Cycles of length three and four in tournaments*

Timothy F. N. Chan[†]   Andrzej Grzesik[‡]   Daniel Král'[§]
Jonathan A. Noel[¶]

Abstract

Linial and Morgenstern conjectured that, among all n-vertex tournaments with $d\binom{n}{3}$ cycles of length three, the number of cycles of length four is asymptotically minimized by a random blow-up of a transitive tournament with all but one part of equal size and one smaller part. We prove the conjecture for $d \geq 1/36$ by analyzing the possible spectrum of adjacency matrices of tournaments. We also demonstrate that the family of extremal examples is broader than expected and give its full description for $d \geq 1/16$.

1 Introduction

One of the oldest theorems in extremal graph theory is Mantel's theorem [14], which asserts that every n-vertex graph with more than $n^2/4$ edges contains a triangle. The Erdős–Rademacher Problem, which can be traced back to the work of Rademacher in the 1940's and the later work of Erdős [4], asks for the minimum possible number of triangles in a graph with a given number of vertices and edges.

It was conjectured that this minimum is asymptotically attained by a complete multipartite graph (i.e. a blow-up of a constant-sized clique) with all but one part of equal size and one smaller part. This conjecture attracted substantial attention for several decades, see e.g. [1,5,8,12], until its solution by Razborov [17] in 2008 using his newly developed flag algebra method. Pikhurko and Razborov [16] described the asymptotic structure of all extremal graphs for this problem and an exact description was obtained by Liu, Pikhurko and Staden [10]. The more general problem of determining the minimum asymptotic density of k-cliques in graphs with given edge-density (the Erdős–Rademacher Problem corresponds to the case $k = 3$) has also been solved by Nikiforov [15] (the case $k = 4$) and by Reiher [18] in full generality.

In this paper, we investigate a related problem for tournaments posed by Linial and Morgenstern [9], who asked for the minimum density of 4-cycles in a large tournament with fixed density of 3-cycles. They conjectured that the tournament asymptotically minimizing this density is a blow-up of a transitive tournament with all but one part of equal size and one smaller part in which the arcs within each part are oriented randomly (they call this construction a *random blow-up*), i.e., the structure of the conjectured extremal examples is akin to those of the Erdős–Rademacher problem.

We confirm this conjecture in the case where the proposed extremal examples have two or three parts and provide a full description of extremal tournaments in the two-part case. In contrast to many of the recent proofs in this area that use the flag algebra method, our approach is based on the analysis of the spectrum of adjacency matrices of tournaments.

We now state the problem that we study in the paper more formally. The *density* of the directed cycle C_ℓ of length ℓ in a tournament T, denoted by $t(C_\ell, T)$, is the probability that a random mapping from $V(C_\ell)$ to $V(T)$ is a homomorphism (i.e. arcs of C_ℓ map to arcs of T). Note that, for fixed ℓ, a tournament T on n vertices contains $t(C_\ell, T)n^\ell/\ell + O(n^{\ell-1})$ cycles of length ℓ. In fact, for $\ell \in \{3, 4, 5\}$, the error term is zero as every homomorphism of C_ℓ to T is injective. A standard application of the Cauchy–Schwarz inequality shows that $t(C_3, T) \leq 1/8$ for every tournament T (see [3, Fact 1] for details). Our focus is on bounding the minimum possible value of $t(C_4, T)$ asymptotically as a function of $t(C_3, T)$.

Next, we describe the family of conjectured tight examples from [9] which will motivate the definition of the function g below. Given $z \in [0, 1]$, we let n be an integer chosen large with respect to z and define an n-vertex tournament T as follows. If $z = 0$, then let T be a transitive tournament. Otherwise, the vertices of T are split into $\lfloor z^{-1} \rfloor + 1$ disjoint parts $V_1, \ldots, V_{\lfloor z^{-1} \rfloor + 1}$ such that $\lfloor z^{-1} \rfloor$ parts contain exactly $\lfloor zn \rfloor$ vertices and the remaining part contains the rest of the vertices (note that if z^{-1} and zn are integers, then the last part is empty). If two vertices v and v' respectively belong to distinct parts V_i and V_j with $i < j$, then the tournament T contains an edge from v to v'. If v and v' instead belong to the *same* part, then the edge between them is oriented from v to v' with probability

2
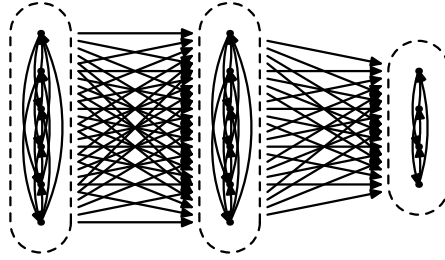
Figure 1: An illustration of the random blow up construction for $z = 3/8$.

$1/2$, i.e., the vertices of each part induce a randomly oriented tournament. See Figure 1 for an illustration. It is easy to see that $t(C_3, T) = t(C_4, T) = 0$ if $z = 0$ and, if $z \in (0, 1]$, then, with high probability, it holds that

$$t(C_3, T) = \frac{1}{8} \left(\lfloor z^{-1} \rfloor z^3 + \left(1 - \lfloor z^{-1} \rfloor z \right)^3 \right) + o(1)$$

and

$$t(C_4, T) = \frac{1}{16} \left(\lfloor z^{-1} \rfloor z^4 + \left(1 - \lfloor z^{-1} \rfloor z \right)^4 \right) + o(1)$$

because of the concentration around the expected values.

The conjecture of Linial and Morgenstern [9] asserts that the above construction is asymptotically optimal. In light of this, we write *the regime of k parts* to denote the set of values of $t(C_3, T)$ between $1/(8k^2)$ and $1/(8(k-1)^2)$, corresponding to the range of values for which the above construction has its vertices split into k parts. In particular, the focus of this paper is on the regimes of two and three parts, which refer to values of $t(C_3, T)$ in the ranges $[1/32, 1/8]$ and $[1/72, 1/32]$, respectively.

To formally state the conjecture, define $g : [0, 1/8] \to [0, 1]$ as follows: $g(0) = 0$ and

$$g \left(\frac{1}{8} \left(\lfloor z^{-1} \rfloor z^3 + \left(1 - \lfloor z^{-1} \rfloor z \right)^3 \right) \right) = \frac{1}{16} \left(\lfloor z^{-1} \rfloor z^4 + \left(1 - \lfloor z^{-1} \rfloor z \right)^4 \right)$$

for $z \in (0, 1]$.

Conjecture 1 (Linial and Morgenstern [9, Conjecture 2.2]). *It holds that*

$$t(C_4, T) \geq g(t(C_3, T)) + o(1)$$

for every tournament T.[1]

Although $0 \leq t(C_3, T) \leq 1/8$ for every tournament T, the conjecture is currently only known to hold for tournaments with 3-cycle density asymptotically equal to 0, 1/8, or 1/32 [9].

---

[1]Linial and Morgenstern phrased their conjecture equivalently in terms of subgraph counts as opposed to homomorphism densities; the quantities c_3 and c_4 in their paper are asymptotically equal to $2t(C_3, T)$ and $6t(C_4, T)$, respectively.
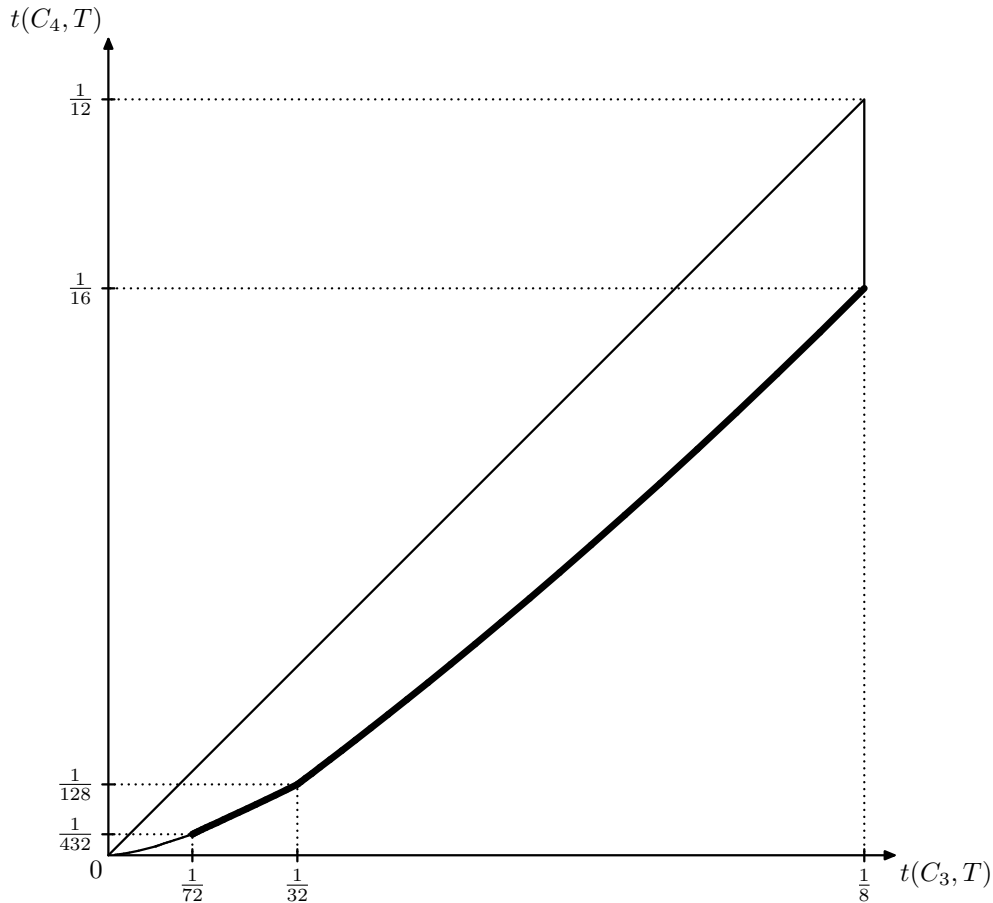
Figure 2: The conjectured region of asymptotically feasible densities of C_3 and C_4 in tournaments. The lower bound for $t(C_3, T) \in \{1/8, 1/32\}$ and the upper bound were proved in [9]. The rest of the lower bound is conjectured except for the part depicted in bold, which we prove in this paper.

We confirm the conjecture for all 3-cycle densities in the range $[1/72, 1/8]$ (Theorems 5 and 11, see the discussion after Conjecture 2 in Section 2) and characterize the asymptotic structure of the extremal tournaments for densities in $[1/32, 1/8]$ (Corollary 6). We refer the reader to Figure 2 for the visualization of the conjectured feasible region of 3-cycle and 4-cycle densities.

Conjecture 1 appears to be resistant to the flag algebra method and we follow a different approach based on spectral analysis of adjacency matrices of tournaments. We believe that the difficulty in applying the flag algebra method is rooted in the fact that random blow-ups of transitive tournaments are far from being the unique extremal examples for Conjecture 1. In particular, a rather complicated family of extremal examples T is described as follows. Denote the vertices of T by $v_1, \ldots, v_n$ and associate v_i with a real number $p_i \in [0, 1/2]$, $i = 1, \ldots, n$.

4

Then, direct the edge $v_i v_j$ from v_i to v_j with probability $1/2 + p_i - p_j$. Note that, if all the values of p_i are either 0 or $1/2$, then this construction is nothing more than a random blow-up of a 2-vertex tournament, i.e., it is identical to the examples of [9] for 3-cycle density in $[1/32, 1/8]$. For large n, this tournament satisfies $t(C_4, T) = g(t(C_3, T)) + o(1)$ with high probability (this follows from Theorem 5). In particular, all tournaments obtained in this way are extremal with respect to Conjecture 1 in the regime of two parts. In Corollary 6, we prove that these are asymptotically the only extremal constructions in this regime. As we have said, we believe that this complex structure of extremal examples makes it challenging to apply the flag algebra method. Indeed, the method relies on finding suitable positive semidefinite matrices such that the vectors of rooted homomorphism densities belong to their kernels for all choices of roots in any extremal example. However, the richness of the structure of extremal examples restricts significantly which matrices could possibly appear in the flag algebra argument.

We conclude this introductory section by summarizing the previous results on the interplay between the densities of C_3 and C_4 in tournaments. Linial and Morgenstern [9] proved $t(C_4, T) \geq \frac{12t(C_3,T)^2}{1+16t(C_3,T)}$ which confirmed Conjecture 1 in the case $t(C_3, T) = 1/32 + o(1)$. They also proved that, for $d \in [0, 1/8]$, the asymptotically feasible densities of cycles of length four in tournaments with $t(C_3, T) = d + o(1)$ form an interval [9, Proof of Lemma 1.3]. For the related problem of *maximizing* the 4-cycle density relative to $t(C_3, T)$, they proved that $t(C_4, T) \leq 2t(C_3, T)/3$ for all T where equality holds if and only if $|V(T)| \geq 4$ and every set of 4 vertices either induces a transitive tournament or contains a 4-cycle.

For completeness, we briefly describe a tight family of tournaments from [9] for the upper bound $t(C_4, T) \leq 2t(C_3, T)/3$. Let $\xi \in [0, 1/2]$ and denote the vertices by $v_1, \ldots, v_n$ with n large. The edge between v_i and v_j, $i \leq j$, is directed from v_i to v_j if and only if $j \leq i + (1 - \xi)n$. If $\xi = 0$, we obtain a transitive tournament and if $\xi = 1/2$, we obtain the "circular" n-vertex tournament, i.e., the tournament contains edges from v_i to $v_{i+1}, \ldots, v_{i+\lfloor n/2 \rfloor}$ (indices modulo n). These tournaments achieve all possible values of $t(C_3, T)$ in the limit as n tends to infinity and their 4-vertex subtournaments are either transitive or contain a 4-cycle (see [9, Proof of Observation 2.1]); therefore, they show that the upper boundary in Figure 2 is asymptotically feasible.

2 Preliminaries

In this section, we introduce the notation used throughout the paper. The set of all positive integers is denoted by $\mathbb{N}$ and the set of integers $1, \ldots, n$ by $[n]$. Some of the matrices that we consider have complex eigenvalues and the complex unit will be denoted by ι. If A is a matrix (or a vector), then we write A^T for

its transpose and A^* for its conjugate transpose; in particular, if A is real, then $A^T = A^*$. The *trace* of a square matrix A is the sum of the entries in its diagonal and is denoted by $\operatorname{Tr} A$. We let $\langle \cdot \mid \cdot \rangle$ denote the standard inner (dot) product on $\mathbb{R}^n$. We use $\mathbb{J}_n$ to denote the square matrix of order n such that each entry of $\mathbb{J}_n$ is equal to one; if n is clear from the context, we will omit the subscript. Note that $\mathbb{J}_n$ has one eigenvalue equal to n and the remaining $n-1$ eigenvalues are equal to zero. The n-dimensional column vector with all entries equal to one is denoted by $\vec{j}_n$ and we again omit the subscript when n is clear from the context. Note that $\mathbb{J}_n = \vec{j}_n \vec{j}_n^T$.

2.1 Tournament matrices

We say that a square matrix A of order n is a *tournament matrix* if A is non-negative and $A + A^T = \mathbb{J}$; in particular, if A is a tournament matrix, then each diagonal entry of A is equal to $1/2$. Every n-vertex tournament T can be associated with a tournament matrix A of order n, which we refer to as the *adjacency matrix* of T, in the following way. Each diagonal entry A is equal to $1/2$ and, for $i \neq j$, the entry of A in the i-th row and the j-th column (denoted $A_{i,j}$) is equal to 1 if T contains an arc from the i-th vertex to the j-th vertex, and it is equal to 0 otherwise. The following proposition readily follows.

Proposition 1. *Let T be a tournament on n vertices, A be the adjacency matrix of T and $\ell \geq 3$. The number of homomorphisms of C_ℓ to T is $\operatorname{Tr} A^\ell + O(n^{\ell-1})$.*

Recall that the trace of a matrix is equal to the sum of its eigenvalues and that the eigenvalues of the ℓ-th power of a matrix are the ℓ-th powers of its eigenvalues. In view of Proposition 1, for $\ell \geq 1$, we define $\sigma_\ell(A)$ for a square matrix A of order n to be

$$\sigma_\ell(A) = \frac{1}{n^\ell} \sum_{i=1}^{n} \lambda_i^\ell = \frac{1}{n^\ell} \operatorname{Tr} A^\ell$$

where $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ are the eigenvalues of A. Note that the normalization of the sum is chosen in such a way that $\sigma_1(A) = 1/2$ for every tournament matrix A.

Next, we argue that Conjecture 1 is equivalent to the following.

Conjecture 2. *If A is a tournament matrix, then $\sigma_4(A) \geq g(\sigma_3(A))$.*

Indeed, Conjecture 2 implies Conjecture 1 by Proposition 1. In the other direction, suppose that there exists a tournament matrix A of order n such that $\sigma_4(A) < g(\sigma_3(A))$. We consider the following (random) tournament T with $k \cdot n$ vertices, $k \in \mathbb{N}$: the vertices of T are split into n sets $V_1, \ldots, V_n$, each containing k vertices, and a vertex $v \in V_i$ is joined by an arc to a vertex $v' \in V_j$ with probability $A_{i,j}$; note that v' is joined by an arc to v with probability $A_{j,i} = 1 - A_{i,j}$, i.e., the tournament T is well defined. Since n is fixed, for $\ell \in \{3, 4\}$ and large k,

the number of homomorphisms from C_ℓ to T is $\sigma_\ell(A)(nk)^\ell + O(k^{\ell-1})$ with high probability and so Conjecture 1 fails for $t(C_3, T) \approx \sigma_3(A)$.

We conclude this subsection by establishing the following lemma. A similar result appears in Brauer and Gentry [2], but for a slightly different definition of a tournament matrix.

Lemma 2. *If A is a tournament matrix, then every eigenvalue of A has non-negative real part.*

Proof. Let λ be any eigenvalue of A and let v be a corresponding eigenvector. Observe that the following holds:

$$0 \leq \overline{(\vec{j}^T v)}(\vec{j}^T v) = v^* \mathbb{J} v = v^*(A + A^T)v = v^*(Av) + (v^* A^T)v$$
$$= v^*(\lambda + \overline{\lambda})v = (\lambda + \overline{\lambda})v^* v.$$

Since $v^* v$ is a non-negative real number, it follows that $\lambda + \overline{\lambda}$ is a non-negative real. In particular, the real part of λ is non-negative. $\square$

2.2 Tournament limits

One of the substantial recent developments in combinatorics is the theory of graph limits which aims to provide analytic tools to represent and analyze large graphs. In an analogous way, one can develop a limit theory for tournaments, in which essentially all of the foundational results for graph limits, which can be found, e.g., in the monograph on graph limits by Lovász [11], translate to similar statements for tournament limits with essentially the same proofs. Below, we define tournament limits and outline some of the basic results that we will use.

A *tournament limit* is a measurable function $W : [0, 1]^2 \to [0, 1]$ such that $W(x, y) + W(y, x) = 1$ for all $(x, y) \in [0, 1]^2$. One can define the density of the cycle C_ℓ in W as follows:

$$t(C_\ell, W) = \int_{x_1, \ldots, x_\ell \in [0,1]} W(x_1, x_2) W(x_2, x_3) \cdots W(x_{\ell-1}, x_\ell) W(x_\ell, x_1) \, \mathrm{d}x_1 \cdots x_\ell.$$

Note that any n-dimensional tournament matrix A can be represented by a tournament limit W_A by dividing $[0, 1]$ into sets $I_1, \ldots, I_n$ of measure $1/n$ and setting W equal to $A_{i,j}$ on the set $I_i \times I_j$. It is easily observed that $t(C_\ell, W_A)$ is precisely $\sigma_\ell(A)$. The following proposition links densities of cycles in tournament limits and in tournaments.

Proposition 3. *The following two statements are equivalent for every sequence $(s_\ell)_{\ell \geq 3}$ of non-negative reals:*

- *There exists a tournament limit W such that $t(C_\ell, W) = s_\ell$ for every $\ell \geq 3$.*

- *There exists a sequence $(T_i)_{i \in \mathbb{N}}$ of tournaments with increasing orders such that*

$$\lim_{i \to \infty} t(C_\ell, T_i) = s_\ell$$

  for every $\ell \geq 3$.

The first statement easily implies the second by letting T_i be a W-*random tournament* of order i; that is, we let $x_1, \ldots, x_i$ be independent uniformly random points of $[0, 1]$ and join the ith vertex to the jth with probability $W(x_i, x_j)$. For the other direction, the tournament limit W can be constructed by adapting one of the existing arguments in the graph case, e.g., the argument of Lovász and Szegedy [13] based on weak regularity and the Martingale Convergence Theorem. In light of Proposition 3, Conjecture 1 is equivalent to the following.

Conjecture 3. *For every tournament limit W, it holds that $t(C_4, W) \geq g(t(C_3, W))$.*

The notion of regularity decompositions of graphs readily extends to tournaments. We present here the notion of weak regular partitions introduced by Frieze and Kannan in [6] adapted to the setting of tournament limits. We use $|X|$ to denote the measure of a measurable subset X of $[0, 1]$. Given a tournament limit W and $\varepsilon \in (0, 1)$, a partition $Z_1, \ldots, Z_n$ of $[0, 1]$ into sets of measure $1/n$ is *weak ε-regular* for W if

$$\left| \int_{(x,y) \in X \times Y} W(x, y) \, \mathrm{d}x \, \mathrm{d}y - \sum_{i,j=1}^{n} A_{i,j} \cdot |Z_i \cap X| \cdot |Z_j \cap Y| \right| \leq \varepsilon$$

for all measurable subsets X and Y of $[0, 1]$, where A is the tournament matrix defined by

$$A_{i,j} = \frac{\int_{(x,y) \in Z_i \times Z_j} W(x, y) \, \mathrm{d}x \, \mathrm{d}y}{|Z_i| \cdot |Z_j|} \, .$$

We say that a tournament limit W' is a *weak ε-regular approximation* of W if there exists a weak ε-regular partition $\{Z_1, \ldots, Z_n\}$ such that $W'(x, y) = A_{i,j}$ for $(x, y) \in Z_i \times Z_j$, $i, j \in [n]$, where A is the tournament matrix associated with the partition.

The results of Frieze and Kannan [6] adapted to the setting of tournament limits and the corresponding arguments for graph limits [13] yield the following: for every tournament limit W and $k \geq 2$, there exists a weak $1/k$-regular partition $\{Z_{k,1}, \ldots, Z_{k,n_k}\}$ with the following properties: (a) n_k is bounded by a function of k, and (b) the partitions are *refining* in the sense that, for every $k < k'$ and $i' \in [n_{k'}]$, the set $Z_{k',i'}$ is contained in $Z_{k,i}$ for some $i \in [n_k]$. It can be shown analogously to the graph case that the corresponding weak $1/k$-regular approximations converge to W in L_1. In particular, it holds that

$$\lim_{k \to \infty} \sigma_\ell(A_k) = t(C_\ell, W)$$

for every $\ell \geq 3$, where, for $k \in \mathbb{N}$, A_k is the tournament matrix associated with the partition $Z_{k,1}, \ldots, Z_{k,n_k}$.

We conclude with a proposition on the density of C_3 in a weak regular approximation of a tournament limit. The proof of the proposition is also valid in a more general setting of step approximations of tournament limits, which need not be weak regular, however, we prefer stating the proposition in the restricted setting of weak regular approximations to avoid introducing additional notation not needed for our exposition.

Proposition 4. *Let W be a tournament limit and W' a weak ε-regular approximation of W. It holds that $t(C_3, W) \leq t(C_3, W')$.*

Proof. We begin by showing that any tournament limit U satisfies

$$t(C_3, U) = \frac{1}{2} - \frac{3}{2} \int_{x \in [0,1]} \left(\int_{y \in [0,1]} U(x,y) \,\mathrm{d}y \right)^2 \,\mathrm{d}x\,. \tag{1}$$

To do this, we derive two identities based on the symmetry of variables x, y and z. Firstly,

$$1 = \int_{x,y,z \in [0,1]} (U(x,y) + U(y,x))(U(x,z) + U(z,x))(U(y,z) + U(z,y)) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

$$= 2 \int_{x,y,z \in [0,1]} U(x,y)U(y,z)U(z,x) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

$$+ 6 \int_{x,y,z \in [0,1]} U(x,y)U(y,z)U(x,z) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

$$= 2t(C_3, U) + 6 \int_{x,y,z \in [0,1]} U(x,y)U(y,z)U(x,z) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,, \tag{2}$$

and similarly,

$$\int_{x \in [0,1]} \left(\int_{y \in [0,1]} U(x,y) \,\mathrm{d}y \right)^2 \,\mathrm{d}x = \int_{x,y,z \in [0,1]} U(x,y)U(x,z) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

$$= \int_{x,y,z \in [0,1]} U(x,y)U(x,z)(U(y,z) + U(z,y)) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

$$= 2 \int_{x,y,z \in [0,1]} U(x,y)U(x,z)U(y,z) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,. \tag{3}$$

Noticing that the integrals on the last lines in (2) and (3) are the same, the equality (1) is obtained. Hence, the inequality from the statement of the proposition is equivalent to

$$\int_{x \in [0,1]} f'(x)^2 \,\mathrm{d}x \leq \int_{x \in [0,1]} f(x)^2 \,\mathrm{d}x\,, \tag{4}$$

9

where for brevity we have set

$$f'(x) = \int_{y \in [0,1]} W'(x, y) \, \mathrm{d}y \qquad \text{and} \qquad f(x) = \int_{y \in [0,1]} W(x, y) \, \mathrm{d}y \,.$$

Since

$$f'(x) = \frac{1}{|Z_i|} \int_{x' \in Z_i} f(x') \, \mathrm{d}x'$$

for every x in a part Z_i of the weak ε-regular partition defining the tournament limit W', it holds that

$$\int_{x \in Z_i} f'(x)^2 \, \mathrm{d}x = |Z_i| \left(\frac{1}{|Z_i|} \int_{x' \in Z_i} f(x') \, \mathrm{d}x' \right)^2$$

$$\leq \int_{x \in Z_i} f(x)^2 \, \mathrm{d}x \,, \tag{5}$$

where the last line follows from the Cauchy–Schwarz inequality. Summing the inequalities obtained from applying (5) to each Z_i yields (4). $\qquad \square$

3 Regime of two parts

Our goal in this section is to prove Conjecture 2 in the case that $\sigma_3(A) \geq 1/32$, as well as describe the tournament matrices which achieve equality. We then apply this result to characterise the extremal tournament limits for Conjecture 3 for $t(C_3, W) \geq 1/32$. Throughout the proof of the next theorem, we will frequently use the property that the trace of a product of matrices is invariant under "cyclic permutations", i.e., $\mathrm{Tr}\,(M_1 M_2 \cdots M_k) = \mathrm{Tr}\,(M_2 \cdots M_k M_1)$.

Theorem 5. *Let A be a tournament matrix of order n. If $\sigma_3(A) \geq 1/32$, then $\sigma_4(A) \geq g(\sigma_3(A))$ and equality holds if and only if there exists a vector $z \in \mathbb{R}^n$ such that $A_{i,j} = 1/2 + z_i - z_j$ for $i, j \in [n]$.*

Proof. Fix a tournament matrix A of order n. Let $B = \mathbb{J} - 2A$. Note that B is a skew-symmetric matrix, i.e., $B = -B^T$. It follows (see, e.g., [7, p. 293]) that B can be written as $B = ULU^T$ where the columns $v_1, v_2, \ldots, v_n$ of U form an orthonormal basis of $\mathbb{R}^n$ and L has the form

$$L = \begin{bmatrix} 0 & \lambda_1 n & 0 & 0 & \cdots & 0 & 0 \\ -\lambda_1 n & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 n & \cdots & 0 & 0 \\ 0 & 0 & -\lambda_2 n & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_k n \\ 0 & 0 & 0 & 0 & \cdots & -\lambda_k n & 0 \end{bmatrix}$$

10

if n is even, and

$$
L =
\begin{bmatrix}
0 & \lambda_1 n & 0 & 0 & \cdots & 0 & 0 & 0 \\
-\lambda_1 n & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & 0 & \lambda_2 n & \cdots & 0 & 0 & 0 \\
0 & 0 & -\lambda_2 n & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & \lambda_k n & 0 \\
0 & 0 & 0 & 0 & \cdots & -\lambda_k n & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0
\end{bmatrix}
$$

otherwise, where $k = \lfloor n/2 \rfloor$ and $\lambda_1, \ldots, \lambda_k$ are real numbers. (Note that they are not the eigenvalues of B.) Since replacing v_{2i-1} and v_{2i} with $v_{2i-1} \cos \beta + v_{2i} \sin \beta$ and $v_{2i} \cos \beta - v_{2i-1} \sin \beta$, respectively, does not change the matrix B (this corresponds to rotating the basis inside the plane spanned by v_{2i-1} and v_{2i}), we can assume that the vectors $v_2, v_4, \ldots, v_{2k}$ are orthogonal to the vector $\vec{j}$. Set $\alpha_i = \cos^{-1} \left\langle v_{2i-1} \mid n^{-1/2} \vec{j} \right\rangle$ for $i \in [k]$, and additionally set $\alpha_{k+1} = \cos^{-1} \left\langle v_{2k+1} \mid n^{-1/2} \vec{j} \right\rangle$ if n is odd.

We next examine $\operatorname{Tr} A^3$ and $\operatorname{Tr} A^4$ in terms of $\mathbb{J}$ and B. We start with the trace of A^3:

$$
8 \operatorname{Tr} A^3 = \operatorname{Tr}(\mathbb{J} - B)^3 = \operatorname{Tr} \mathbb{J}^3 - 3 \operatorname{Tr} \mathbb{J}^2 B + 3 \operatorname{Tr} \mathbb{J} B^2 - \operatorname{Tr} B^3 .
$$

Since both B and B^3 are skew-symmetric, it follows that $\operatorname{Tr} \mathbb{J}^2 B = 0$ and $\operatorname{Tr} B^3 = 0$. We next analyze the term $\operatorname{Tr} \mathbb{J} B^2$. Since $v_1, \ldots, v_n$ are mutually orthogonal and $v_2, v_4, \ldots, v_{2k}$ are orthogonal to $\vec{j}$, we have

$$
\operatorname{Tr} \mathbb{J} B^2 = \frac{1}{n} \operatorname{Tr} \mathbb{J}^2 B^2 = \frac{1}{n} \operatorname{Tr} \mathbb{J} B^2 \mathbb{J} = \frac{1}{n} \operatorname{Tr} \mathbb{J} (U L U^T)^2 \mathbb{J} = \frac{1}{n} \operatorname{Tr} \mathbb{J} U L^2 U^T \mathbb{J}
$$

$$
= -n^2 \sum_{i=1}^{k} \lambda_i^2 \left\langle v_{2i-1} \mid \vec{j} \right\rangle^2 = -n^3 \sum_{i=1}^{k} \lambda_i^2 \cos^2 \alpha_i .
$$

Hence, we obtain that

$$
8 \sigma_3(A) = 1 - 3 \sum_{i=1}^{k} \lambda_i^2 \cos^2 \alpha_i. \tag{6}
$$

Similarly, we can express the trace of A^4 as follows:

$$
16 \operatorname{Tr} A^4 = \operatorname{Tr} \mathbb{J}^4 - 4 \operatorname{Tr} \mathbb{J}^3 B + 4 \operatorname{Tr} \mathbb{J}^2 B^2 + 2 \operatorname{Tr} \mathbb{J} B \mathbb{J} B - 4 \operatorname{Tr} \mathbb{J} B^3 + \operatorname{Tr} B^4 .
$$

Since B and B^3 are skew-symmetric, it follows that $\operatorname{Tr} \mathbb{J}^3 B = 0$, $\operatorname{Tr} \mathbb{J} B \mathbb{J} B = 0$ and $\operatorname{Tr} \mathbb{J} B^3 = 0$. Also, $\operatorname{Tr} B^4 = 2n^4 \sum_{i=1}^{k} \lambda_i^4$ by the cyclic property. Consequently,

11

we get that

$$16\sigma_4(A) = 1 - 4\sum_{i=1}^{k} \lambda_i^2 \cos^2 \alpha_i + 2\sum_{i=1}^{k} \lambda_i^4. \tag{7}$$

Recall that, if $\sigma_3(A) \in [1/32, 1/8]$, then $g(\sigma_3(A)) = \frac{1}{16}(z^4 + (1-z)^4)$ where $z \in [1/2, 1]$ such that $\sigma_3(A) = \frac{1}{8}(z^3 + (1-z)^3)$. So, for $\sigma_3(A)$ in the considered range, we have

$$8\sigma_3(A) = z^3 + (1-z)^3 = 1 - 3(z - z^2).$$

By comparing this equation to (6), it must be the case that

$$z - z^2 = \sum_{i=1}^{k} \lambda_i^2 \cos^2 \alpha_i. \tag{8}$$

It follows that

$$
\begin{aligned}
16g(\sigma_3(A)) &= z^4 + (1-z)^4 = 1 - 4z + 6z^2 - 4z^3 + 2z^4 \\
&= 1 - 4(z - z^2) + 2(z - z^2)^2 \\
&= 1 - 4\sum_{i=1}^{k} \lambda_i^2 \cos^2 \alpha_i + 2\left(\sum_{i=1}^{k} \lambda_i^2 \cos^2 \alpha_i\right)^2.
\end{aligned}
$$

Combining this with (7), we see that the inequality $\sigma_4(A) \geq g(\sigma_3(A))$ holds if and only if

$$\sum_{i=1}^{k} \lambda_i^4 \geq \left(\sum_{i=1}^{k} \lambda_i^2 \cos^2 \alpha_i\right)^2, \tag{9}$$

and $\sigma_4(A) = g(\sigma_3(A))$ if and only if (9) holds with equality.

Since $v_1, \ldots, v_n$ form an orthonormal basis of $\mathbb{R}^n$,

$$\sum_{i=1}^{n} \left\langle v_i \mid n^{-1/2}\vec{j} \right\rangle^2 = \left\langle n^{-1/2}\vec{j} \mid n^{-1/2}\vec{j} \right\rangle^2 = 1.$$

Thus, $\sum_{i=1}^{k} \cos^2 \alpha_i = 1$ if n is even and $\sum_{i=1}^{k+1} \cos^2 \alpha_i = 1$ otherwise. In either case, $\sum_{i=1}^{k} \cos^4 \alpha_i \leq 1$ and the equality holds if and only if exactly one of the values of $\alpha_1, \ldots, \alpha_k$ is equal to zero and the remainder are equal to $\pi/2$. Since the Cauchy–Schwarz inequality implies that

$$\left(\sum_{i=1}^{k} \lambda_i^2 \cos^2 \alpha_i\right)^2 \leq \left(\sum_{i=1}^{k} \lambda_i^4\right) \cdot \left(\sum_{i=1}^{k} \cos^4 \alpha_i\right), \tag{10}$$

we obtain that the inequality (9) indeed holds.

Now, assume that the inequality (9) holds with equality. As we have seen, this can only occur if exactly one of the α_i are zero and the rest are $\pi/2$. By

12

symmetry, we can assume that $\alpha_1 = 0$ and $\alpha_i = \pi/2$ for $i > 1$. It follows that $\lambda_2 = \cdots = \lambda_k = 0$, and $v_1 = n^{-1/2}\vec{j}$. Hence, as $B = ULU^T$, the entry $B_{i,j}$ is equal to $\lambda_1 n^{1/2}(v_{2,j} - v_{2,i})$ for all $i, j \in [n]$. It follows that, for $\sigma_3(A) \in [1/32, 1/8]$, if $\sigma_4(A) = g(\sigma_3(A))$, then $A_{i,j} = 1/2 + z_i - z_j$ where $z_i = \lambda_1 n^{1/2} v_{2,i}/2$. Conversely, any matrix of this form satisfies (9) with equality and therefore satisfies $\sigma_4(A) = g(\sigma_3(A))$. $\qquad\square$

Reinterpreting Theorem 5 in the language of tournament limits, we obtain the following corollary.

Corollary 6. *Let W be a tournament limit. If $t(C_3, W) \geq 1/32$, then $t(C_4, W) \geq g(t(C_3, W))$ and the equality holds if and only if there exists a measurable function $f : [0, 1] \to [0, 1/2]$ such that $W(x, y) = 1/2 + f(x) - f(y)$ for almost every $(x, y) \in [0, 1]^2$.*

Proof. Let $(W_k)_{k \in \mathbb{N}}$ be a sequence of refining weak $1/k$-regular approximations of W and let A_k, $k \in \mathbb{N}$, be the corresponding tournament matrices. Since $t(C_3, W) \geq 1/32$, it holds that $\sigma_3(A_k) = t(C_3, W_k) \geq 1/32$ by Proposition 4. Thus, by Theorem 5, we have that

$$t(C_4, W_k) = \sigma_4(A_k) \geq g\left(\sigma_3(A_k)\right) = g(t(C_3, W_k)).$$

and so $t(C_4, W) \geq g(t(C_3, W))$ by the fact that $(W_k)_{k \in \mathbb{N}}$ converges to W in L_1.

To prove the structure of W in the case of equality, assume that $t(C_4, W) = g(t(C_3, W))$. Let n_k be the order of A_k for $k \in \mathbb{N}$, and let $\lambda_{k,1}, \ldots, \lambda_{k,\lfloor n_k/2 \rfloor}$, $\alpha_{k,1}, \ldots, \alpha_{k,\lfloor n_k/2 \rfloor}$ and B_k be defined as in the proof of Theorem 5 (we may assume that n_k is even and so $\alpha_{k,\lfloor n_k/2 \rfloor + 1}$ is not defined). The analysis of the case of equality in the proof of Theorem 5 implies that

$$\lim_{k \to \infty} \lambda_{k,1} = \sqrt{\frac{1 - 8t(C_3, W)}{3}} \qquad \text{and} \qquad \lim_{k \to \infty} \alpha_{k,1} = 0. \qquad (11)$$

Let $\{Z_{k,1}, \ldots, Z_{k,n_k}\}$ be the partition of the interval $[0, 1]$ corresponding to W_k, and let $w_k = (w_{k,1}, \ldots, w_{k,n_k})$ be the vector v_2 as defined in the proof of Theorem 5. Define a function $f_k : [0, 1] \to \mathbb{R}$ by setting

$$f_k(x) = \frac{\lambda_{k,1} n_k^{1/2}}{2} \left(w_{k,i} - \frac{\sum_{i'=1}^{n_k} w_{k,i'}}{n_k} \right)$$

where $i \in [n_k]$ such that $x \in Z_{k,i}$. It follows from (11) that

$$\lim_{k \to \infty} \int_{x,y \in [0,1]} |W_k(x, y) - (1/2 + f_k(x) - f_k(y))| \; \mathrm{d}x \; \mathrm{d}y = 0. \qquad (12)$$

Also observe that the definition of f_k implies that

$$\int_{x \in [0,1]} f_k(x) = 0. \qquad (13)$$

13

We next define functions $\tilde{f}_k : [0, 1] \to \mathbb{R}$ by setting

$$\tilde{f}_k(x) = \frac{1}{|Z_{k,i}|} \int_{(x',y) \in Z_{k,i} \times [0,1]} W(x', y) \, \mathrm{d}x' \, \mathrm{d}y - \frac{1}{2}$$

for $x \in Z_{k,i}$, $i \in [n_k]$. Note that

$$\tilde{f}_k(x) = \int_{y \in [0,1]} W_k(x, y) \, \mathrm{d}y - \frac{1}{2} \, . \tag{14}$$

In particular, $\tilde{f}_k(x) \in [-1/2, 1/2]$ for all $x \in [0, 1]$. Observe that the just defined functions satisfy that

$$\int_{x \in Z_{k,i}} \tilde{f}_{k'}(x) \, \mathrm{d}x = \int_{x \in Z_{k,i}} \tilde{f}_k(x) \, \mathrm{d}x$$

for every $k \in \mathbb{N}$, $i \in [n_k]$ and $k' \geq k$. In particular, it holds that the sequence $(\tilde{f}_k)_{k \in \mathbb{N}}$ forms a martingale when viewed as a sequence of random variables on $[0, 1]$. So, Doob's Martingale Convergence Theorem yields that the sequence $(\tilde{f}_k)_{k \in \mathbb{N}}$ L_1-converges to a function $\tilde{f} : [0, 1] \to [-1/2, 1/2]$.

We derive by applying the L_1-convergence of $(\tilde{f}_k)_{k \in \mathbb{N}}$, (14), (13), the triangle inequality and (12) (in this order) that

$$\lim_{k \to \infty} \int_{x \in [0,1]} \left| \tilde{f}(x) - f_k(x) \right| \, \mathrm{d}x = \lim_{k \to \infty} \int_{x \in [0,1]} \left| \tilde{f}_k(x) - f_k(x) \right| \, \mathrm{d}x$$

$$= \lim_{k \to \infty} \int_{x \in [0,1]} \left| \int_{y \in [0,1]} W_k(x, y) \, \mathrm{d}y - 1/2 - f_k(x) \right| \, \mathrm{d}x$$

$$= \lim_{k \to \infty} \int_{x \in [0,1]} \left| \int_{y \in [0,1]} W_k(x, y) - (1/2 + f_k(x) - f_k(y)) \, \mathrm{d}y \right| \, \mathrm{d}x$$

$$\leq \lim_{k \to \infty} \int_{x,y \in [0,1]} |W_k(x, y) - (1/2 + f_k(x) - f_k(y))| \, \mathrm{d}x \, \mathrm{d}y = 0 \, .$$

This implies that the sequence $(f_k)_{k \in \mathbb{N}}$ also L_1-converges to the function $\tilde{f}$. It follows from (12) and the L_1-convergence of W_k to W that $W(x, y)$ is equal to $1/2 + \tilde{f}(x) - \tilde{f}(y)$ for almost every $(x, y) \in [0, 1]^2$.

It remains to shift $\tilde{f}$ so that its range lies in $[0, 1/2]$. Let z_0 be the infimum of those values z such that the measure of $\tilde{f}^{-1}((-\infty, z])$ is positive, and define a function $f : [0, 1] \to [0, 1/2]$ as follows:

$$f(x) = \begin{cases} \tilde{f}(x) - z_0 & \text{if } \tilde{f}(x) \in [z_0, z_0 + 1/2], \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Since $W(x, y) = 1/2 + \tilde{f}(x) - \tilde{f}(y)$ for almost every $(x, y) \in [0, 1]^2$ and $W(x, y) \in [0, 1]$ for all $(x, y) \in [0, 1]^2$, the set of $x \in [0, 1]$ such that the second case in the definition of $f(x)$ applies has measure zero. It follows that $W(x, y) = 1/2 + f(x) - f(y)$ for almost every $(x, y) \in [0, 1]^2$ as desired. $\qquad\square$

4 Regime of three parts

Having confirmed Conjecture 1 in the regime of two parts, we now turn towards the next case, namely $1/72 \leq \sigma_3(A) \leq 1/32$ (Theorem 11). Indeed, the proof of Theorem 11 will apply to both regimes, although it does not characterise the extremal tournaments.

We start with analyzing the following optimization problem, which we refer to as the problem Spectrum. This optimization problem is obtained from constraints that (normalized) eigenvalues of a non-negative matrix of order n with trace $n/2$ must satisfy. We state this formally in Lemma 7, which follows the statement of the problem.

Optimization problem Spectrum	
Parameters:	reals $s_3 \in [0, 1/8]$ and $\rho \in [0, 1/2]$
	non-negative integers k and ℓ such that $k + \ell \geq 1$
Variables:	real numbers $r_1, \ldots, r_k$, $a_1, \ldots, a_\ell$ and $b_1, \ldots, b_\ell$
Constraints:	$0 \leq r_1, \ldots, r_k \leq \rho$
	$0 \leq a_1, \ldots, a_\ell$
	$\rho + \sum_{i=1}^{k} r_i + 2 \sum_{i=1}^{\ell} a_i = 1/2$
	$\rho^3 + \sum_{i=1}^{k} r_i^3 + 2 \sum_{i=1}^{\ell} (a_i^3 - 3a_i b_i^2) = s_3$
Objective:	minimize $\rho^4 + \sum_{i=1}^{k} r_i^4 + 2 \sum_{i=1}^{\ell} (a_i^4 - 6a_i^2 b_i^2 + b_i^4)$

Lemma 7. *Let A be a tournament matrix of order n with spectral radius equal to $\rho \cdot n$. Let k be one less than the number of real eigenvalues of A (counting multiplicities) and ℓ the number of conjugate pairs of complex eigenvalues (again counting multiplicities). Further, let $\rho \cdot n, r_1 \cdot n, \ldots, r_k \cdot n$ be the $k+1$ real eigenvalues and $(a_1 \pm \iota b_1)n, \ldots, (a_\ell \pm \iota b_\ell)n$ be the ℓ pairs of complex eigenvalues. Then the numbers $r_1, \ldots, r_k$, $a_1, \ldots, a_\ell$ and $b_1, \ldots, b_\ell$ satisfy all constraints in the optimization problem Spectrum for the parameters $s_3 = \sigma_3(A)$, ρ, k and ℓ.*

Proof. Since $\rho \cdot n$ is the spectral radius of A, it holds that $r_i \leq \rho$ for every $i \in [k]$ and that $\rho \cdot n$ is an eigenvalue of A by the Perron–Frobenius theorem. Since the real part of every eigenvalue of A is non-negative by Lemma 2, all $r_1, \ldots, r_k$ and $a_1, \ldots, a_\ell$ are non-negative. Since the diagonal entries of A are all $1/2$ and trace of A is equal to the sum of its eigenvalues, we have

$$\rho n + \sum_{i=1}^{k} r_i n + 2 \sum_{i=1}^{\ell} a_i n = n/2.$$

Similarly, the trace of A^3 gives us

$$s_3 n^3 = \sigma_3(A)n^3 = \rho^3 n^3 + \sum_{i=1}^{k} r_i^3 n^3 + 2 \sum_{i=1}^{\ell} \left(a_i^3 - 3a_i b_i^2\right) n^3.$$

15

Thus, we conclude that the numbers $r_1, \ldots, r_k$, $a_1, \ldots, a_\ell$ and $b_1, \ldots, b_\ell$ satisfy all constraints in the optimization problem Spectrum. $\qquad\square$

Note that the objective function of Spectrum is precisely $\sigma_4(A)$. Next, we analyze the structure of optimal solutions of the optimization problem Spectrum.

Lemma 8. *Let $r_1, \ldots, r_k$, $a_1, \ldots, a_\ell$ and $b_1, \ldots, b_\ell$ be an optimal solution of the optimization problem Spectrum with the parameters s_3, ρ, k and ℓ. Then, at least one of the following two cases must hold:*

- *There exist positive reals r' and r'' such that $r_1, \ldots, r_k \in \{0, r', r'', \rho\}$, $(a_1, b_1), \ldots, (a_\ell, b_\ell) \in \{(0, 0), (r', 0), (r'', 0)\}$.*

- *There exist reals a' and $b' \neq 0$ such that $r_1, \ldots, r_k \in \{0, \rho\}$ and $(a_1, b_1), \ldots, (a_\ell, b_\ell) \in \{(0, 0), (a', b'), (a', -b')\}$.*

Proof. The method of Lagrange multipliers implies that the gradient of the objective function is a linear combination of the gradient of the two equality constraints when restricted to the entries indexed by $r_i \notin \{0, \rho\}$, by $a_i \neq 0$ and $b_i \neq 0$, i.e., when we are not on the boundary of the feasible set. In particular, the rank of the matrix M with rows being the described restrictions of the three gradient vectors is at most two.

We first analyze the case that one of the numbers $b_1, \ldots, b_\ell$ is non-zero. Our aim is to show that the second case described in the statement of the lemma applies. By symmetry, we can assume that $b_1 \neq 0$. Also note the following holds for every $i \in [\ell]$: if $b_i \neq 0$, then $a_i \neq 0$. Indeed, if $a_i = 0$ and $b_i \neq 0$, then setting $b_i = 0$ does not affect the constraints and decreases the objective function, which contradicts that the solution is optimal. It follows that a_1 is positive.

Suppose that there exist r_i such that $0 < r_i < \rho$. The restriction of the matrix M to the columns corresponding to a_1, b_1 and r_i is the following.

$$\begin{bmatrix} 2 & 0 & 1 \\ 6a_1^2 - 6b_1^2 & -12a_1 b_1 & 3r_i^2 \\ 8a_1^3 - 24a_1 b_1^2 & 8b_1^3 - 24a_1^2 b_1 & 4r_i^3 \end{bmatrix} \tag{15}$$

Dividing the first column by 2 and the second by $4b_1$, dividing the second row by 3 and the third by 2, and subtracting the last column from the first one yields the following matrix, which has the same rank.

$$\begin{bmatrix} 0 & 0 & 1 \\ a_1^2 - b_1^2 - r_i^2 & -a_1 & r_i^2 \\ 2(a_1^3 - 3a_1 b_1^2 - r_i^3) & b_1^2 - 3a_1^2 & 2r_i^3 \end{bmatrix}$$

This matrix is not full rank if and only if the determinant of its submatrix formed by the intersection of the second and third rows with the first and second columns, which is equal to

$$-(a_1^2 + b_1^2)^2 + (3a_1^2 - b_1^2)r_i^2 - 2a_1 r_i^3 ,$$

16

is zero. However, this determinant can be rewritten as

$$-\left(2a_1 r_i + a_1^2 + b_1^2\right)\left((r_i - a_1)^2 + b_1^2\right),$$

which is negative since $a_1 > 0$ and $b_1 \neq 0$. It follows that $r_i \in \{0, \rho\}$ for all $i \in [k]$.

Further suppose that there exists an index $i \in [\ell]$ for which it holds $(a_i, b_i) \notin \{(0,0), (a_1, b_1), (a_1, -b_1)\}$. If $b_i = 0$, then the restriction of the matrix M to the columns corresponding to a_1, b_1 and a_i is the same as the restriction of the matrix M considered in the previous paragraph with r_i replaced by a_i and the corresponding column multiplied by two. In particular, the restriction cannot have rank two in this case. Hence, we can assume that $b_i \neq 0$ and so $a_i > 0$ (the argument is the same as when we argued that $a_1 > 0$). The restriction of the matrix M to the columns corresponding to a_1, b_1, a_i and b_i is the following.

$$\begin{bmatrix} 2 & 0 & 2 & 0 \\ 6a_1^2 - 6b_1^2 & -12a_1 b_1 & 6a_i^2 - 6b_i^2 & -12a_i b_i \\ 8a_1^3 - 24a_1 b_1^2 & 8b_1^3 - 24a_1^2 b_1 & 8a_i^3 - 24a_i b_i^2 & 8b_i^3 - 24a_i^2 b_i \end{bmatrix}$$

The rank of this matrix is the same as the rank of the following matrix (the rows are multiplied by $1/2$, $1/6$ and $1/4$, the columns by 1, $-a_1/2b_1$, 1 and $-a_i/2b_i$, respectively).

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ a_1^2 - b_1^2 & a_1^2 & a_i^2 - b_i^2 & a_i^2 \\ 2a_1^3 - 6a_1 b_1^2 & 3a_1^3 - a_1 b_1^2 & 2a_i^3 - 6a_i b_i^2 & 3a_i^3 - a_i b_i^2 \end{bmatrix}$$

By subtracting twice the second column from the first column and twice the fourth column from the third column, we obtain the following matrix.

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ -(a_1^2 + b_1^2) & a_1^2 & -(a_i^2 + b_i^2) & a_i^2 \\ -4a_1(a_1^2 + b_1^2) & 3a_1^3 - a_1 b_1^2 & -4a_i(a_i^2 + b_i^2) & 3a_i^3 - a_i b_i^2 \end{bmatrix}$$

Since the last row of the matrix is a linear combination of the previous two rows (the operation that we have performed so far has preserved this property of the matrix M), it follows that

$$\frac{3a_1^2 - b_1^2}{a_1} = \frac{3a_i^2 - b_i^2}{a_i} \tag{16}$$

We now subtract the second row multiplied by the value given in (16) from the third row and obtain the following matrix, which has the rank two.

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ -(a_1^2 + b_1^2) & a_1^2 & -(a_i^2 + b_i^2) & a_i^2 \\ -(a_1^2 + b_1^2)^2/a_1 & 0 & -(a_i^2 + b_i^2)^2/a_i & 0 \end{bmatrix}$$

It follows that
$$\frac{(a_1^2 + b_1^2)^2}{a_1} = \frac{(a_i^2 + b_i^2)^2}{a_i},$$

which yields that
$$b_i^2 = -a_i^2 + \sqrt{\frac{a_i}{a_1}}(a_1^2 + b_1^2). \tag{17}$$

On the other hand, we derive from (16) that
$$b_i^2 = 3a_i^2 - \frac{a_i}{a_1}\left(3a_1^2 - b_1^2\right). \tag{18}$$

We obtain by comparing (17) and (18) the following.
$$0 = 4a_i^2 - \frac{a_i}{a_1}\left(3a_1^2 - b_1^2\right) - \sqrt{\frac{a_i}{a_1}}(a_1^2 + b_1^2)$$
$$= \sqrt{a_i}(\sqrt{a_i} - \sqrt{a_1})\left(4a_i + 4\sqrt{a_1 a_i} + \frac{a_1^2 + b_1^2}{a_1}\right)$$

Since both a_1 and a_i are positive, this expression can be equal to zero only if $a_i = a_1$. Consequently, the equality (17) implies that $b_i = b_1$ or $b_i = -b_1$, which contradicts the choice of (a_i, b_i). Hence, we have established that if at least one of $b_1, \ldots, b_\ell$ is non-zero, then the second case indeed applies.

We now consider the case that $b_1 = \cdots = b_\ell = 0$. Suppose that the first case in the statement of the lemma does not apply. This implies that there exist three distinct positive reals α, β and γ such that at least one of the values $r_1, \ldots, r_k, a_1, \ldots, a_\ell$ is α, at least one is β and at least one is γ. Consequently, the matrix M contains the following submatrix possibly after dividing some columns by two (the columns correspond to those of the variables $r_1, \ldots, r_k, a_1, \ldots, a_\ell$ that are equal to α, β and γ, respectively; the columns corresponding to the variables $a_1, \ldots, a_\ell$ are divided by two).

$$\begin{bmatrix} 1 & 1 & 1 \\ 3\alpha^2 & 3\beta^2 & 3\gamma^2 \\ 4\alpha^3 & 4\beta^3 & 4\gamma^3 \end{bmatrix}$$

The determinant of this matrix is equal to
$$12(\alpha^2\beta^3 + \beta^2\gamma^3 + \gamma^2\alpha^3 - \alpha^2\gamma^3 - \beta^2\alpha^3 - \gamma^2\beta^3),$$

which is equal to
$$12(\alpha - \beta)(\beta - \gamma)(\gamma - \alpha)(\alpha\beta + \alpha\gamma + \beta\gamma).$$

Since this expression is non-zero for all distinct positive reals α, β and γ, we conclude that the rank of M is three, which contradicts our assumption that the first case as described in the statement of the lemma does not apply. $\qquad\square$

Before we can prove the main result of this section, we need two additional auxiliary lemmas.

Lemma 9. *Let* $z \in (0,1]$, *and let* $x_1, \ldots, x_n$ *be non-negative reals such that* $x_1 + \cdots + x_n = 1/2$ *and* $x_i \le z$ *for every* $i \in [n]$. *Then it holds that*

$$\sum_{i=1}^{n} x_i^3 \le \lfloor z^{-1}/2 \rfloor \cdot z^3 + (1/2 - \lfloor z^{-1}/2 \rfloor \cdot z)^3,$$

and the equality holds if and only if all but at most one of $x_1, \ldots, x_n$ *are equal to* 0 *or* z.

Proof. Consider any n-tuple $x_1, \ldots, x_n$ that maximizes the sum $x_1^3 + \cdots + x_n^3$ among all n-tuples of non-negative reals $x_1, \ldots, x_n$ such that $x_1 + \cdots + x_n = 1/2$ and $x_i \le z$, $i \in [n]$. If $x_i \in \{0, z\}$ for all but at most one $i \in [n]$, then the sum of $x_1^3 + \cdots + x_n^3$ is equal to $\sum_{i=1}^{n} x_i^3 \le \lfloor z^{-1}/2 \rfloor \cdot z^3 + (1/2 - \lfloor z^{-1}/2 \rfloor \cdot z)^3$ and the lemma holds. Otherwise, there exist x_i and x_j such that $0 < x_i \le x_j < z$. Choose $\varepsilon > 0$ such that $\varepsilon < x_i$ and $\varepsilon < z - x_j$, and replace x_i with $x_i - \varepsilon$ and x_j with $x_j + \varepsilon$. This preserves the sum $x_1 + \cdots + x_n$ and increases the sum $x_1^3 + \cdots + x_n^3$, which contradicts the choice of the n-tuple $x_1, \ldots, x_n$. $\square$

Linial and Morgenstern [9] proved that, among the random blow-ups of transitive tournaments with the fixed density of C_3, the density of C_4 is minimized if all parts have the same size except possibly for a single smaller part. This statement is equivalent to the following.

Lemma 10 (Linial and Morgenstern [9, Lemma 2.7]). *Let* $x_1, \ldots, x_n$ *be any non-negative reals such that their sum is* $1/2$. *It holds that*

$$x_1^4 + \cdots + x_n^4 \ge g(x_1^3 + \cdots + x_n^3).$$

We are now ready to prove the main result of this section.

Theorem 11. *Let* A *be a tournament matrix of order* n. *If* $\sigma_3(A) \ge 1/72$, *then* $\sigma_4(A) \ge g(\sigma_3(A))$.

Proof. Let $s_3 = \sigma_3(A)$. We start with lower bounding the spectral radius of A. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A. By Lemma 2, the real parts of all the eigenvalues are non-negative, which implies that

$$s_3 = \sum_{i=1}^{n} \left(\frac{\lambda_i}{n}\right)^3 \le \sum_{i=1}^{n} \left(\frac{\operatorname{Re} \lambda_i}{n}\right)^3.$$

By Lemma 9, the last sum is at most

$$\lfloor \rho_A^{-1}/2 \rfloor \cdot \rho_A^3 + (1/2 - \lfloor \rho_A^{-1}/2 \rfloor \cdot \rho_A)^3$$

19

where ρ_A is the spectral radius of A divided by n. Consequently, ρ_A is at least z where z is the unique real between 0 and $1/2$ satisfying that

$$s_3 = \lfloor z^{-1}/2 \rfloor \cdot z^3 + (1/2 - \lfloor z^{-1}/2 \rfloor \cdot z)^3 \,.$$

Note that $z \geq 1/6$ since $s_3 \geq 1/72$.

Lemma 7 now yields that the theorem will be proven if we show that the optimal solution of the problem Spectrum is at least $g(s_3)$ for s_3, any $\rho \geq z$ and all non-negative integers k and ℓ. By Lemma 8, this would be implied by the following two claims, which correspond to the two cases described in the statement of Lemma 8.

Claim 1. If $r_1, \ldots, r_k$ are any positive real numbers that have at most three distinct values and that satisfy $r_1 + \cdots + r_k = 1/2$ and $r_1^3 + \cdots + r_k^3 = s_3$, then $r_1^4 + \cdots + r_k^4 \geq g(s_3)$.

Claim 2. If m and m' are positive integers, $\rho \geq z$, a is a non-negative real and b is a real such that $m\rho + 2m'a = 1/2$ and $m\rho^3 + 2m'(a^3 - 3ab^2) = s_3$, then $m\rho^4 + 2m'(a^4 - 6a^2b^2 + b^4) \geq g(s_3)$.

Claim 1 follows from Lemma 10 (even without the restriction on the number of the distinct values that $r_1, \ldots, r_k$ may have). So, we focus on proving Claim 2 in the remainder of the proof. Note that it holds that $m \in \{1, 2\}$ in this case since $\rho \geq z \geq 1/6$.

To prove Claim 2, we fix m and m' and consider the following optimization problem: minimize the sum $m\rho^4 + 2m'(a^4 - 6a^2b^2 + b^4)$ subject to $m\rho + 2m'a = 1/2$, $m\rho^3 + 2m'(a^3 - 3ab^2) = s_3$, $\rho \geq z$ and $a \geq 0$. The method of Langrange multipliers implies that the following matrix is not full rank

$$\begin{bmatrix} m & 2m' & 0 \\ 3m\rho^2 & 6m'a^2 - 6m'b^2 & -12m'ab \\ 4m\rho^3 & 8m'a^3 - 24m'ab^2 & 8m'b^3 - 24m'a^2b \end{bmatrix} \tag{19}$$

for the values of ρ, a and b that minimize the sum unless $\rho = z$ or $a = 0$. However, dividing the first column by m and the remaining two by m', permuting the columns and renaming the variables yields the same matrix as in (15), which we have analyzed in the proof of Lemma 8. In particular, the matrix in (19) has full rank unless $a = 0$ or (possibly) $b = 0$. We conclude that the expression $m\rho^4 + 2(a^4 - 6a^2b^2 + b^4)$ is minimized when at least one of the following holds: $\rho = z$, $a = 0$ or $b = 0$. We next analyze these three cases.

The case $\rho = z$. In this case, Lemma 9 implies that $m\rho^3 + 2m'a^3 < s_3$ unless $a \geq \rho$, i.e., there is no such feasible solution unless $a \geq \rho$. If indeed $a \geq \rho$, then since $m\rho + 2m'\rho = 1/2$ and $\rho \geq 1/6$, it must be that $m = m' = 1$, $\rho = a = 1/6$, $s_3 = 1/72$ and $b = 0$, in which case $m\rho^4 + 2m'(a^4 - 6a^2b^2 + b^4) = 1/432 = g(s_3)$.

The case $a = 0$. If $a = 0$, then, as noted in the proof of Lemma 8, setting $b = 0$ does not affect the constraints, but does decrease the objective function. Hence this case reduces to the final case $b = 0$.

The case $b = 0$. If $b = 0$, then using Lemma 10 we get that $m\rho^4 + 2m'a^4 \geq g(m\rho^3 + 2m'a^3) = g(s_3)$.

Hence, we have shown that $m\rho^4 + 2m'(a^4 - 6a^2b^2 + b^4) \geq g(s_3)$ for all ρ, a and b such that $m\rho + 2m'a = 1/2$, $m\rho^3 + 2m'(a^3 - 3ab^2) = s_3$, $\rho \geq z$ and $a \geq 0$. The proof of Claim 2 is now finished and so is the proof of the theorem. $\qquad\square$

Proposition 3 yields the following corollary in the tournament limit setting.

Corollary 12. *Let W be a tournament limit. If $t(C_3, W) \geq 1/72$, then it holds that $t(C_4, W) \geq g(t(C_3, W))$.*

5 Concluding remarks

Unfortunately, the proof of Theorem 11 does not immediately work in higher regimes because, in the case of four or more parts, the solution to the optimization problem Spectrum beats the conjectured minimum. Of course, as solutions need not be realizable as the eigenvalues of a tournament matrix, this does not invalidate Conjecture 1. It is possible that the current method can be pushed further by introducing to the optimization problem additional constraints that reflect the properties that eigenvalues of tournament matrices must satisfy.

Meanwhile, in the regime of two parts, we have been able to fully determine the asymptotic structure of extremal examples. These constructions can be extended to the remaining regimes as follows. Fix $k \in \mathbb{N}$, $z \in [1/(k+1), 1/k]$ and $i, i' \in [k+1]$ such that $|i - i'| = 1$. We construct a tournament T with n vertices as follows. The vertices of T are split into $k + 1$ parts $V_1, \ldots, V_{k+1}$ such that $|V_i| = n - k\lfloor zn \rfloor$ and $|V_j| = \lfloor zn \rfloor$ if $j \neq i$. If two vertices v and v' respectively belong to distinct parts V_j and $V_{j'}$ with $j < j'$ and $\{j, j'\} \neq \{i, i'\}$, then the tournament T contains an arc from v to v'. If, instead, v and v' belong to the *same* part V_j, where $j \notin \{i, i'\}$, then the edge between v and v' is oriented from v to v' with probability $1/2$, i.e., the vertices of every such part induce a randomly oriented tournament. Finally, each vertex $v \in V_i \cup V_{i'}$ is assigned a real number $p_v \in [0, 1/2]$ and the edge between v and $v' \in V_i \cup V_{i'}$ is directed from v to v' with probability $1/2 + p_v - p_{v'}$. If the expected number of triangles in T is equal to $\frac{1}{8}\left(k\lfloor zn \rfloor^3 + (n - k\lfloor zn \rfloor)^3\right)$, then the expected value of the density $t(C_4, T)$ is

$$g\left(\frac{1}{8}\left(k\lfloor zn \rfloor^3 + (n - k\lfloor zn \rfloor)^3\right)\right) + o(1)$$

and both of these random variables are concentrated. In particular, unless z^{-1} is a positive integer, there are infinitely many different types of extremal tournaments.

As discussed in the introduction, we believe this is the reason why the problem is resistant to the standard flag algebra techniques.

It is also interesting to note that the problem of determining the set of feasible densities of cycles of length three and four is equivalent to the analogous problems for transitive tournaments of order three and four and for the cycle and transitive tournament of order four [9, Proposition 1.1]. To see this, let T_k be the transitive tournament of order k and let $t(T_k, T)$ be the probability that a random mapping from $V(T_k)$ to $V(T)$ is a homomorphism. The following holds for every n-vertex tournament T:

$$8t(C_3, T) + 24t(T_4, T) - 6t(C_4, T) = 1 - O(n^{-1}).$$

Thus, the problem of minimizing the density of C_4 when the density of C_3 is fixed is equivalent to minimizing the density of T_4 when the density of T_3 is fixed, and also equivalent to minimizing the density of C_4 when T_4 is fixed, in the sense that a complete solution to any of these three problems yield complete solutions to the remaining two.

References

[1] B. Bollobás: *Relations between sets of complete subgraphs*, in: Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen, 1975) (1976), 79–84.

[2] A. Brauer and I. C. Gentry: *On the characteristic roots of tournament matrices*, Bull. Amer. Math. Soc. **74** (1968), 1133–1135.

[3] F. R. K. Chung and R. L. Graham: *Quasi-random tournaments*, J. Graph Theory **15** (1991), 173–198.

[4] P. Erdős: *Some theorems on graphs*, Riveon Lematematika **9** (1955), 13–17.

[5] D. C. Fisher: *Lower bounds on the number of triangles in a graph*, J. Graph Theory **13** (1989), 505–512.

[6] A. Frieze and R. Kannan: *Quick approximation to matrices and applications*, Combinatorica **19** (1999), 175–200.

[7] F. R. Gantmacher: The theory of matrices. Vol. 1, 1998, translated from the Russian by K. A. Hirsch, Reprint of the 1959 translation.

[8] A. W. Goodman: *On sets of acquaintances and strangers at any party*, Amer. Math. Monthly **66** (1959), 778–783.

[9] N. Linial and A. Morgenstern: *On the number of 4-cycles in a tournament*, J. Graph Theory **83** (2016), 266–276.

[10] H. Liu, O. Pikhurko and K. Staden: *The exact minimum number of triangles in graphs of given order and size*, preprint arXiv:1712.00633 (2017).

[11] L. Lovász: Large Networks and Graph Limits, *Colloquium Publications*, volume 60, 2012.

[12] L. Lovász and M. Simonovits: *On the number of complete subgraphs of a graph. II*, in: Studies in pure mathematics (1983), 459–495.

[13] L. Lovász and B. Szegedy: *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), 933–957.

[14] W. Mantel: *Problem 28*, Wiskundige Opgaven **10** (1907), 60–61.

[15] V. S. Nikiforov: *The number of cliques in graphs of given order and size*, Trans. Amer. Math. Soc. **363** (2011), 1599–1618.

[16] O. Pikhurko and A. Razborov: *Asymptotic structure of graphs with the minimum number of triangles*, Combin. Probab. Comput. **26** (2017), 138–160.

[17] A. A. Razborov: *On the minimal density of triangles in graphs*, Combin. Probab. Comput. **17** (2008), 603–618.

[18] C. Reiher: *The clique density theorem*, Ann. of Math. (2) (2016), 683–707.