the number of path components of length at least 1. Of course, here n is restricted to odd integers. The proof of the following theorem uses an integral formulation of the number of regular n-tournaments which was used by McKay [6] to count them asymptotically. For another application of this formulation to a somewhat related problem, see McKay and Robinson [7].

Lemma 3 *For each $n \geq 1$ let F_n be either a directed graph on n labelled vertices with each component a directed path, or a directed Hamilton cycle. Then the probability that F_n appears in a random n-vertex regular tournament is asymptotically equal to $2^{-a(F_n)}e^{a(F_n)/n-s(F_n)/n}$ as $n \to \infty$.*

Proof Following [6], the number of regular tournaments on n vertices is precisely

$$RT(n) = \frac{1}{(2\pi i)^n}\oint\cdots\oint \frac{\prod_{1\leq j<k\leq n}(x_j^{-1}x_k + x_j x_k^{-1})}{x_1 x_2\cdots x_n}\, dx_1 dx_2\cdots dx_n, \tag{2.3}$$

where each integral is around the circle $|z| = 1$ in the anticlockwise direction. The numerator of the integrand is the generating function for tournaments enumerated by the excess of out-degree over in-degree at each vertex. Under the substitution $x_j = e^{i\theta_j}$ this becomes

$$RT(n) = \frac{2^{n(n-1)/2}}{(2\pi i)^n}\int_{U_n(\pi)}\prod_{1\leq j<k\leq n}\cos(\theta_j-\theta_k)\, d\theta_1 d\theta_2\cdots d\theta_n \tag{2.4}$$

where $U_n(x)$ denotes $[-x,x]^n$. For the general situation it suffices to consider the case that F_n is a subdigraph of the directed path $1,2,\ldots,n$, with the ordering imposed on the paths components being the natural one. (The Hamilton cycle case requires only a trivial modification to the argument for the case $a(F)=n-1$, and we make no further remark on it.) Then all arcs of F_n are of the form $(j,j+1)$. Let $E=\{j:(j,j+1)\in E(F)\}$. Then the number $N(F_n)$ of tournaments containing F_n is expressed as an integral like (2.3), but with the following factors altered: if $j\in E$ and $k=j+1$, the factor $(x_j^{-1}x_k+x_jx_k^{-1})$ is replaced by $x_jx_k^{-1}$. This means that the expression in (2.4) for the current problem becomes

$$N(F_n)=\frac{2^{n(n-1)/2-a(F_n)}}{(2\pi i)^n}\int_{U_n(\pi)}\frac{\prod_{1\le j<k\le n}\cos(\theta_j-\theta_k)}{\prod_{j\in E}\cos(\theta_j-\theta_{j+1})}\frac{\prod_{j\in E^+}e^{i\theta_j}}{\prod_{j\in E^-}e^{i\theta_j}}\,d\theta_1d\theta_2\cdots d\theta_n \tag{2.5}$$

where E^+ (E^-) is the set of vertices which begin (end) a directed path. Thus $E^+=\{j:(j,j+1)\in E(F),(j-1,j)\notin E(F)\}$, and for defining E^-, switch $\in$ and $\notin$.

We continue to follow the argument in [6] and split the integral in (2.5) into three major parts. For some $\epsilon>0$, let I_1 denote the contribution coming from those points at which $|\theta_j-\theta_n|\le n^{-1/2+\epsilon/4}$ or $|\theta_j-\theta_n+\pi|\le n^{-1/2+\epsilon/4}$ for $1\le j\le n-1$, where θ_j values are mod 2π. Still valid in the present context is the argument in [6] that we may set $\theta_n=0$ and assume all the θ_j are clustered near 0, which in this case gives

$$I_1=2^n\pi\int_{U_{n-1}(n^{-1/2+\epsilon/4})}\frac{\prod_{1\le j<k\le n}\cos(\theta_j-\theta_k)}{\prod_{j\in E}\cos(\theta_j-\theta_{j+1})}\frac{\prod_{j\in E^+}e^{i\theta_j}}{\prod_{j\in E^-}e^{i\theta_j}}\,d\theta_1d\theta_2\cdots d\theta_{n-1}. \tag{2.6}$$

Expanding (most of) the integrand gives (dropping terms dominated by the $o(1)$ error term),

$$\exp\left(o(1)+\sum_{1\le j<k\le n}-\frac{1}{2}(\theta_j-\theta_k)^2-\frac{1}{12}(\theta_j-\theta_k)^4\right) \tag{2.7}$$

$$\times\exp\left(\sum_{j\in E^+}i\theta_j\;+\sum_{j\in E^-}-i\theta_j\;+\sum_{j\in E}\frac{1}{2}(\theta_j-\theta_{j+1})^2\right), \tag{2.8}$$

where the first line contains the terms in the original calculation in [6].

The next step is to apply the linear transformation

$$y_j = \theta_j - \sum_{k=1}^{n-1} \theta_k/(n+\sqrt{n}),$$

for $1 \le j \le n-1$, which implies using [6, Equation (2)] that $\theta_j = y_j + \phi$ where

$$\phi = \sum_{k=1}^{n-1} y_k/(1+\sqrt{n}).$$

The transformed logarithm of (2.7) is dealt with in [6], the significant terms being

$$-\frac{n}{2}\sum_{k=1}^{n-1} y_k^2 - \frac{n}{12}\sum_{k=1}^{n-1} y_k^4 - \frac{n}{4}(\sum_{k=1}^{n-1} y_k^2)^2. \tag{2.9}$$

For the logarithm of (2.8), note that $|E^+| = |E^-|$ and so every ϕ which might appear is actually cancelled out. This gives

$$\sum_{j\in E^+} iy_j + \sum_{j\in E^-} -iy_j + \sum_{j\in E} \frac{1}{2}(y_j - y_{j+1})^2. \tag{2.10}$$

However, y_n (if it appears) must be defined as $-\phi$, which is a function of the other y_j. To handle the integral (of the exponential of all these terms) we may use the averaging argument leading to [8, Equation (2.4)]. This converts these terms into

$$\sum_{j\in E^+} -y_j^2 + \sum_{j\in E} \frac{1}{2}(y_j^2 + y_{j+1}^2) + o(1).$$

(The effect of the terms iy_j can be gauged by considering the effect of J in the proof of [8, Lemma 3]. The terms involving y_n, if they appear, similarly have negligible affect on the result of the averaging. The only different type of term this might produce appears if $n-1 \in E$, so that y_n^2 appears. A term like the fourth power of y_n appears in [8, Lemma 3] with coefficient I, and has negligible effect.) To evaluate (2.6) we must integrate the exponential of the sum of this and (2.9). The result is easily seen to be $W \exp(-|E^+|/n + |E|/n) = W \exp(a(F_n)/n - s(F_n)/n)$ where $W = (2\pi/n)^{(n-1)/2}\sqrt{n/e}$ is the original integral due to the terms in (2.9). Note that together with the extra factors outside the integral in (2.5) as compared to (2.4), this gives the result desired for the theorem.

It remains only to show that the part of the integral outside I_1 is negligible. The factors $e^{i\theta_j}$ in (2.6) have absolute value 1 so can be ignored. For the rest, the argument from [6] needs to be modified in a few small ways to be able to cope with the extra factor $\prod_{j\in E}\cos(\theta_j-\theta_{j+1})^{-1}$. We now refer heavily to the argument in [6], which splits the remaining part of the integral up into pieces denoted $I_2(r)$ and $I_3(h)$. At the point where values θ_j satisfying $|\theta_j|<\pi/8$ are singled out for evaluating $I_2(r)$, one may instead consider $|\theta_j|<1/\eta$ for a sufficiently small $\eta>0$. Then the value of $\prod_{j\in E}\cos(\theta_j-\theta_{j+1})^{-1}$ can be made sufficiently small that it is dominated by the exponentially small bound obtained on $I_2(r)$. For $I_3(h)$, a similar modification is required: by considering, in addition to what is already in [6], the number h' of θ_j satisfying $n^{-1/2+\eta\epsilon}\le|\theta_j|\le n^{-1/2+\epsilon/4}$, this case is also seen to be negligible. The lemma follows. ∎

Proof of Theorem 1 We first consider $\mathcal{R}(T^*)$. If T^* has only one vertex, then $\mathcal{T}(T^*)$ is just a random tournament chosen from S_n, so $\mathbf{E}[H(\mathcal{T}(T^*))]=g(n)$. By Lemma 3, the probability that a given Hamilton cycle occurs in a random regular tournament is $2^{-n}e(1+o(1))$. Since there are $(n-1)!$ possible Hamilton cycles, the expected number occurring is $(n-1)!2^{-n}e(1+o(1))\sim eg(n)$, as required.

If T^* has more than one vertex, then by the first statement in Lemma 2, with probability tending to 1, a random Hamilton cycle in Q induces $(d-1)n/d+o(n)$ paths in any given T_i. These paths come in a linear order, namely the order traversed in Q when it has a special initial vertex (see the proof of Lemma 1) and so they induce a forest $F_{i,n}$ as in Lemma 3. Thus $a(F_{i,n})\sim n/d$. Given $a(F_{i,n})$, one may define $F_{i,n}$ by choosing a Hamilton path and specifying which of its arcs are to be included. Random choice will choose two consecutive edges asymptotically $n/d^2+o(n)$ times (with probability tending to 1). Note that fixing $F_{i,n}$, the probability it appears in T_i is exactly $2^{-a(F_{i,n})}$. So with probability tending to 1, $s(F_{i,n})\sim(d-1)n/d^2$. For any set of forests $F_{1,n},\dots,F_{d,n}$ satisfying $a(F_{i,n})\sim n/d$ and $s(F_{i,n})\sim(d-1)n/d^2$ for $1\le i\le d$, Lemma 3 implies that switching from random tournaments T_i to random regular tournaments T_i, which produces $\mathcal{R}(T^*)$, will increase the probability of containment by

$$\prod_{i=1}^{d} e^{a(F_{i,n})/n-s(F_{i,n})/n} \sim e^{n/d}.$$

Thus $\mathbf{E}[H(\mathcal{R}(T^*))] \sim e^{1/d}\mathbf{E}[H(\mathcal{T}(T^*))]$, as required.

For $\mathcal{R}_0(T^*)$, we may first modify the calculation in [1] to see that the analogue of Lemma 3 applies for the tournaments considered there. This is rather more routine than the proof of Lemma 3 so we omit the details. The rest of the argument above for $\mathcal{R}$ therefore applies to $\mathcal{R}_0$, and we obtain the desired result. ∎

Proof of Theorem 2 We use the formula in Lemma 2. As noted in the proof of Lemma 1, the coefficient in this formula counts walks in T^*, which is in this case C_3. Thus, the coefficient is 0 unless $j_1 = j_2 = j_3 = j$ for some j, in which case it is 1. Hence,

$$\mathbf{E}[H(\mathcal{T}(C_3))] \sim \frac{3(n!2^{-n})^3}{2n} \sum_{j\geq 1} \left(\binom{n-1}{j-1} 2^j \right)^3.$$

Routine analysis shows that the maximum contribution to the sum comes from terms with $j \sim 2n/3$ (as indicated in Lemma 2) and, applying Stirling's formula and expanding, one obtains

$$\begin{aligned} \mathbf{E}[H(\mathcal{T}(T^*))] &\sim \frac{3\sqrt{2}}{n} \left(\frac{3n}{2e}\right)^n \sum_{x\in\mathbb{Z}} e^{-27x^2/4n} \\ &\sim \sqrt{\frac{8\pi}{3n}} \left(\frac{3n}{2e}\right)^n \\ &\sim 2g(3n). \end{aligned}$$ ∎

Proof of Theorem 3 Note that

$$\binom{n-1}{j-1} 2^j = 2^n [y^n](2y/(2-y))^j$$

and thus by applying Lemma 2

$$\mathbf{E}[H(\mathcal{T}(T^*))] \sim f_n \frac{d(n!)^d}{(d-1)n} \sim f_n \frac{d}{(d-1)n} (n/e)^{dn} (2\pi n)^{d/2}.$$

Dividing this by $g(dn)$ and applying Stirling's formula gives the expression for $\mathbf{E}[H(\mathcal{T}(T^*))]$ stated in the theorem. To show that $\hat{f}_n \sim f_n$, we may just note that the same formula results, by the same argument, with $f(x_1, \ldots, x_n)$ defined to be any particular diagonal element of the matrix $(I_d - AD)^{-1}$. ∎

To estimate f_n, we will apply the part of [9, Theorem 3.5] referring to the leading asymptotic term of a ratio of analytic functions (in the case of non-zero numerator at the singularity of interest). This requires some preliminary setting. Suppose that G and K are analytic functions of $\mathbf{y} = (y_1, \ldots, y_d)$ in a neigbourhood of the toroidal region

$$U_{\mathbf{z}} = \{\mathbf{y} \, : \, |y_j| \le |z_j|, \, 1 \le j \le d\}$$

where $\mathbf{z} = (z_1, \ldots, z_d)$, is the unique singular point of G/K in U. We use the suffix j to denote partial differentiation with respect to y_j. Suppose that $K_d \neq 0$. Then (using the implicit function theorem) solving $K(\mathbf{y}) = 0$ yields $y_d = h(y_1, \ldots, y_{d-1})$ with h analytic near $\mathbf{z}$. As in Section 1 before the statement of Theorem 4, let $\tilde{f}(\theta_1, \ldots, \theta_{d-1}) = \log h(z_1 e^{i\theta_1}, \ldots, z_{d-1} e^{i\theta_{d-1}})$, and let $\mathcal{H}$ be the determinant of the Hessian matrix of $\tilde{f}$ at $(\theta_1, \ldots, \theta_{d-1}) = (0, \ldots, 0)$. Define $\mathrm{dir}(\mathbf{z})$ to be the set of complex scalar multiples of $(z_1 K_1(\mathbf{z}), \ldots, z_d K_d(\mathbf{z}))$. (The reader may notice that $\tilde{f}$ differs from that in [9] by a function which is linear in the θ_i. The definitions are equivalent for the current purpose because $\tilde{f}$ enters the following formula only through $\mathcal{H}$.)

Theorem 6 *[9] With the assumptions and definitions in the paragraph above, if $G(\mathbf{z}) \neq 0$ and $\mathcal{H} \neq 0$ then for $\mathbf{r} \in dir(\mathbf{z})$*

$$[y_1^{r_1} \cdots y_d^{r_d}] \, \frac{G(\mathbf{y})}{K(\mathbf{y})} \sim (2\pi)^{(1-d)/2} \frac{G(\mathbf{z})}{-z_d K_d(\mathbf{z}) \sqrt{\mathcal{H}}} \prod_{j=1}^{d} z_j^{-r_j}.$$

Proof of Theorem 4 Define $\tilde{B} = I_d - AD$. Write $C_{[kl]}$ for the result of deleting row k and column l from a matrix C, and let ϕ denote the formal power series homomorphism generated by

$$\phi(x_j) = \frac{y_j}{1 - \frac{1}{2} y_j}$$

for all j. Extend ϕ in the natural way to act on matrices by acting on all elements of the matrix. Since $\tilde{B}^{-1} = (\det \tilde{B})^{-1} \mathrm{adj}\ \tilde{B}$ and $f(x_1, \ldots, x_d)$ is the $(1,1)$ entry of $\tilde{B}^{-1}$, we have

$$f(x_1, \ldots, x_d) = \frac{\det \tilde{B}_{[11]}}{\det \tilde{B}}.$$

Applying ϕ to both sides,

$$\begin{aligned} \phi(f(x_1, \ldots, x_d)) &= \frac{\det \phi(\tilde{B}_{[11]})}{\det \phi(\tilde{B})} \\ &= \frac{(1 - \frac{1}{2}y_1) \det B_{[11]}}{\det B} \end{aligned}$$

where the last step uses the fact that

$$B = \phi(\tilde{B}) \operatorname{diag}(1 - \tfrac{1}{2}y_1, \ldots, 1 - \tfrac{1}{2}y_d). \tag{2.11}$$

It thus follows from (1.1) that

$$f_n = [y_1^n \cdots y_d^n] \frac{(1 - \frac{1}{2}y_1) \det B_{[11]}}{\det B}. \tag{2.12}$$

To apply this result, we must find a suitable singular point $\mathbf{z}$ of the function in (2.12), so that $\mathrm{dir}(\mathbf{z})$ contains $(n, n, \ldots, n)$. First note that $\det B_{[11]}$ and $\det B$ are both polynomials, and hence entire functions. Moreover, since T^* is regular, all the row sums of B are equal to 0 when $(y_1, \ldots, y_d) = \mathbf{z}$ where $\mathbf{z}$ was defined in (1.2). Thus $\det B(\mathbf{z}) = 0$, which makes $\mathbf{z}$ a singularity of the rational function in (2.12) unless the numerator is also 0 at $\mathbf{z}$.

Note that $\frac{d}{2}B(\mathbf{z})$ is the Laplacian matrix of T^*, and so by the Matrix-Tree Theorem [11], $\det \frac{d}{2}B_{[11]}(\mathbf{z})$ is the number of directed spanning trees in T^* rooted at vertex 1. Since T^* is a regular tournament, it is strongly connected. (This follows because, for any partition of the vertex set into two nonempty parts U and V, there must be an equal number of edges directed from U to V as from V to U.) Hence the number of directed spanning trees with any given root vertex is non-zero. Thus the numerator in (2.12) is non-zero at $\mathbf{z}$, and so $\mathbf{z}$ is indeed a singularity.

To show that there are no other singularities in an appropriate toroidal region, it suffices to show that $\det B \neq 0$ at an arbitrary point $\mathbf{y} = (y_1, \ldots, y_d)$

where $|y_j| \le z$ for each j, and $|y_k| < z_k$ for some k. To this end, let $D = \mathrm{diag}(y_1, \dots, y_d)$, $\hat{D} = \mathrm{diag}(|y_1|, \dots, |y_d|)$, $D_z = \mathrm{diag}(z, \dots, z)$ and $\hat{A} = \frac{1}{2}I_d + A$. Then $B = I_d - \hat{A}D$. Assume that $\det B = 0$. It follows that 1 is an eigenvalue of $\hat{A}D$. Let $\mathbf{v}$ be an associated eigenvector. Then

$$|\hat{A}D\mathbf{v}| = |\mathbf{v}|$$

where $|C|$ denotes the matrix whose entries are the absolute values of the entries in C. Hence

$$\hat{A}\hat{D}|\mathbf{v}| \ge |\mathbf{v}|,$$

where $\ge$ denotes elementwise comparison. That is, $\mathbf{v}$ is a 1-subharmonic vector for $\hat{A}\hat{D}$. So by [4, Lemma 8.7.1] the largest eigenvalue of $\hat{A}\hat{D}$ is at least 1. Note that $\hat{A}D_z$, $\hat{A}\hat{D}$ and $\hat{A}D_z - \hat{A}\hat{D}$ are all nonegative real matrices, $\hat{A}D_z \ne \hat{A}\hat{D}$, and furthermore the underlying directed graph of $\hat{A}$ consists of T^* with a loop at each vertex, which is strongly connected. Thus by the Perron-Frobenius Theorem (see [4, Theorem 8.8.1] in particular), the maximum eigenvalue of $\hat{A}\hat{D}$ is strictly less than that of $\hat{A}D_z$, which must therefore be strictly greater than 1. On the other hand, it is exactly 1 because all row sums of $\hat{A}$ are $d/2 = 1/z$. This is a contradiction, so $\det B \ne 0$ at $\mathbf{y}$.

Furthermore, with $K(\mathbf{y}) = \det B$, we need to find $K_j(\mathbf{z})$ for each j. Note that $K_j(\mathbf{z}) = \det \frac{\partial}{\partial y_j} B(\mathbf{z})$. But $\frac{\partial}{\partial y_j} B(\mathbf{z})$ and $B(\mathbf{z})$ only differ in column j, which in the former has -1 in the off-diagonal entries where the latter has $-2/d$, and $-\frac{1}{2}$ in the diagonal entry where the latter has $1 - 1/d$. Thus, expanding the determinants down column j shows that

$$\begin{aligned} K_j(\mathbf{z}) &= \frac{d}{2}\det B(\mathbf{z}) - \frac{d}{2}\det B_{[jj]}(\mathbf{z}) \\ &= -\frac{d}{2}\det B_{[jj]}(\mathbf{z}). \end{aligned} \tag{2.13}$$

Since T^* is regular it is balanced, and so by the BEST theorem (see [5] for example), the number of directed spanning trees rooted at vertex j is independent of j (this independence can alternatively be derived from $A \cdot \mathrm{adj}A = \mathbf{0}$), and is equal to $\frac{d-1}{2}\det B_{[jj]}(\mathbf{z})$. Thus $K_j(\mathbf{z})$ is independent of j, in particular $K_d(\mathbf{z}) \ne 0$ which completes verification of the hypotheses of Theorem 6. This independence of j also shows that $\mathbf{dir}(\mathbf{z})$ contains the vector $(1,1,\dots,1)$ as desired. By (2.13), it also shows that $-zK_d(\mathbf{z})$ cancels with $\det B_{[11]}(\mathbf{z})$ in the numerator (here $G(\mathbf{z}) = (1-z/2)\det B_{[11]}(\mathbf{z})$). Thus by Theorem 6

$$f_n \sim \left(\frac{d}{2}\right)^{dn} (2\pi n)^{-(d-1)/2} \frac{(1-z/2)}{\sqrt{\mathcal{H}}}. \tag{2.14}$$

The theorem now follows from Theorem 3 on recalling that $z = 2/d$. ∎

Proof of Theorem 5 We modify the treatment of the random variable H, which counts cycles, for H', which counts paths. Define $f^{(kl)}(x_1,\dots,x_n)$ to be the (k,l) entry of $(I_d - AD)^{-1}$. Then the argument in Lemma 1 applies for Hamilton paths which begin at a vertex of T_k and end at T_l, with the modifications that W is no longer necessarily closed, there is no overcounting by j_1, and there is an initial path in T_k corresponding to the first vertex of W. This applies even if W happens to be closed; the initial and final vertices of W correspond to paths which are essentially independent. So we have

$$\mathbf{E}[H'(\mathcal{R}(T^*))] = (n!2^{-n})^d \sum_{1\le k,l\le d} \sum_{j_1,\dots,j_d\ge 1} [x_1^{j_1}\cdots x_d^{j_d}] x_k f^{(kl)}(x_1,\dots,x_d) \prod_{i=1}^{d} \binom{n-1}{j_i-1} 2^{j_i}.$$

Thus (c.f. Theorem 3 and (2.12))

$$\mathbf{E}[H'(\mathcal{R}(T^*))] \sim g(dn)\left(\frac{2}{d}\right)^{dn} (2\pi n)^{(d-1)/2} d^{1/2} n \sum_{1\le k,l\le d} f_n^{(kl)}$$

where

$$\begin{aligned} f_n^{(kl)} &= [y_1^n\cdots y_d^n]\frac{y_k}{1-\frac{1}{2}y_k} f\big(1-\tfrac{1}{2}y_1,\dots,1-\tfrac{1}{2}y_d\big) \\ &= [y_1^n\cdots y_d^n]\frac{y_k \det B_{[kl]}}{\det B}. \end{aligned}$$

The Matrix-Tree Theorem (or, again, linear algebra) implies that $\det B_{[kl]} = \det B_{[kk]}$, which, as already noted in the proof of Theorem 4, is equal to $B_{[11]}$. Hence, applying the proof of Theorem 4, we may replace y_k by z and we find $f_n^{(kl)} \sim zf_n/(1-\frac{1}{2}z)$. So in summary, $\mathbf{E}[H'(\mathcal{R}(T^*))]$ (in comparison with $\mathbf{E}[H(\mathcal{R}(T^*))]$) gains factors of $n(d-1)/d$ (from the missing j_1^{-1}) and

$z/(1-\frac{1}{2}z) = 2/(d-1)$, and the factor of d^2 for summing over k and l. Since $g'(dn)/g(dn) \sim 2dn$, the claimed path analogue of Theorem 4 follows.

For the path analogue of Theorem 1, the proof goes through with no more changes, using Lemma 3. ∎

References

[1] I. Adler, N. Alon and S. M. Ross, On the maximum number of Hamiltonian paths in tournaments, *Random Structures and Algorithms* **18** (2001), 291-296.

[2] N. Alon, The maximum number of Hamiltonian paths in tournaments, *Combinatorica* **10** (1990), 319-324.

[3] E. Friedgut and J. Kahn, On the number of Hamiltonian cycles in a tournament (manuscript).

[4] C. Godsil and G. Royle, *Algebraic Graph Theory.* Graduate Texts in Mathematics, 207. Springer-Verlag, New York, 2001.

[5] I.P. Goulden and D.M. Jackson, *Combinatorial Enumeration.* Wiley, New York (1983).

[6] B.D. McKay, The asymptotic numbers of regular tournaments, eulerian digraphs and eulerian oriented graphs, *Combinatorica* **10** (1990), 367–377.

[7] B.D. McKay and R.W. Robinson, Asymptotic enumeration of Eulerian circuits in the complete graph, *Combin. Probab. Comput.* **7** (1998), 437–449.

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

[9] R. Pemantle and M. Wilson, Asymptotics of multivariate sequences, part I: smooth points of the singular variety, *J. Combinatorial Theory, Series A* **97** (2001), 129–161.

[10] T. Szele, Kombinatorische Untersuchungen ber den gerichteten vollstndigen Graphen (Hungarian), *Mat. Fiz. Lapok* **50** (1943), 223–256.

[11] W.T. Tutte, The dissection of equilateral triangles into equilateral triangles, *Proc. Cambridge Philos. Soc.* **44** (1948), 463–482.