On the longest path of a randomly weighted tournament

Raphael Yuster *

Abstract

Finding a maximum weight path in a directed or undirected graph is a basic combinatorial and algorithmic problem. We consider this problem for tournaments in the well-studied random weighted model. Denote by $\mathcal{P}(G,\mathcal{D})$ the probability space obtained by independently assigning weights to the edges of a tournament G according to a nonnegative probability distribution $\mathcal{D}$. Denote by $\ell(G,\mathcal{D})$ the expected weight of a path with maximum weight.

If $\mathcal{D}$ has finite mean μ, then $\ell(G,\mathcal{D}) \ge \mu(n-1)$ is a trivial lower bound, with equality if $\mathcal{D}$ is constant, as by Redei's Theorem, every tournament has a Hamilton path. However, already for very simple nontrivial distributions, it is challenging to determine $\ell(G,\mathcal{D})$ even asymptotically, and even if the tournament is small and fixed. We consider the two natural distributions of the random weighted model, the continuous uniform distribution $U[0,1]$ and the symmetric Bernoulli distribution $U\{0,1\}$. Our first result is that for *any* tournament, both $\ell(G,U\{0,1\})$ and $\ell(G,U[0,1])$ are larger than the above trivial $0.5(n-1)$ lower bound in the sense that 0.5 can be replaced by a larger *constant*. To this end we prove the existence of dense partial squares of Hamilton paths in any tournament, a combinatorial result which seems of independent interest. Regarding upper bounds, while for some tournaments one can prove that both $\ell(G,U\{0,1\})$ and $\ell(G,U[0,1])$ are $n-o(n)$, we prove that there are other tournaments for which both $\ell(G,U\{0,1\})$ and $\ell(G,U[0,1])$ are significantly smaller. In particular, for every n, the are n-vertex tournaments for which $\ell(G,U\{0,1\}) \le 0.614(n-1)$. Finally, we state several natural open problems arising in this setting.

1 Introduction

Finding the longest (simple) path in a weighted/unweighted graph/digraph is a central problem in graph theory both in the extremal setting and the algorithmic setting. In particular, already deciding if a given graph or digraph has a Hamilton path is an NP-Hard problem. In this paper we look at this problem in the random-weighted model. In this model we assign weights to the edges where the weights are i.i.d. random variables with a given distribution $\mathcal{D}$. This is a well-studied model that has been used by many researchers in various settings, see [3, 8]. We next mention a few notable results in this area.

Walkup [16] proved that the expected value of a minimum weight perfect matching in a randomly weighted complete bipartite graph is bounded by a constant. By "randomly weighted" one assumes

*Department of Mathematics, University of Haifa, Haifa 31905, Israel. Email: raphy@math.haifa.ac.il

that $\mathcal{D} = U[0,1]$ is the uniform distribution in $[0,1]$. Later, Aldous [1] proved that the expected value converges to $\zeta(2) = \sum_{k=1}^{\infty} k^{-2} = \pi^2/6$, see also [17] and the references therein. Frieze [6] determined the expected value of a minimum spanning tree of a randomly weighted complete graph where $\mathcal{D}$ belongs to a large class of distributions. In particular, for $\mathcal{D} = U[0,1]$ it turns out that this value converges to $\zeta(3) = \sum_{k=1}^{\infty} k^{-3} \sim 1.202$. See also [5] for a sharpened result. In [12], Karp considered the expected weight of a minimum length Hamilton cycle in the complete directed graph. Frieze considered the same problem in the complete (undirected) graph [7]. Hassin and Zemel [10] considered the case of weights of shortest paths in complete undirected or directed graphs where here again $\mathcal{D} = U[0,1]$. They proved that the length of the longest shortest path is $O(\log n/n)$ almost surely. Resolving a long-standing open problem, Peres et al. [13] designed a probabilistic algorithm for the all-pairs shortest paths problem which runs in $O(n^2)$ expected time, in the case of a randomly weighted complete directed graph and $\mathcal{D} = U[0,1]$. See their paper for many other references to related results.

While in all of the results mentioned above one tries to minimize an objective function (be it defined on a path, spanning tree, or Hamilton cycle), in our problem we try to maximize the length of a path. It is straightforward to verify that if $\mathcal{D} = U[0,1]$, then the expected length of a longest path in the complete directed or undirected graph with n vertices is $n - o(n)$. This is not surprising as in both of these graphs, any permutation of vertices corresponds to a Hamilton path. However, things become considerably more involved when one limits the set of possible Hamilton paths. An interesting class of graphs for this problem is the class of all orientations of complete graphs, namely the class of tournaments. In fact, as we show in this paper, tournaments form a very rich class with respect to the longest path in the random weighted model. We next give the formal definitions.

A *tournament* $G = (V,E)$ is a digraph such that for every pair of distinct vertices u, v, the edge set E contains exactly one edge with ends $\{u,v\}$. In other words, either $(u,v) \in E$ or $(v,u) \in E$ is present, but not both. If $G = (V,E)$ is a tournament, we say that $X \subseteq V$ is *transitive* if the sub-tournament $G[X]$ induced on X has no directed cycle. We denote by T_n the (unique) transitive tournament on n vertices. We shall usually denote the set of vertices by $\{1,\ldots,n\}$ and if the tournament is transitive, we shall assume that $(i,j) \in E$ whenever $i < j$. Denote by $\mathcal{P}(G,\mathcal{D})$ the probability space obtained by independently assigning weights to the edges of G according to a nonnegative probability distribution $\mathcal{D}$. The weight of a path is the sum of the weights of its edges. Denote by $X(G,\mathcal{D})$ the random variable corresponding to the weight of a path with maximum weight and let $\ell(G,\mathcal{D}) = \mathbb{E}[X(G,\mathcal{D})]$ be the expectation of $X(G,\mathcal{D})$. To avoid trivial cases we assume that $\mathbb{E}[\mathcal{D}]$ is finite, and so $\ell(G,\mathcal{D})$ is well-defined. Let $\mathcal{D}_{max}$ be the supremum of all values attained by $\mathcal{D}$. Thus, we have the following bounds:

$$(n-1)\mathbb{E}[\mathcal{D}] \le \ell(G,\mathcal{D}) \le (n-1)\mathcal{D}_{max} \ .$$

The upper bound follows trivially from the fact that any path has at most $n-1$ edges and the lower bound follows from Redei's Theorem [15] that every tournament has a Hamilton path, so by linearity of expectation, the expected weight of a Hamilton path is $(n-1)\mathbb{E}[\mathcal{D}]$. Notice that if $\mathcal{D}$ is constant, then the bounds are tight. As is common in the random weight model, we shall consider the most natural distributions $U[0,1]$ (the continuous uniform distribution on the unit interval) and $U\{0,1\}$, the Bernoulli distribution with success probability 0.5. So the above inequalities translate to

$$\frac{n-1}{2} \le \ell(G, U[0,1]), \ell(G, U\{0,1\}) \le (n-1) \ . \tag{1}$$

It turns out that computing $\ell(G, U[0,1])$ and $\ell(G, U\{0,1\})$ is difficult already for small fixed tournaments, moreover determining such values asymptotically. As we shall see, distinct tournaments with n vertices may exhibit very different values. Hence, we are interested in the extremal setting. Let $\ell(n, \mathcal{D}) = \min_{|G|=n} \ell(G, \mathcal{D})$ and let $\ell_{max}(n, \mathcal{D}) = \max_{|G|=n} \ell(G, \mathcal{D})$. It is not difficult to prove that $\ell_{max}(n, U[0,1])$ and $\ell_{max}(n, U\{0,1\})$ get close to the upper bound in (1) as shown in the following proposition.

Proposition 1.1

$$\lim_{n\to\infty} \frac{\ell_{max}(n, U\{0,1\})}{n-1} = \lim_{n\to\infty} \frac{\ell_{max}(n, U[0,1])}{n-1} = 1 \ .$$

It is more intriguing, however, to determine the behavior of $\ell(n, U[0,1])$ and $\ell(n, U[0,1])$. To this end, define

$$\begin{aligned} \beta(\mathcal{D}) &= \liminf_{n\to\infty} \frac{\ell(n, \mathcal{D})}{n-1} \ . \\ \beta_{tr}(\mathcal{D}) &= \lim_{n\to\infty} \frac{\ell(T_n, \mathcal{D})}{n-1} \ . \end{aligned}$$

By their definition and by (1) we have $0.5 \le \beta(U[0,1]) \le \beta_{tr}(U[0,1]) \le 1$ and similarly for $U\{0,1\}$. We can, however, significantly improve these bounds.

Theorem 1 *For $U[0,1]$ we have:*

$$\frac{85}{168} \le \beta(U[0,1]) \quad , \quad 0.525 \le \beta_{tr}(U[0,1]) < \frac{2}{3} \ .$$

For $U\{0,1\}$ we have:

$$\frac{29}{56} \le \beta(U\{0,1\}) \quad , \quad 0.595 \le \beta_{tr}(U\{0,1\}) \le 0.614 \ .$$

The proof of the lower bounds for β in both cases means that for *every* tournament we can significantly improve upon the $0.5(n-1)$ lower bound of (1). To the proofs end, we establish an extremal result concerning the existence of dense partial squares of Hamilton paths that exist in every tournament. This result, which may be of independent interest, as well as the lower bound for β, are established in Section 2. One may stipulate that $\beta_{tr} = \beta$ as it is plausible (but not obvious) that transitive tournaments are the "worst" w.r.t. minimizing the expected maximum length. Thus, it seems of interest to determine β_{tr}. For the lower bound of $\beta_{tr}(U[0,1])$ and $\beta_{tr}(U\{0,1\})$, proved in Section 3, we first compute precisely some exact values of $\ell(T_n, U[0,1])$ and $\ell(T_n, U\{0,1\})$. This task turns out to be rather involved already for very small n. In Section 4 we prove the upper bound of $\beta_{tr}(U[0,1])$ and $\beta_{tr}(U\{0,1\})$ as well as an upper bound for a wider class of distributions. To this end we use large deviation inequalities, the properties of the Irwin-Hall distribution, and some recursive analysis. Notice in particular that our lower and upper bounds for $\beta_{tr}(U\{0,1\})$ are very close. Proposition 1.1 is proved in Section 5. We end the paper by stating some conjectures and open problems that arise naturally in our setting.

2 Partial squares of Hamilton paths and a lower bound for β

For an integer $k \geq 1$, a *k'th power* of a Hamilton path is a Hamilton path ordered by $v_1, \ldots, v_n$ where for all $1 \leq i < j \leq n$ with $j \leq i + k$, (v_i, v_j) is an edge. In other words, every sequence of $k+1$ consecutive vertices on the Hamilton path induces the transitive tournament T_{k+1}. The case $k = 2$ is also called a *square* of a Hamilton path.

It has been proved by Bollobás and Häggkvist [4] that for every $\epsilon > 0$ and $k \in \mathbb{N}$, every sufficiently large n-vertex tournament with minimum in-degree and minimum out-degree at least $n(\frac{1}{4} + \epsilon)$ has a k'th power of a Hamilton cycle (and hence a k'th power of a Hamilton path). Unfortunately, however, while every tournament has a Hamilton path, it is no longer true that every tournament (no matter how large) has a square of a Hamilton path, as shown by the following construction. Suppose that 3 divides n and consider the tournament on vertex sets $V_1, \ldots, V_{n/3}$. For all $1 \leq i < j \leq n/3$, there is an edge from each vertex of V_i to each vertex of V_j. Furthermore, each V_i consists of three vertices inducing a directed triangle. This tournament has n vertices, and it is straightforward that in each Hamilton path, all vertices of V_i precede all vertices of V_j whenever $i < j$. But for any such Hamilton path, the partial path on V_i "misses" an edge in the square. Hence, altogether, at least $n/3$ edges are missed. If 3 does not divide n, then the last set may have fewer than three vertices and $\lfloor n/3 \rfloor$ edges are missed.

Thus, the best we can ask for in general is a partial square of a Hamilton path, formally defined as follows. An *α-square* of a Hamilton path is a Hamilton path such that at least an α fraction of the consecutive sequences of three vertices on the path induce a T_3. So, a 1-square is just a (usual) square of a Hamilton path. The extremal function of interest here is to determine α^*, the supremum over all α's such that any sufficiently large tournament has an α-square of a Hamilton path. The construction in the previous paragraph shows that $\alpha^* \leq 2/3$. The next lemma supplies a nontrivial lower bound for α^*. This lower bound will be useful in the rest of this paper.

A *shortcut triple* on a path is a set of three consecutive vertices on the path which induces a T_3.

Lemma 2.1 *Every tournament has a Hamilton path with at least $n/7 - 1$ pairwise edge-disjoint shortcut triples.*

Proof. Let t be the largest integer such that there is a path P with t vertices and with at least $t/3$ shortcut triples. We will first prove that P (which is not necessarily a Hamilton path) has at least $n/7 - 1$ pairwise edge-disjoint shortcut triples. Let $v_1, v_2, \ldots, v_t$ denote the vertices of P and let $W = V(G) \setminus \{v_1, \ldots, v_t\}$ where G denotes the tournament. Consider any $w \in W$. We say that w has *type* t if $(v_t, w) \in E(G)$. We say that w has type 0 if $(w, v_1) \in E(G)$. We say that w has type

i if $(v_i, w) \in E(G)$ and $(w, v_{i+1}) \in E(G)$ for some $i = 1, \ldots, t-1$. Observe that each $w \in E(G)$ has at least one type out of the $t+1$ possible types.

If there are two vertices of type t we are done. Indeed, say these are w, w' and suppose that $(w, w') \in E(G)$. By defining $v_{t+1} = w$ and $v_{t+2} = w'$ we increase the number of shortcut triples by 1 and the number of vertices on P only by 2, contradicting the maximality of t. Similarly, if there are two vertices of type 0 we are done. Say these are w, w' and $(w, w') \in E(G)$. So defining $v_{-1} = w$ and $v_0 = w'$ we increase the number of shortcut triples by 1 and the number of vertices on P only by 2, again contradicting the maximality of t.

We may now assume that there is at most one vertex in W of type 0 and at most one vertex in W of type t. Hence at least $|W| - 2$ vertices of W have types that are distributed over the $t-1$ types $1, \ldots, t-1$. Let Q be a maximum set of pairwise edge-disjoint shortcut triples on the path. Trivially, $|Q| \leq (t-1)/2$ as the triples are pairwise edge-disjoint. By the definition of t we also have $|Q| \geq t/3$. Let $X \subset \{v_1, \ldots, v_{t-1}\}$ where $v_i \in X$ if (v_i, v_{i+1}) does not belong to a shortcut triple of Q. Notice that $|X| = t - 1 - 2|Q|$.

If some $w \in W$ has type i where $v_i \in X$ we are done. Indeed, placing w between v_i and v_{i+1} increases the length of P by 1 and increases the number of shortcut triples by 1 since now v_i, w, v_{i+1} also form a shortcut triple. This contradicts the maximality of t. We may therefore assume that there are $|W| - 2$ vertices of W distributed over $t - 1 - |X| = 2|Q|$ types, where these types correspond to indices $1 \leq i \leq t-1$ and $v_i \notin X$. Now, assume there are three vertices with the same type i, say x, y, z, where, without loss of generality, $(x, y) \in E(G)$ and $(y, z) \in E(G)$. Then we can extend P by adding x, y, z between v_i and v_{i+1} so that the sub-path is now (v_i, x, y, z, v_{i+1}), and we have formed two edge-disjoint shortcut triples v_i, x, y and y, z, v_{i+1} whereas we have deleted at most one shortcut triple of Q (the shortcut triple that contains the edge (v_i, v_{i+1})). Notice that we have added only three vertices and gained one shortcut triple, contradicting the maximality of t.

So, there are at most two vertices with the same given type out of at most $2|Q|$ possible types, and hence $|W| - 2 = n - 2 - t \leq 4|Q| < 4(n/7 - 1)$ where the latter inequality assumes $|Q| < n/7 - 1$ as otherwise we are done. We therefore obtain that $t > 3n/7 + 2$ and thus $t/3 > n/7$ and we are done.

The only remaining issue is that P is not necessarily a Hamilton path (namely, it may be that $t < n$). Let W and Q be defined as above. If $W = \emptyset$ we are done since then P is a Hamilton path. Otherwise, let $w \in W$ and consider its type, as defined above. If w's type is 0 we may extend the path by adding w before v_1, without decreasing the number of shortcut triples. If w's type is t we may extend the path by adding w after v_t, without decreasing the number of shortcut triples. If w's type is i where $1 \leq i \leq t-1$, we may extend the path by adding w between v_i and v_{i+1}. We may have destroyed at most one shortcut triple in Q (one that contains the edge (v_i, v_{i+1})) but we have introduced instead a new shortcut triple v_i, w, v_{i+1} which is edge disjoint from all other remaining shortcut triples in Q. Thus, we have not decreased the number of shortcut triples. Continuing in this fashion, we extend P until it becomes a Hamilton path. ■

We note that the constant $\frac{1}{7}$ in Lemma 2.1 can be slightly improved using a somewhat more involved proof. This is because we can start constructing P with a denser fraction of shortcut triples using the fact that for any k, any sufficiently large tournament contains T_k, so we may pick the vertices of such a T_k as vertices of the initial P and also continue to extend this P somewhat using a slightly better than $1:3$ ratio of vertices versus disjoint shortcut triples. However, as the proof becomes more involved and the improvement will not be dramatic, we prefer the simpler proof. Observe also that Lemma 2.1 guarantees a lower bound for the stronger property of having pairwise edge-disjoint shortcut triples, while in the definition of α^* the shortcut triples are not required to be pairwise edge-disjoint. In any case, we have the following corollary.

Corollary 2.2

$$\frac{1}{7} < \alpha^* \le \frac{2}{3} \,.$$

Corollary 2.3

$$\frac{85}{168} \le \beta(U[0,1]) \quad , \quad \frac{29}{56} \le \beta(U\{0,1\}) \,.$$

Proof. Let G be an arbitrary tournament with n vertices, and let P be a Hamilton path with k pairwise edge-disjoint shortcut triples. By Lemma 2.1 we may assume that $k \ge n/7 - 1$. Then this path can be partitioned into $n-1-k$ sub-paths, say $P_1, \ldots, P_{n-1-k}$ where each P_i is either a single edge (u_i, v_i) or a shortcut triple (u_i, x_i, v_i). The last vertex of each P_i is the first vertex of P_{i+1}, namely $v_i = u_{i+1}$, for $i = 1, \ldots, n-2-k$. Now, clearly, $\ell(G, \mathcal{D})$ is at least as large as the expected weight of P. The latter, by linearity of expectation, is at least the sum of the expected weights of the P_i. As there are $n-1-2k$ of the P_i's that are isomorphic to T_2 and there are k of the P_i's that are isomorphic to T_3 we have that

$$\ell(G, \mathcal{D}) \ge (n-1-2k)\mathbb{E}[\mathcal{D}] + k \cdot \ell(T_3, \mathcal{D}) \,.$$

Consider first $\mathcal{D} = U[0,1]$. In this case, $\mathbb{E}[\mathcal{D}] = \frac{1}{2}$ and, as shown in the next section, $\ell(T_3, U[0,1]) = \frac{25}{24}$. Using $k \ge n/7 - 1$ we obtain

$$\ell(G, U[0,1]) \ge \frac{n-1-2k}{2} + \frac{25k}{24} \ge \frac{85}{168}(n-1) - \frac{1}{28} \,.$$

By the definition of $\beta(U[0,1])$ as a limit, it follows that $\beta(U[0,1]) \ge \frac{85}{168}$.

Consider next $\mathcal{D} = U\{0,1\}$. In this case, $\mathbb{E}[\mathcal{D}] = \frac{1}{2}$ and, as shown in the next section, $\ell(T_3, U\{0,1\}) = \frac{9}{8}$. Using $k \ge n/7 - 1$ we obtain

$$\ell(G, U\{0,1\}) \ge \frac{n-1-2k}{2} + \frac{9k}{8} \ge \frac{29}{56}(n-1) - \frac{3}{28} \,.$$

By the definition of $\beta(U\{0,1\})$ as a limit, it follows that $\beta(U\{0,1\}) \ge \frac{29}{56}$. ■

n	3	4	5	6	7	8
$\ell(T_n, U\{0,1\})$	$\frac{9}{8}$	$\frac{111}{64}$	$\frac{2399}{1024}$	$\frac{96735}{32768}$	$\frac{7468479}{2097152}$	$\frac{1119481727}{268435456}$
$\frac{\ell(T_n, U\{0,1\})}{n-1} \geq$	0.562	0.578	0.585	0.59	0.593	0.595

Table 1: Small values of $\ell(T_n, U\{0,1\})$.

3 Lower bounds for transitive tournaments

Lemma 3.1 *Let $\mathcal{D}$ be a nonnegative distribution with finite expectation. If $k-1$ divides $n-1$, then $\ell(T_n, \mathcal{D}) \geq \frac{n-1}{k-1}\ell(T_k, \mathcal{D})$. In particular, regardless of divisibility,*

$$\ell(T_n, \mathcal{D}) \geq (1-o(1))(n-1)\frac{\ell(T_k, \mathcal{D})}{k-1} .$$

Proof. Since $\mathcal{D}$ is nonnegative, there is a maximum weight path in T_n that starts at vertex 1 and ends at vertex n. To see this, notice that any path that starts at vertex i and ends at vertex j can be extended to a path from 1 to n by starting with a path from vertex 1 to i and ending with a path from vertex j to n. The weight of the extended path is at least as large as the original one.

Now, suppose that $k-1$ divides $n-1$. Let Y be the random variable corresponding to the maximum weight of all paths that go through all the vertices $i(k-1)+1$ for $i = 0, \ldots, (n-1)/(k-1)$. By definition, $Y \leq X(T_n, \mathcal{D})$ and thus $\mathbb{E}[Y] \leq \ell(T_n, \mathcal{D})$. By linearity of expectation $\mathbb{E}[Y] = \frac{n-1}{k-1}\ell(T_k, \mathcal{D})$. ■

Observe that Lemma 3.1 implies, in particular, that $\beta_{tr}(\mathcal{D}) = \lim_{n\to\infty} \frac{\ell(T_n,\mathcal{D})}{n-1}$ exists, justifying the comment in the introduction.

We next establish $\ell(T_n, U\{0,1\})$ and $\ell(T_n, U[0,1])$ for some small values of n. Lemma 3.1 shows that any such bound can be used as a lower bound for the corresponding β_{tr}.

We start with the discrete case, as it is simpler and one can explicitly compute $\ell(T_n, U\{0,1\})$ for some small n. For convenience, set $X_n = X(T_n, U\{0,1\})$. We notice that $\Pr[X_n = n-1] = 2^{1-n}$ as the only option to obtain the weight $n-1$ is through the unique Hamilton path of T_n. Also, $\Pr[X_n = 0] = 2^{-\binom{n}{2}}$ as for this to happen, all edges must receive weight 0. These observations already give $\ell(T_3, U\{0,1\}) = 9/8$.

The case $n = 4$ is determined manually as follows. By the above, $\Pr[X_4 = 0] = 1/64$ and $\Pr[X_4 = 3] = 1/8$. In order to have $X_4 = 2$ exactly one of the following must occur:

(i) Precisely two of the edges of the Hamilton path have weight 1.

(ii) Edge $(1,2)$ is the only edge on the Hamilton path with weight 1 and $(2,4)$ also has weight 1.

(iii) Edge $(3,4)$ is the only edge on the Hamilton path with weight 1 and $(1,3)$ also has weight 1.

Hence $\Pr[X_4 = 2] = 3/8+1/16+1/16 = 1/2$. This leaves $\Pr[X_4 = 1] = 23/64$ hence $\ell(T_4, U\{0,1\}) = 111/64 \sim 1.734$.

Larger exact values of $\ell(T_n, U\{0,1\})$ are hard to compute manually, but can be easily evaluated by a computer, by considering all possible $\{0,1\}$ weighings of T_n. This is feasible for all $n \leq 8$ (where T_8 has 2^{28} weighings). The values are summarized in Table 1.

The continuous case is harder already for very small n. Again, for convenience, set $X_n = X(T_n, U[0,1])$. We start by determining the distribution of X_3.

Lemma 3.2 $\ell(T_3, U[0,1]) = \frac{25}{24}$.

Proof. Clearly, $X_3 = \max\{Y, Z_3\}$ where $Y = Z_1 + Z_2$ and Z_1, Z_2, Z_3 are independent random variables distributed uniformly in $[0,1]$. The probability density function of Y is the result of an easy convolution (for example, see [9], page 291) and is:

$$f_Y(t) = \begin{cases} t & \text{if } 0 \le t \le 1 \\ 2-t & \text{if } 1 \le t \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

Since

$$\Pr[X_3 \le \alpha] = \Pr[Y \le \alpha] \cdot \Pr[Z_3 \le \alpha]$$

we have that for $0 \le \alpha < 1$,

$$\Pr[X_3 \le \alpha] = \left[\int_0^\alpha t\,dt\right] \cdot \left[\int_0^\alpha 1\,dt\right] = \frac{\alpha^3}{2}$$

and for $1 \le \alpha \le 2$,

$$\Pr[X_3 \le \alpha] = \left[\int_0^1 t\,dt + \int_1^\alpha (2-t)\,dt\right] = 2\alpha - 1 - \frac{\alpha^2}{2}\ .$$

Hence, the probability density function of X_3 is:

$$f_{X_3}(t) = \begin{cases} 1.5t^2 & \text{if } 0 < t < 1 \\ 2-t & \text{if } 1 < t < 2 \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$\ell(T_3, U[0,1]) = \int_0^1 1.5t^3\,dt + \int_1^2 (2-t)t\,dt = \frac{25}{24}\ .$$

■

The case $n = 4$ is quite involved, as demonstrated in the proof of the following lemma.

Lemma 3.3 $\ell(T_4, U[0,1]) = \frac{7949}{5040}$.

Proof. Notice that $X_4 = \max\{Z_1+Z_2+Z_3, Z_1+Z_4, Z_3+Z_5, Z_6\}$ where $Z_1, \ldots, Z_6$ are independent random variables distributed uniformly in $[0,1]$. We first observe that the density of $Z_1+Z_2+Z_3$ is the case $n=3$ of the Irwin-Hall distribution (also known as the Uniform Sum distribution, see [11]) and is:

$$f_{Z_1+Z_2+Z_3}(t) = \begin{cases} \frac{t^2}{2} & \text{if } 0 \le t \le 1 \\ \frac{1}{2}\left(-2t^2+6t-3\right) & \text{if } 1 \le t \le 2 \\ \frac{1}{2}\left(t^2-6t+9\right) & \text{if } 2 \le t \le 3 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

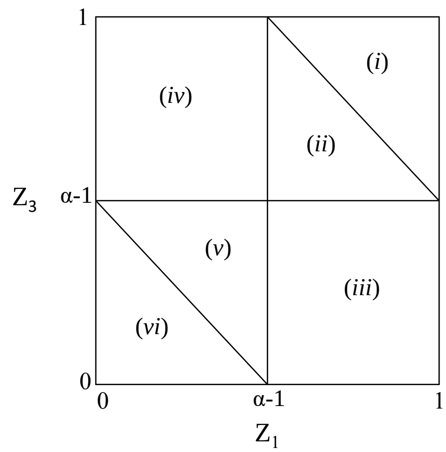

Figure 1: Partitioning $Z_1 + Z_3$ into regions in the case $1 \leq \alpha \leq 2$. The aspect ratio in the figure is for $\alpha = 1.5$.

This covers the case $2 \leq \alpha \leq 3$ since in this case we have $\Pr[X_4 \leq \alpha] = \Pr[Z_1 + Z_2 + Z_3 \leq \alpha]$. We next consider the case $1 \leq \alpha \leq 2$. It will be convenient to look at conditional probabilities given Z_1 and Z_3. We partition into six regions, as depicted in Figure 1.
(i) If $Z_1 + Z_3 \geq \alpha$, then $\Pr[X_4 \leq \alpha] = 0$.
(ii) If $Z_1 + Z_3 \leq \alpha$ and $Z_1 \geq \alpha - 1$ and $Z_3 \geq \alpha - 1$, then $\Pr[X_4 \leq \alpha] = (\alpha - Z_1)(\alpha - Z_3)(\alpha - Z_1 - Z_3)$.
(iii) If $Z_1 + Z_3 \leq \alpha$ and $Z_1 \geq \alpha - 1$ and $Z_3 \leq \alpha - 1$, then $\Pr[X_4 \leq \alpha] = (\alpha - Z_1)(\alpha - Z_1 - Z_3)$.
(iv) If $Z_1 + Z_3 \leq \alpha$ and $Z_1 \leq \alpha - 1$ and $Z_3 \geq \alpha - 1$, then $\Pr[X_4 \leq \alpha] = (\alpha - Z_3)(\alpha - Z_1 - Z_3)$.
(v) If $Z_1 + Z_3 \geq \alpha - 1$ and $Z_1 \leq \alpha - 1$ and $Z_3 \leq \alpha - 1$, then $\Pr[X_4 \leq \alpha] = \alpha - Z_1 - Z_3$.
(vi) If $Z_1 + Z_3 \leq \alpha - 1$, then $\Pr[X_4 \leq \alpha] = 1$.
It follows that for $1 \leq \alpha \leq 2$,

$$\begin{aligned}
\Pr[X_4 \leq \alpha] &= \int_{\alpha-1}^{1} \int_{\alpha-1}^{\alpha-t} (\alpha - t)(\alpha - s)(\alpha - t - s)\, ds\, dt \\
&+ 2 \int_{0}^{\alpha-1} \int_{\alpha-1}^{1} (\alpha - s)(\alpha - t - s)\, ds\, dt \\
&+ \int_{0}^{\alpha-1} \int_{\alpha-1-t}^{\alpha-1} (\alpha - t - s)\, ds\, dt \\
&+ \int_{0}^{\alpha-1} \int_{0}^{\alpha-1-t} 1\, ds\, dt\ .
\end{aligned}$$

Computing these integrals we obtain

$$\begin{aligned}
\Pr[X_4 \le \alpha] &= \frac{(2-\alpha)^3}{120}(\alpha^2+6\alpha+4) \\
&+ \frac{1}{6}(-\alpha^4+4\alpha^3-9\alpha^2+14\alpha-8) \\
&+ \frac{1}{6}(4-\alpha)(\alpha-1)^2 \\
&+ \frac{(\alpha-1)^2}{2} \\
&= -\frac{1}{120}(\alpha^5+20\alpha^4-80\alpha^3+40\alpha^2+20\alpha-12)\ . \qquad (3)
\end{aligned}$$

We last consider the case $0 \le \alpha \le 1$. Here, if $Z_1 + Z_3 \ge \alpha$, then surely $\Pr[X_4 \le \alpha] = 0$. Otherwise,

$$\begin{aligned}
\Pr[X_4 \le \alpha] &= \Pr[Z_6 \le \alpha] \cdot \int_0^{\alpha}\int_0^{\alpha-t}(\alpha-t)(\alpha-s)(\alpha-t-s)\,ds\,dt \\
&= \frac{11\alpha^6}{120}\ . \qquad (4)
\end{aligned}$$

Summarizing (2), (3), (4) we get that the probability density function of X_4 is:

$$f_{X_4}(t) = \begin{cases} \frac{11t^5}{20} & \text{if } 0 < t < 1 \\ -\frac{1}{24}(t^4+16t^3-48t^2+16t+4) & \text{if } 1 < t < 2 \\ \frac{1}{2}\left(t^2-6t+9\right) & \text{if } 2 < t < 3 \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$\ell(T_4, U[0,1]) = \int_0^1 \frac{11t^6}{20}\,dt + \int_1^2 \frac{-t}{24}(t^4+16t^3-48t^2+16t+4)\,dt + \int_2^3 \frac{t}{2}\left(t^2-6t+9\right)\,dt = \frac{7949}{5040}\ .$$

■

Determining $\ell(T_5, U[0,1])$ precisely seems far more complicated.

Problem 1 *Determine* $\ell(T_5, U[0,1])$.

We can use Lemma 3.1 and the exact values of $\ell(T_8, U\{0,1\})$ and $\ell(T_4, U[0,1])$ that were determined here and obtain:

Corollary 3.4

$$\begin{aligned}
\beta_{tr}(U\{0,1\}) &\ge \frac{159925961}{268435456} > 0.595\ , \\
\beta_{tr}(U[0,1]) &\ge \frac{7949}{15120} > 0.525\ .
\end{aligned}$$

■

4 Upper bounds for transitive tournaments

4.1 An upper bound via a large deviation inequality

Throughout this subsection we shall assume that $\mathcal{D}$ is any distribution taking values only in $[0,1]$ and with mean $1/2$ (so in particular this is satisfied by $U[0,1]$ and $U\{0,1\}$). For $m = 1,\dots,n-1$, let $P_{m,n}$ be the set of paths of T_n starting at vertex 1 and ending at vertex n, and which contain m edges. As a path in $P_{m,n}$ corresponds to choosing $m-1$ vertices out of the vertices $2,\dots,n-1$, we have that $|P_{m,n}| = \binom{n-2}{m-1}$. We shall assume here that m goes to infinity and upper bound the probability that a given path in $P_{m,n}$ has a large weight. We will then use this upper bound to give an upper bound for $\beta_{tr}(\mathcal{D})$.

Let $X(T_n, m, \mathcal{D})$ be the random variable corresponding to the weight of a maximum weighted path in $P_{m,n}$. Clearly,

$$X(T_n, \mathcal{D}) = \max_{m=1}^{n-1} X(T_n, m, \mathcal{D}) \ .$$

In order to prove that $\ell(T_n, \mathcal{D}) \le 0.85(n-1)(1+o(1))$, which in turn implies that $\beta_{tr}(\mathcal{D}) \le 0.85$, it suffices to prove the following lemma.

Lemma 4.1 *For all sufficiently large n we have that for all $m = 1,\dots,n-1$.*

$$\Pr[X(T_n, m, \mathcal{D}) > 0.85(n-1)] < \frac{1}{n^2} \ .$$

In particular, for all sufficiently large n, $\Pr[X(T_n, \mathcal{D}) > 0.85(n-1)] < \frac{1}{n}$ and hence $\ell(T_n, \mathcal{D}) \le 0.85(n-1)(1+o(1))$.

Proof. We may assume that $m \ge 0.85(n-1)$ as otherwise $\Pr[X(T_n, m, \mathcal{D}) > 0.85(n-1)] = 0$. As $|P_{m,n}| = \binom{n-2}{m-1}$, it suffices to prove that the probability that a given element of $P_{m,n}$ has weight larger than $0.85(n-1)$ is less than $n^{-2}\binom{n-2}{m-1}^{-1}$.

Fix a path p of length m and let $w(p)$ denote the random variable corresponding to its weight. Observe that $w(p)$ is the sum of m i.i.d. random variables each with distribution $\mathcal{D}$ and recall that $\mathcal{D}$ has mean 0.5 and takes values only in $[0,1]$. Hence, by a Chernoff large deviation inequality (see [2] A.1.18),

$$\Pr[w(p) - \frac{m}{2} > \alpha m] < \exp(-2\alpha^2 m) \ .$$

Let $m = c(n-1)$ where $c \ge 0.85$. Hence $(\alpha + 0.5)m = (\alpha+0.5)c(n-1)$ so by the last inequality,

$$\Pr[w(p) > 0.85(n-1)] < \exp(-2c(\frac{0.85}{c} - \frac{1}{2})^2(n-1)) \ .$$

It thus remains to prove that

$$\exp(-2c(\frac{0.85}{c} - \frac{1}{2})^2(n-1)) < n^{-2}\binom{n-2}{c(n-1)-1}^{-1} \ .$$

Taking logarithms in both sides, this amounts to proving that for every $0.85 \le c \le 1$,

$$2c(\frac{0.85}{c} - \frac{1}{2})^2 > -c\ln c - (1-c)\ln(1-c)$$

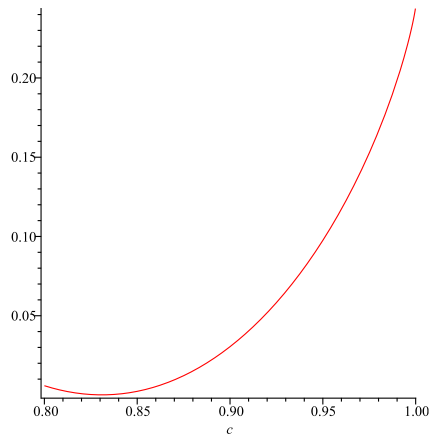

Figure 2: The function $f(c) = 2c(\frac{0.85}{c} - \frac{1}{2})^2 + c\ln c + (1-c)\ln(1-c)$

which indeed holds as the function $f(c) = 2c(\frac{0.85}{c} - \frac{1}{2})^2 + c\ln c + (1-c)\ln(1-c)$ is positive in this range (see Figure 2). Formally, $f(0.85)$ is positive and its derivative is $f'(c) = 0.5 + \ln(c/(1-c)) - 1.445/c^2$ which proves that it is also monotone increasing for $0.85 \le c \le 1$. ■

Corollary 4.2 *Let $\mathcal{D}$ be any distribution with mean $1/2$ taking values only in $[0,1]$. Then $\beta_{tr}(\mathcal{D}) \le 0.85$.* ■

4.2 An upper bound for $\beta_{tr}(U[0,1])$

For the case of the distribution $U[0,1]$ we can do better than the general bound supplied by Corollary 4.2. An edge $e = (i,j) \in E(T_n)$ is said to have *width* $j-i$. Notice that there are precisely $n-k$ edges with width k. Consider the probability space $\mathcal{P}(T_n, U[0,1])$ and let $w(e)$ denote the weight of an edge. We say that an edge $e = (i,j)$ is a *candidate* if $w(e)$ is the weight of a maximum weight path of the sub-tournament induced by $i, i+1, \ldots, j$. In particular, all edges of width 1 are candidates. Let $Y(T_n, U[0,1])$ be the random variable which is the sum of the weights of all candidate edges and recall that $X(T_n, U[0,1])$ is the random variable which is the weight of a maximum weighted path. We claim that $Y(T_n, U[0,1]) \ge X(T_n, U[0,1])$. Suppose not, and let p be a maximum weight path. Then if p contains a non-candidate edge $e = (i,j)$, then replacing e with a maximum weight path in the sub-tournament $i, i+1, \ldots, j$ results in a path with larger weight, a contradiction. So, all edges in p are candidates. Thus, the sum of weights of all candidates is at least as large as the weight of p.

Now, since $Y(T_n, U[0,1]) \ge X(T_n, U[0,1])$ then also $\mathbb{E}[Y(T_n, U[0,1])] \ge \mathbb{E}[X(T_n, U[0,1])] = \ell(T_n, U[0,1])$. So, it suffices to upper bound $\mathbb{E}[Y(T_n, U[0,1])]$.

Lemma 4.3 $\mathbb{E}[Y(T_n, U[0,1])] < \frac{2}{3}(n-1)$. *Consequently, $\beta_{tr}(U[0,1]) \le \frac{2}{3}$.*

Proof. We start by estimating the probability that a given edge $e = (i, i+k)$ is a candidate. Clearly, if $k = 1$ then e is of width 1 so it is trivially a candidate, so assume $k > 1$. Consider the following random variable:

$$Z(e) = \begin{cases} w(e) & \text{if } \sum_{j=i}^{i+k-1} w(j,j+1) \le w(e) \\ 0 & \text{otherwise.} \end{cases}$$

Notice that if e is a candidate, then, in particular, $\sum_{j=i}^{i+k-1} w(j,j+1) \le w(e)$ so its contribution to the sum of weights of all candidates is $w(e) = Z(e)$. If e is not a candidate, then its contribution to the sum of weights of all candidates is $0 \le Z(e)$. So, in particular, if Y_1 is the sum of the weights of all edges with width 1, then

$$Y(T_n, U[0,1]) \le Y_1 + \sum_{e \in E(T)\ ,\ width(e) \ge 2} Z(e) \ .$$

Notice that the distribution of $\sum_{j=i}^{i+k-1} w(j,j+1)$ is the sum of k i.i.d. random variables with distribution $U[0,1]$ so its distribution is the Irwin-Hall distribution with parameter k which, in $[0,1]$, has probability density function $x^{k-1}/(k-1)!$. Also notice that $w(e)$ has distribution $U[0,1]$ and is independent of $\sum_{j=i}^{i+k-1} w(j,j+1)$. Thus, the cumulative distribution function of $Z(e)$ in $[0,1]$ is

$$\Pr[Z(e) \le z] = \int_0^z \int_0^t \frac{x^{k-1}}{(k-1)!} dx\, dt = \frac{z^{k+1}}{(k+1)!} \ .$$

Hence

$$\mathbb{E}[Z(e)] = \int_0^1 z \cdot \frac{z^k}{k!} dz = \frac{k+1}{(k+2)!} \ .$$

Thus,

$$\mathbb{E}[Y(T_n, U[0,1])] = \mathbb{E}[Y_1] + \sum_{k=2}^{n-1} (n-k)\frac{k+1}{(k+2)!} < \frac{n-1}{2} + \frac{1}{6}(n-1) = \frac{2}{3}(n-1) \ .$$

In the last inequality we have used the fact that Y_1 is the sum of $n-1$ i.i.d. variables with mean 0.5 each and that $\sum_{k=2}^{\infty} \frac{k+1}{(k+2)!} = \frac{1}{6}$. Finally, we note that for $k \ge 3$, $\mathbb{E}[Z(e)]$ is strictly larger than the expected contribution of a candidate edge $e = (i, i+k)$ with width k to $Y(T_n, U[0,1])$ since in the definition of $Z(e)$ we only considered the Hamilton path from i to $i+k$ while there can be other paths that void the candidacy of e. This shows that, in fact $\beta_{tr}(U[0,1]) < \frac{2}{3}$. ■

4.3 An upper bound via recurrence

The discrete nature of the random variable $X(T_n, U\{0,1\})$ makes it possible to obtain good estimates for $\beta_{tr}(U\{0,1\})$ via recurrence.

Consider the set of $n-1$ edges of the (unique) Hamilton path of T_n. For any assignment of weights from $\{0,1\}$ to the edges of T_n, there always exists a maximum weight path that contains all the edges of the Hamilton path that received weight 1. Indeed, take any maximum weight path p and assume that $(i,i+1)$ received weight 1 and $(i,i+1) \notin p$. Then let $j \le i$ be the largest index such that $j \in p$, and let (j,j') be the edge of p emanating from j. Notice that $j' \ge i+1$. Then we

k	3	4	5	6	7	8
$c(k)$	$\frac{1}{2}$	$\frac{7}{8}$	$\frac{79}{64}$	$\frac{1663}{1024}$	$\frac{65535}{32768}$	$\frac{4986879}{2097152}$

Table 2: Small values of $c(k)$.

can replace (j, j') with the sub-path $j, j+1, \ldots, j'$ and the new path now contains $(i, i+1)$ and its weight has not decreased, so it is still a maximum weight path. Using this observation, we can now state and prove our recurrence. We require some additional notation. Denote by $c(k)$ the expected weight of a random $U\{0,1\}$ assignment to the edges of T_k under the additional constraint that all the edges of the Hamilton path of T_k receive weight 0. Let $w(e)$ denote the weight of an edge. Thus $c(1) = c(2) = 0$. Notice that $c(3) = 1/2$ since the maximum path weight of T_3 is either 1 if $w(1,3) = 1$ or 0 is $w(1,3) = 0$. Observe that such a weighing of T_4 has maximum path weight at most 1 and it is 0 if all three edges $(1,4), (1,3), (2,4)$ have weight 0. Thus, $c(4) = 7/8$. Some further easily computed values are given in Table 2. To simplify notation denote $\ell(n) = \ell(T_n, U\{0,1\})$.

Lemma 4.4 *For any $n > k \geq 1$ we have*

$$\ell(n) \leq \frac{\ell(n)}{2^k} + \sum_{t=1}^{k} \frac{1}{2^t}(1 + \ell(n-t) + c(t)) \ .$$

Proof. We will bound $\ell(n)$ (which, recall, is the expected value of $X(T_n, U\{0,1\})$) using conditional expectations. Divide the probability space of the $U\{0,1\}$ assignment of T_n into the following $k+1$ events $A_0, A_1, \ldots, A_k$. For $1 \leq t \leq k$, event A_t is the event that edge $(t, t+1)$ received weight 1, and all $t-1$ edges preceding it on the Hamilton path of T_n received weight 0. Event A_0 is the event that all the first k edges on the Hamilton path received weight 0. Clearly, this partitions the probability space, and $\Pr[A_t] = 2^{-t}$ for $t = 1, \ldots, k$, while $\Pr[A_0] = 2^{-k}$. Using conditional expectations we have

$$\ell(n) = \sum_{t=0}^{k} \mathbb{E}[X(T_n, U\{0,1\}) \mid A_t] \cdot \Pr[A_t] \ .$$

As for the first term corresponding to $t = 0$, we use the trivial upper bound $\mathbb{E}[X(T_n, U\{0,1\}) \mid A_0] \leq \mathbb{E}[X(T_n, U\{0,1\})] = \ell(n)$ Since the occurrence of A_0 only decreases the expected largest weight. This corresponds to the first term of the recurrence in the statement of the lemma. As for a general term corresponding to some $1 \leq t \leq k$ we proceed as follows. Suppose event A_t occurred. By the observation preceding the lemma, we know that a maximum weight path can be obtained by taking a maximum weight path on the vertices induced by $1, \ldots, t$, then concatenating it with the edge $(t, t+1)$ that received weight 1 and then concatenating it with a maximum weight path on the vertices induced by $t+1, \ldots, n$. The latter of the three parts is just a weighing of T_{n-t} and hence contributes in expectation a weight of $\ell(n-t)$. The first of the three parts is a weighing of T_t but under the additional constraint that the Hamilton path of this T_t has all its edges with weight 0, as we assume the event A_t now holds. Thus, it contributes in expectation a weight of $c(t)$. The edge $(t, t+1)$ contributes its unit weight. Hence,

$$\mathbb{E}[X(T_n, U\{0,1\}) \mid A_t] = c(t) + 1 + \ell(n-t) \ .$$

This corresponds to a general term in the sum of the recurrence in the statement of the lemma. ■

Corollary 4.5 *$\ell(n) \le 0.614(n-1)$ for all $n \ge 1$.*

Proof. Suppose $\ell(n) \le \alpha(n-1)$ for all $n \ge 1$. We prove that $\alpha \le 0.614$. Notice that this clearly holds for all $n \le 8$ by Table 1. Suppose we know that $\ell(n) \le \alpha(n-1)$ for all $n \le k$. By Lemma 4.4 we know that

$$\ell(n)\left(1-\frac{1}{2^k}\right) \le \sum_{t=1}^{k} \frac{1}{2^t}(1+\ell(n-t)+c(t)) \ .$$

So, to prove that $\ell(n) \le \alpha(n-1)$ for all $n > k$ using induction, we must prove that

$$\sum_{t=1}^{k} \frac{1}{2^t}(1+\alpha(n-t-1)+c(t)) \le \left(1-\frac{1}{2^k}\right)\alpha(n-1) \ .$$

Rearranging the terms, the last inequality is equivalent to showing that

$$-\alpha\left(1-\frac{1}{2^k}\right)+\alpha\sum_{t=1}^{k}\frac{t+1}{2^t} \ge \sum_{t=1}^{k}\frac{1+c(t)}{2^t}$$

which simplifies to

$$\alpha\frac{2^{k+1}-k-2}{2^k} \ge \sum_{t=1}^{k}\frac{1+c(t)}{2^t} \ .$$

Since we know $c(k)$ for all $k \le 8$ using Table 2 we can use $k=8$ in the last inequality and obtain that α must satisfy

$$\alpha\frac{502}{256} \ge \frac{645396351}{536870912} \ .$$

In particular $\alpha = 0.614$ satisfies the last inequality. ■

By 2.3, 3.4, 4.3, 4.5, Theorem 1 is now established. ■

5 Proof of Proposition 1.1

In this section we assume that $\mathcal{D}$ is a nonnegative distribution with $\Pr[\mathcal{D} \ge 1-o(1)] \ge n^{-1/3}$. Notice that $U\{0,1\}$ trivially satisfies this assumption and so does $U[0,1]$ since for $X \sim U[0,1]$ we have $\Pr[X \ge 1-n^{-1/3}] = n^{-1/3}$. We prove that $\ell_{max}(n,\mathcal{D}) \ge n-o(n)$ implying Proposition 1.1.

Lemma 5.1 *$\ell_{max}(n,\mathcal{D}) \ge n-o(n)$.*

Proof. Consider a random tournament G with n vertices and the probability space $\mathcal{P}(G,\mathcal{D})$. By our assumption on $\mathcal{D}$, there exists $t_n \ge 1-o(1)$ such that $\Pr[\mathcal{D} \ge t_n] \ge n^{-1/3}$. Let G' be the subgraph of G obtained by keeping only the edges that received weight at least t_n. We will prove that with high probability (i.e. with probability tending to 1 as n goes to infinity), G' has a Hamilton cycle and in particular, w.h.p., $X(G,\mathcal{D}) \ge (n-1)t_n \ge n-o(n)$. Thus, its expectation $\ell(G,\mathcal{D})$ also satisfies $\ell(G,\mathcal{D}) \ge n-o(n)$. In particular, $\ell_{max}(n,\mathcal{D}) \ge n-o(n)$.

Stated otherwise, we must show that if G' is a random oriented graph where each edge appears with probability $n^{-1/3}$, and the direction of an edge is chosen uniformly from both possible directions, then w.h.p. G' has a Hamilton cycle. One can observe that in the undirected case this occurs already if the edge probability is $\Omega(\log n/n)$ [14] (and in fact, it is possible to modify Pósa's proof using a slightly more complicated argument to handle the case of random orientations – however, we settle here for $n^{-1/3}$ as it makes the proof much simpler, and suffices for our purposes).

We will first prove that w.h.p. for every ordered sequence of vertices $v_1, \ldots, v_k$ with $3 \le k < n$, there exists a vertex u outside of the sequence such that $(v_i, u) \in E(G')$ and also $(u, v_{i+1}) \in E(G')$ for some $1 \le i \le k-1$. For a given sequence $S = v_1, \ldots, v_k$, a given $1 \le i \le k-1$ and a given $u \notin S$, let $A(S,i,u)$ be the event that $(v_i, u) \notin E(G')$ or $(u, v_{i+1}) \notin E(G')$. Hence,

$$\Pr[A(S,i,u)] = 1 - \left(\frac{n^{-1/3}}{2}\right)^2 = 1 - \frac{1}{4n^{2/3}}\ .$$

Let $A(S,u)$ be the event $\cap_{i=1}^{k-1} A(S,i,u)$. Since $A(S,i,u)$ is independent of $A(S,j,u)$ unless $|j-i| \le 1$, we have that

$$\Pr[A(S,u)] \le \left(1 - \frac{1}{4n^{2/3}}\right)^{(k-1)/2}\ .$$

Let $A(S) = \cap_{u \notin S} A(S,u)$. As for $u \ne u'$ the events $A(S,u)$ and $A(S,u')$ are independent we have

$$\Pr[A(S)] \le \left(1 - \frac{1}{4n^{2/3}}\right)^{(k-1)(n-k)/2}\ .$$

There are $k!\binom{n}{k}$ sequences of length k, so taking the union of $A(S)$ over all possible sequences we obtain

$$\Pr[\cup_S A(S)] \le \sum_{k=3}^{n-1} \binom{n}{k} k! \left(1 - \frac{1}{4n^{2/3}}\right)^{k(n-k)/2} \ll \frac{1}{n}\ .$$

We have thus proved that with probability at least $1 - 1/n$, for every ordered sequence of vertices $v_1, \ldots, v_k$ with $3 \le k < n$, there exists a vertex u outside of the sequence such that $(v_i, u) \in E(G')$ and also $(u, v_{i+1}) \in E(G')$ for some $1 \le i \le k-1$. So, assuming this holds our Hamilton cycle can be easily constructed as follows. Take an arbitrary directed cycle (trivially, the probability that G' is acyclic is extremely small). Suppose it is on the vertices $v_1, \ldots, v_k$. If $k = n$ we are done. Otherwise, there exists a vertex u outside of the cycle such that $(v_i, u) \in E(G')$ and also $(u, v_{i+1}) \in E(G')$. Place u between v_i and v_{i+1} to obtain a larger cycle and continue accordingly until a Hamilton cycle is formed. ■

6 Some open problems and conjectures

As mentioned in the introduction, it seems plausible that $\beta = \beta_{tr}$ (and hence that the limsup in the definition of β is a limit).

Conjecture 1

$$\beta(U[0,1]) = \beta_{tr}(U[0,1]) \quad , \quad \beta(U\{0,1\}) = \beta_{tr}(U\{0,1\}) \ .$$

In fact, it is plausible that the conjecture holds for any nonnegative distribution with finite mean.

Notice that the upper bound of $\beta_{tr}(U\{0,1\})$ via recurrence is very close to the lower bound of $\beta_{tr}(U\{0,1\})$. It is better than the more general upper bound given in Corollary 4.2 which applies to any distribution $\mathcal{D}$ with mean 0.5 taking values only in $[0,1]$. On the other hand, $U\{0,1\}$ is easily shown to have the largest possible variance (1/4) among all such distributions $\mathcal{D}$. Hence it is plausible to conjecture that

Conjecture 2 *Let $\mathcal{D}$ be an arbitrary distribution taking values only in* $0,1$ *and with mean* 0.5. *Then,*

$$\beta_{tr}(\mathcal{D}) \le \beta_{tr}(U\{0,1\}) \ .$$

Finally, obtaining exact formulas for $\ell(T_k, U\{0,1\})$ and $\ell(T_k, U[0,1])$ seems like a challenging open problem already for small k.

Acknowledgments

The author especially thanks Safwat Nassar for many useful discussions and his invaluable help in the transitive case.

This research was supported in part by the Israel Science Foundation (Grant No. 1082/16).

References

[1] D. Aldous. The $\zeta(2)$ limit in the random assignment problem. *Random Structures and Algorithms*, 18(4):381–418, 2001.

[2] N. Alon and J.H. Spencer. *The Probabilistic Method, 2nd edition.* Wiley-Interscience, 2nd edition, 2000.

[3] D. Bertsimas, P. Jaillet, and A. Odoni. A priori optimization. *Operations Research*, 38(6):1019–1033, 1990.

[4] B. Bollobás and R. Häggkvist. Powers of hamilton cycles in tournaments. *Journal of Combinatorial Theory, Series B*, 50(2):309–318, 1990.

[5] C. Cooper, A. Frieze, N. Ince, S. Janson, and J. Spencer. On the length of a random minimum spanning tree. *Combinatorics, Probability and Computing*, 25(1):89–107, 2016.

[6] A. Frieze. On the value of a random minimum spanning tree problem. *Discrete Applied Mathematics*, 10(1):47–56, 1985.

[7] A. Frieze. On random symmetric travelling salesman problems. *Mathematics of Operations Research*, 29(4):878–890, 2004.

[8] A. Frieze and C. McDiarmid. Algorithmic theory of random graphs. *Random Structures and Algorithms*, 10(1-2):5–42, 1997.

[9] C. Grinstead and L. Snell. *Introduction to Probability.* American Mathematical Society, 1997.

[10] R. Hassin and E. Zemel. On shortest paths in graphs with random weights. *Mathematics of Operations Research*, 10(4):557–564, 1985.

[11] N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Vol. 2 (Wiley Series in Probability and Statistics)*. Wiley-Interscience, 1995.

[12] R.M. Karp. A patching algorithm for the nonsymmetric traveling-salesman problem. *SIAM Journal on Computing*, 8(4):561–573, 1979.

[13] Y. Peres, D. Sotnikov, B. Sudakov, and U. Zwick. All-pairs shortest paths in $o(n^2)$ time with high probability. *Journal of the ACM*, 60(4):26, 2013.

[14] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

[15] L. Rédei. Ein kombinatorischer satz. *Acta Litt. Szeged*, 7:39–43, 1934.

[16] D. Walkup. On the expected value of a random assignment problem. *SIAM Journal on Computing*, 8(3):440–442, 1979.

[17] J. Wästlund. An easy proof of the $\zeta(2)$ limit in the random assignment problem. *Electronic Communications in Probability*, 14:261–269, 2009.