Tournament quasirandomness from local counting

Matija Bucić[*]   Eoin Long[†]   Asaf Shapira[‡]   Benny Sudakov[§]

Abstract

A well-known theorem of Chung and Graham states that if $h \geq 4$ then a tournament T is quasirandom if and only if T contains each h-vertex tournament the 'correct number' of times as a subtournament. In this paper we investigate the relationship between quasirandomness of T and the count of a *single* h-vertex tournament H in T. We consider two types of counts, the global one and the local one.

We first observe that if T has the correct *global* count of H and $h \geq 7$ then quasirandomness of T is only forced if H is transitive. The next natural question when studying quasirandom objects asks whether possessing the correct *local* counts of H is enough to force quasirandomness of T. A tournament H is said to be locally forcing if it has this property.

Variants of the local forcing problem have been studied before in both the graph and hypergraph settings. Perhaps the closest analogue of our problem was considered by Simonovits and Sós who looked at whether having 'correct counts' of a fixed graph H as an induced subgraph of G implies G must be quasirandom, in an appropriate sense. They proved that this is indeed the case when H is regular and conjectured that it holds for all H (except the path on 3 vertices). Contrary to the Simonovits-Sós conjecture, in the tournament setting we prove that a constant proportion of all tournaments are not locally forcing. In fact, any locally forcing tournament must itself be strongly quasirandom. On the other hand, unlike the global forcing case, we construct infinite families of non-transitive locally forcing tournaments.

1   Introduction

A combinatorial structure is said to be 'quasirandom' if it behaves in a similar manner to a random structure, where the comparison is made with respect to some deterministic property. The systematic study of quasirandomness was initiated by Thomason [37] [36], and Chung, Graham and Wilson [8], who examined notions of quasirandomness arising from various graph properties. One of the surprising conclusions of these papers is that a wide-range of natural graph properties all lead to essentially the same notion of quasirandomness, in the sense that a graph satisfying one of the properties necessarily satisfies them all. Since then, notions of quasirandomness have been extensively studied in a wide variety of contexts, including hypergraphs [7], [18], [28], [30], permutations [11], [25] and groups [19]. The reader is referred to the survey [26] for an overview of this extensive topic.

In this paper we will study notions of quasirandomness for tournaments. The first paper on this topic is due to Chung and Graham [6] who proved that, as in the graph case, a wide range of natural tournament properties give rise to the same notion of quasirandomness. Before stating some of their results we require a little notation. A tournament $T = (V, E)$ consists of a set of vertices $V = V(T)$,

together with a set of edges $E = E(T) \subset V \times V$, with the property that (i) $(u, u) \notin E$ for all $u \in V$ and (ii) exactly one of (u, v) or (v, u) lies in E for all distinct $u, v \in V$. We often write $\overrightarrow{uv}$ to denote an edge (u, v). A tournament H appears as a subtournament of T if there is a map $\phi : V(H) \to V(T)$ such that $\overrightarrow{uv} \in E(H)$ if and only if $\overrightarrow{\phi(u)\phi(v)} \in E(T)$. The map ϕ is said to be a labelled embedding of H into T. Let

$$N_T^*(H) = \left| \left\{ \phi : V(H) \to V(T) : \phi \text{ is a labelled embedding of } H \text{ into } T \right\} \right|.$$

Given $U \subset V(T)$ let $T[U]$ denote the subtournament of T induced by the vertex set U. Also let $N_T^*(H; U) := N_{T[U]}^*(H)$. For $u \in V(T)$ let $d_T^+(u) = |\{v \in V(T) : \overrightarrow{uv} \in E(T)\}|$ and $d_T^-(u) = |\{v \in V(T) : \overrightarrow{vu} \in E(T)\}|$. A tournament T is regular if $d_T^+(u) = d_T^-(u)$ for all $u \in V(T)$. For $U, W \subseteq V(T)$ we denote by $e(U, W)$ the number of edges starting in U and ending in W.

We say that T is an n-vertex tournament if $|V(T)| = n$. An ordering of T is a bijective map $\sigma : V \to [n]$ and the set of all orderings of T is naturally identified with S_n, the symmetric group on n elements. An edge $\overrightarrow{uv} \in E$ is a σ-forward edge if $\sigma(u) < \sigma(v)$ and we write $F_{\sigma, T} \subset E$ to denote the set of σ-forward edges of T. Let $B_{\sigma, T} = E \setminus F_{\sigma, T}$, the set of σ-backward edges. T is said to be transitive if $F_{\sigma, T} = E$ for some $\sigma \in S_n$ and write Tr_n to denote the unique (up to isomorphism) n-vertex transitive tournament. Lastly, we will write $a \pm b$ to denote some value c with $a - b \le c \le a + b$.

Our starting point in this paper is a result due to Chung and Graham [6], which gives two equivalences of tournament quasirandomness.

Theorem 1.1 (Chung–Graham). *Let $h \in \mathbb{N}$ with $h \ge 4$. Then for any n-vertex tournament T, the following properties are equivalent:*
- $\mathcal{P}_1$: $|F_{\sigma, T}| = \frac{1}{2}\binom{n}{2} \pm o(n^2)$ for every ordering σ of T.
- $\mathcal{P}_2(h)$: $N_T^*(H) = 2^{-\binom{h}{2}} n^h \pm o(n^h)$ for every h-vertex tournament H.

In fact, there are nine further equivalences given in this paper (see also [20], [23]). Equivalence here is understood in the following sense, focusing on the implication $\mathcal{P}_1 \implies \mathcal{P}_2(h)$: given $\varepsilon > 0$ there is $\delta > 0$ and $n_0 \in \mathbb{N}$ so that if T is an n-vertex tournament which satisfies $|F_{\sigma, T}| = \frac{1}{2}\binom{n}{2} \pm \delta n^2$ for every ordering σ of T and $n \ge n_0$ then $N_T^*(H) = 2^{-\binom{h}{2}} n^h \pm \varepsilon n^h$ for every h-vertex tournament H.

It is easily seen that both properties $\mathcal{P}_1$ and $\mathcal{P}_2(h)$ hold for a random n-vertex tournament T with high probability. We will say T is *quasirandom* if it satisfies $\mathcal{P}_1$, that is $|F_{\sigma, T}| = \frac{1}{2}\binom{n}{2} \pm o(n^2)$. In light of Theorem 1.1, this notion is equivalent to the analogous notion arising from $\mathcal{P}_2(h)$.

1.1 Globally forcing tournaments

Although $\mathcal{P}_2(h)$ guarantees that T is quasirandom, it is natural to ask whether this can already be deduced from the count of a *single* tournament H. That is, does quasirandomness of T already follow if $N_T^*(H) = 2^{-\binom{h}{2}} n^h \pm o(n^h)$ for a single h-vertex tournament H?

Definition. *Let H be an h-vertex tournament. Given an n-vertex tournament T, consider the following property:*
- $\mathcal{P}_2(H)$: $N_T^*(H) = 2^{-\binom{h}{2}} n^h \pm o(n^h)$.

The tournament H is said to be globally forcing *if $\mathcal{P}_2(H) \implies \mathcal{P}_1$.*

It follows easily from exercise 10.44(b) of [27] that for $h \ge 4$ each transitive tournament Tr_h is globally forcing. This statement was recently reproved in the language of flag-algebras by Coregliano and Razborov [13]. Furthermore, Coregliano, Parente and Sato [12] constructed a 5-vertex non-transitive globally forcing tournament T_5. Our first observation is that there are no further globally forcing tournaments.

Proposition 1.2. *A tournament H is globally forcing if and only if it is either transitive with 4 or more vertices or if $H = T_5$.*

In this paper we establish the above proposition for tournaments on 7 or more vertices. After the first version of our paper appeared on arXiv, Hancock, Kabela, Kral, Martins, Parente, Skerman and Volec [22] managed to resolve the remaining cases of 5 and 6 vertex tournaments building on some of our ideas.

1.2 Locally forcing tournaments

We have seen that having 'correct count' of a fixed tournament H is not enough to guarantee quasirandomness, for essentially any non-transitive H. This is a fairly common situation when studying quasirandom properties in general and the key insight to understand why came from Simonovits and Sós [32] who observed that quasirandomness is a hereditary property, in the sense that any large subgraph of a random structure must also be random-like.

This leads us to the natural question of whether requiring that T contains the 'correct count' of H in all large subsets of $V(T)$ is sufficient to guarantee quasirandomness of T. To be more precise, consider the following definition.

Definition. *Let H be an h-vertex tournament. Given an n-vertex tournament T, consider the following property:*
- $\mathcal{P}_2^*(H)$: *$N_T^*(H; U) = 2^{-\binom{h}{2}} |U|^h \pm o(n^h)$ for every set $U \subset V(T)$.*

The tournament H is said to be locally forcing *if $\mathcal{P}_2^*(H) \implies \mathcal{P}_1$.*

An analogous property for graphs was studied by Simonovits and Sós for both induced [32] and not-necessarily induced subgraphs [33] as well as for hypergraphs by Conlon, Hàn, Person and Schacht [10] and Dellamonica and Rödl [14]. In the case of graphs the not-necessarily induced case turned out to be much easier and was resolved by Simonovits and Sós [33] who showed that all non-empty graphs must be 'locally forcing'. This was recently reproved, without use of the regularity lemma, by Conlon, Fox and Sudakov [9]. On the other hand the induced case is still very much open and seems to be the closer analogue to our problem. Indeed, in the case of embedding a tournament one needs to ensure that each pair of vertices get mapped to an edge with one of two possible orientations, while in the induced case each pair needs to be sent to one of two possible states (is an edge or is not an edge). One of the main conjectures in the area, due to Simonovits and Sós [33] says that all graphs on 4 or more vertices must be 'locally forcing' in the appropriate induced sense. They prove their conjecture holds for regular graphs and for various other families of graphs.

In the rest of this subsection we present our main results, which give a good understanding of locally forcing tournaments and show a surprisingly different behaviour compared to the one conjectured by Simonovits and Sós in the graph case. Our first result shows that in order for a tournament to be locally forcing it must be quite strongly quasirandom in several ways.

Theorem 1.3. *Any non-transitive, locally forcing h-vertex tournament H satisfies*
(i)
$$\sum_{v \in V(H)} \left(d_H^+(v) - d_H^-(v) \right)^2 \leq h(h-1).$$

(ii) $|F_{\sigma,H}| = \frac{1}{2}\binom{h}{2} \pm h^{3/2}\sqrt{\log h}$ for every ordering σ of H.
(iii) For any disjoint subsets $U, W \subseteq H$ we must have $e(U, W) \leq \frac{1}{2}|U||W| + 2h^{3/2}$

Given that Theorem 1.3 says that any locally forcing graph must be strongly quasirandom a natural guess for an example might be to take the random tournament, obtained by orienting each edge of a complete graph uniformly at random, independently between edges. Our next result shows that quite surprisingly with positive probability the random tournament fails condition (i) of Theorem 1.3.

Corollary 1.4. *The random tournament is not locally forcing with positive probability.*

This tells us that a positive proportion of all tournaments are in fact not locally forcing, in stark contrast to the induced graph problem in which Simonovits-Sós conjecture states that all graphs on 4 or more vertices should be locally forcing. In particular, one can easily find a regular tournament

which fails conditions (ii) or (iii) whereas in the induced graph case any regular graph is known to be locally forcing.

This might suggest that there are essentially no non-transitive locally forcing tournaments, as in the global forcing case. We show that this is also not the case and find many non-transitive examples of locally forcing tournaments.

Theorem 1.5. *For any large enough h there exist non-transitive, h-vertex, locally forcing tournaments.*

Our examples of locally forcing tournaments come from the following random construction. We take a Steiner triple system on h vertices, that is a partition of the edge set of the complete graph on h vertices into edge disjoint triangles. Now we orient each of these triangles into one of two cycles of length three uniformly at random and independently between triangles. We show that this produces with high probability a locally forcing tournament. It is well known (see [24]) that Steiner triple systems exist provided $h \equiv 1, 3 \pmod 6$ and our arguments allow us a lot of freedom to modify the above procedure in order to obtain examples for other h as well.

All our results are based on a characterisation of locally forcing tournaments in terms of certain graph polynomials. In order to motivate how these polynomials arise let us describe two natural candidates for tournaments with correct local counts of a tournament H, which are not quasirandom.

Definition. *Given $\alpha \in [0,1]$, let $\mathcal{T}(n, \alpha)$ denote the distribution on the set of tournaments with vertex set $\{1, \ldots, n\}$ in which each pair $\{i, j\}$ with $1 \leq i < j \leq n$ appears independently as $\overrightarrow{ij}$ with probability α. We denote by $\mathcal{T}_{cliq}$ the collection of all distributions $\mathcal{T}(n, \alpha)$ as α varies.*

Loosely speaking we say that a graph H is $\mathcal{T}_{cliq}$-forcing if no $\mathcal{T}(n, \alpha)$ shows that H is not locally forcing. In other words if there is no $\alpha \neq 1/2$ such that $\mathcal{T}(n, \alpha)$, which is clearly not quasirandom, has the same local counts of H as $\mathcal{T}(n, 1/2)$ (which *is* quasirandom).

Definition. *Given $\alpha \in [0,1]$, let $\mathcal{T}(n, n, \alpha)$ denote the distribution on the set of tournaments with vertex set $\{1, \ldots, 2n\}$ in which both $\{1, \ldots, n\}$ and $\{n+1, \ldots, 2n\}$ are oriented according to $\mathcal{T}(n, 1/2)$ and each pair $\{i, j\}$ with $1 \leq i \leq n, n+1 \leq j \leq 2n$ appears as $\overrightarrow{ij}$ with probability α, with all choices made independently. We denote by $\mathcal{T}_{bip}$ the collection of all distributions $\mathcal{T}(n, n, \alpha)$ as α varies.*

Loosely speaking a graph H is $\mathcal{T}_{bip}$-forcing if no $\mathcal{T}(n, n, \alpha)$ shows that H is not locally forcing. I.e. if there is no $\alpha \neq 1/2$ such that $\mathcal{T}(n, n, \alpha)$, which is not quasirandom, has the same local counts of H as $\mathcal{T}(n, n, 1/2) = \mathcal{T}(2n, 1/2)$.

Both properties $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing can be expressed in terms of certain polynomial equations depending on the graph H. Our actual definitions of both properties go directly via these polynomial equations and we refer the reader to Section 3 for exact definitions.

Quite remarkably it turns out that if H is not locally forcing one must be able to find an α such that either $\mathcal{T}(n, \alpha)$ or $\mathcal{T}(n, n, \alpha)$ show H is not locally forcing. In particular, we show:

Theorem 1.6. *H is locally forcing if and only if it is $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing.*

This result allows us to answer the question of whether a fixed graph H is locally forcing or not in terms of whether certain polynomial equations have a common real root in a bounded interval. This can be answered using a combination of Sturm's Theorem (see [5]) and the Euclidean Algorithm for computing the greatest common divisor of a sequence of polynomials. See the concluding remarks for more details.

Organisation of the paper: The short proof of Proposition 1.2 is given in the next section. In Section 3 a counting polynomial associated with $\mathcal{T}_{cliq}$-forcing is introduced. In the same section we show that a locally forcing tournament must be $\mathcal{T}_{cliq}$-forcing and that this property imposes many restrictions on H, which gives us a proof of Theorem 1.3. Despite this in Section 4 we find many $\mathcal{T}_{cliq}$-forcing tournaments. In Sections 5 and 6 we introduce the polynomials related to the property of $\mathcal{T}_{bip}$-forcing, show that any nearly regular tournament must be $\mathcal{T}_{cliq}$-forcing and show Theorem 1.6. Using this we prove that many of our examples of $\mathcal{T}_{cliq}$-forcing tournaments are in fact locally forcing.

4

Notation: Before closing this section we introduce some further notation. Given $m, n \in \mathbb{N}$, let $[n] = \{1, \ldots, n\}$ and $(n)_m$ denote the falling factorial $(n)_m = n(n-1) \cdots (n - m + 1)$. Given a set X we write $\binom{X}{k}$ for the collection of k-element subsets of X. Given a probability distribution $\mathcal{D}$ on X we write $x \sim \mathcal{D}$ to denote an element $x \in X$ selected randomly according to $\mathcal{D}$. We simply write $x \sim X$ when $\mathcal{D}$ is taken to be the uniform distribution on X. All our logarithms are natural so in base e.

2 Globally forcing tournaments

In this section we will give the short proof of Proposition 1.2. We first show that asymptotically the number of transitive subtournaments of order h in a tournament of order n is minimised by the random tournament. This is precisely the content of Exercise 10.44(b) of [27]. We need a variation of this proof, which allows us to also show the fact that any asymptotic minimiser must be quasirandom. This implies that Tr_h is globally forcing for $h \geq 4$ showing the 'if' part of the statement of Proposition 1.2.

Proof of Proposition 1.2. We will first prove that $N_T^*(h) := N_T^*(Tr_h)$ satisfies $N_T^*(h) \geq f_h(n) = (2^{-\binom{h}{2}} - o(1))n^h$ for all n-vertex tournaments T where

$$f_h(n) = \begin{cases} \prod_{j=0}^{h-1} \left(\frac{n+1}{2^j} - 1 \right) & \text{if } n \geq 2^{h-1} - 1 \\ 0 & \text{else.} \end{cases}$$

We will prove this by induction on h. For $h = 1$ and 2 the result is immediate. Furthermore,

$$N_T^*(3) = \sum_{v \in V(T)} \binom{d^+(v)}{2} \geq n \binom{\frac{n-1}{2}}{2} = n(n-1)(n-3)/8 = f_3(n). \tag{1}$$

Now assume $h \geq 4$ and $n \geq 2^{h-1} - 1$ as otherwise the result is trivial. For each edge $e = \overrightarrow{uv} \in E(T)$ let $N(e) = \{w \in V(T) : \overrightarrow{uw}, \overrightarrow{vw} \in E(T)\}$. Clearly

$$N_T^*(3) = \sum_{e \in E(T)} |N(e)|. \tag{2}$$

Letting $T[N(e)]$ be the subtournament of T with vertex set $N(e)$, by induction on h

$$N_T^*(h) = \sum_{e \in E(T)} N_{T[N(e)]}^*(h-2) \geq \sum_{e \in E(T)} f_{h-2}(|N(e)|).$$

It is easy to see that f_h is convex so Jensen's inequality together with (2) implies

$$N_T^*(h) \geq \binom{n}{2} f_{h-2} \left(\frac{N_T^*(3)}{\binom{n}{2}} \right) = n \cdot \frac{n-1}{2} \cdot f_{h-2} \left(\frac{n-3}{4} \right)$$

$$= n \cdot \frac{n-1}{2} \cdot \prod_{j=0}^{h-3} \left(\frac{n+1}{2^{j+2}} - 1 \right)$$

$$= \prod_{j=0}^{h-1} \left(\frac{n+1}{2^j} - 1 \right). \tag{3}$$

Where in the second equality we used that $\frac{n-3}{4} \geq \frac{2^{h-1}-1-3}{4} = 2^{h-3} - 1$.

We now show that Tr_h is globally forcing for $h \geq 4$. To see this, suppose that $N_T^*(h) \leq (2^{-\binom{h}{2}} + o(1))n^h$. From (3) this gives $N_T^*(3) \leq (1 + o(1))n^3/8$. As $h \geq 4$, the function x^{h-2} is strictly convex and by the application of Jensen in (3) we find $|N(e)| = (1 \pm o(1))\frac{n}{4}$ for almost all $e \in E(T)$. However, by a result of Chung and Graham (see property P_4 and Theorem 1 in [6]) this property implies that T is quasirandom. Thus Tr_h is globally forcing for $h \geq 4$.

To complete the proof of the proposition, it remains to show that any non-transitive h-vertex tournament H with $h \geq 7$ is not globally forcing. To see this, let $V(H) = \{u_1, \ldots, u_h\}$. For each $n \in \mathbb{N}$ construct an n-vertex tournament T_1 as follows. Let $V(T_1) = V = \{v_1, \ldots, v_n\}$ denote a set of order n and let $V = \cup_{i \in [h]} V_i$ denote a partition with $|V_i| = \lfloor n/h \rfloor$ or $\lceil n/h \rceil$ for all $i \in [h]$. For each edge $\overrightarrow{u_i u_{i'}}$ of H, orient all edges from V_i to $V_{i'}$ in T_1. Orient the remaining edges of T_1 arbitrarily. As each map $\phi : V(H) \to V(T_1)$ with $\phi(u_i) \in V_i$ for all $i \in [h]$ is a labelled embedding, we see that

$$N_{T_1}^*(H) \geq \left(h^{-h} - o(1)\right)n^h. \tag{4}$$

Now, starting from T_1, construct a sequence of tournaments $T_1, \ldots, T_n$ on vertex set $V(T_1)$, where for $i \geq 1$ each T_{i+1} is obtained from T_i by reorienting all edges to go from v_i to $\{v_{i+1}, \ldots, v_n\}$. Using that $h^{-h} > 2^{-\binom{h}{2}}$ for $h \geq 7$, by (4) there is $\delta > 0$ such that $N_{T_1}^*(H) > (2^{-\binom{h}{2}} + \delta)n^h$ for large n. On the other hand $T_n = Tr_n$ and as H is not transitive we have $N_{T_n}^*(H) = 0$. Since $N_{T_i}^*(H) = N_{T_{i+1}}^*(H) \pm hn^{h-1}$, by an intermediate value property $N_{T_i}^*(H) = (2^{-\binom{h}{2}} + o(1))n^h$ for some $i \in [n]$. However, T_i is not quasirandom. Indeed as $(2^{-\binom{h}{2}} + o(1))n^h = N_{T_i}^*(H) \geq (2^{-\binom{h}{2}} + \delta)n^h - ihn^{h-1}$ we have $i \geq \delta n/2h$. As $|d_{T_i}^+(v_j) - d_{T_i}^-(v_j)| = n - 2j + 1 \geq n/2$ for all $j \in [\min(i-1, n/4)]$, this contradicts the quasirandom property $\mathcal{P}_5$ from [6] which requires that

$$\sum_{v \in V(T_i)} |d_{T_i}^+(v) - d_{T_i}^-(v)| = o(n^2).$$

Thus T_i is not quasirandom and so H cannot be globally forcing. $\qquad\square$

Remark: Note that our argument shows that in order for H to be globally forcing $N_T^*(H)$ must be asymptotically maximised when T is the random tournament. In other words if we can find a sequence of n-vertex tournaments T_n such that $N_{T_n}^*(H) \geq (2^{-\binom{h}{2}} + \delta)n^h$ our argument above (switching edges incident to one vertex at a time) shows that H is not globally forcing. In fact it is enough to find a single such T_m for some $m \geq h$ since if we let T_n' be the blow up of T_m with parts of size $\frac{n}{m}$ then $N_{T_n'}^*(H) \geq (2^{-\binom{h}{2}} + \delta)m^h \left(\frac{n}{m}\right)^h = (2^{-\binom{h}{2}} + \delta)n^h$ so T_n' provides us with our sequence.

3   Local forcing and the counting polynomial

To study locally forcing tournaments we introduce the following polynomial.

Definition. *The* counting polynomial *of an h-vertex tournament H is given by*

$$p_H(x) := \mathbb{E}_{\sigma \sim S_h}\left(x^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|}\right) = \frac{1}{h!}\sum_{\sigma \in S_h}\left(x^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|}\right).$$

The following lemma collects a number of useful basic facts concerning $p_H(x)$.

Lemma 3.1. *Given an h-vertex tournament H, the following hold:*
  (i) *For $\alpha \in [0,1]$ we have $\mathbb{E}_{T \sim \mathcal{T}(n,\alpha)}\left(N_T^*(H)\right) = (n)_h \times p_H(\alpha)$.*
  (ii) *$p_H(1/2) = 2^{-\binom{h}{2}}$ and $p_H(1) = 0$ if H is not transitive.*
  (iii) *For all $\alpha \in [0, 1/2]$ we have $p_H(1/2 + \alpha) = p_H(1/2 - \alpha)$.*

Proof. Let $V(H) = \{u_1, \ldots, u_h\}$. A set $\{i_1, \ldots, i_h\} \subset [n] = V(T)$ with $i_1 < i_2 < \ldots < i_h$ forms a σ-copy of H if $\overrightarrow{i_j i_k} \in E(T)$ if and only if $\overrightarrow{u_{\sigma(j)} u_{\sigma(k)}} \in E(H)$. For $T \sim \mathcal{T}(n, \alpha)$, the probability that a fixed h-set forms a σ-copy of H is $\alpha^{|F_{\sigma,H}|}(1 - \alpha)^{|B_{\sigma,H}|}$. Letting $N_T^*(H, \sigma)$ denote the number of σ-copies of H in T gives

$$\mathbb{E}_{T \sim \mathcal{T}(n,\alpha)}(N_T^*(H, \sigma)) = \binom{n}{h}\alpha^{|F_{\sigma,H}|}(1 - \alpha)^{|B_{\sigma,H}|}.$$

As $N_T^*(H) = \sum_{\sigma \in S_h} N_T^*(H, \sigma)$ we have

$$\mathbb{E}_{T \sim \mathcal{T}(n,\alpha)}\big(N_T^*(H)\big) = \sum_{\sigma \in S_h} \mathbb{E}_{T \sim \mathcal{T}(n,\alpha)}\big(N_T^*(H, \sigma)\big) = (n)_h \times p_H(\alpha).$$

This gives *(i)*.

To see *(ii)*, note that $p_H(1/2) = 2^{-\binom{h}{2}}$ from the definition of $p_H(x)$ since $|F_{\sigma,H}| + |B_{\sigma,H}| = \binom{h}{2}$ for every $\sigma \in S_h$. Also note that if T is non-transitive then $B_{\sigma,H} > 0$ for all σ. So, each term of $p_H(x)$ is zero for $x = 1$ implying $p_H(1) = 0$.

Lastly, we show *(iii)*. For each $\sigma \in S_h$ let $\overline{\sigma}$ denote the 'reversal of σ', given by $\overline{\sigma}(i) = \sigma(h + 1 - i)$ for all $i \in [h]$. As this gives a bijection from S_h to itself and $F_{\sigma,H} = B_{\overline{\sigma},H}$, for all $\sigma \in S_h$, we have

$$p_H(x) = \frac{1}{2h!} \sum_{\sigma \in S_h} \left(x^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|} + x^{|F_{\overline{\sigma},H}|}(1-x)^{|B_{\overline{\sigma},H}|} \right)$$

$$= \frac{1}{2h!} \sum_{\sigma \in S_h} \left(x^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|} + x^{|B_{\sigma,H}|}(1-x)^{|F_{\sigma,H}|} \right)$$

As $x^a(1-x)^b + x^b(1-x)^a$ is symmetric around $1/2$ for any $a, b \in [0, 1/2]$ so is $p_H(x)$. $\qquad\square$

We are now ready to give the formal definition of $\mathcal{T}_{cliq}$-forcing which plays a central role in our arguments.

Definition. *A tournament H is said to be $\mathcal{T}_{cliq}$-forcing if $p_H(x) \neq 2^{-\binom{h}{2}}$ for $x \in [0,1] \setminus \{\frac{1}{2}\}$.*

Note that if H is not locally forcing then there is a tournament T which is not quasirandom but has the same local counts of H as the random tournament $\mathcal{T}(n, 1/2)$. We will see that not being $\mathcal{T}_{cliq}$-forcing implies that there exists an $\alpha \neq 1/2$ such that $\mathcal{T}(n, \alpha)$ has the same local counts of H as the random tournament $\mathcal{T}(n, 1/2)$. Indeed, in the next theorem we prove that any locally forcing tournament must be $\mathcal{T}_{cliq}$-forcing, which will allow us to pass any restrictions imposed on the tournament by being $\mathcal{T}_{cliq}$-forcing to locally forcing tournaments and in particular prove Theorem 1.3.

Theorem 3.2. *Every locally forcing tournament is $\mathcal{T}_{cliq}$-forcing.*

Proof. Suppose that H is not $\mathcal{T}_{cliq}$-forcing. Then by definition $p_H(\alpha) = 2^{-\binom{h}{2}}$ for some $\alpha \in [0, 1] \setminus \{1/2\}$. As $p_H(x)$ is symmetric about $1/2$ by Lemma 3.1 *(iii)* we can assume $\alpha > 1/2$. Now select $T \sim \mathcal{T}(n, \alpha)$. For a set $U \subset V(T)$ let X_U denote the random variable $X_U(T) = N_T^*(H; U)$. Noting that for $T \sim \mathcal{T}(n, \alpha)$ we have $T[U] \sim \mathcal{T}(|U|, \alpha)$, by Lemma 3.1 *(i)* we have

$$\mathbb{E}_{T \sim \mathcal{T}(n,\alpha)}(X_U) = (|U|)_h \times p_H(\alpha) = (|U|)_h \times 2^{-\binom{h}{2}}.$$

Since each edge belongs to at most $\binom{h}{2}n^{h-2}$ copies of H, X_U is sharply concentrated by Azuma's inequality (see Chapter 7 [4])

$$\mathbb{P}(|X_U - \mathbb{E}(X_U)| > \varepsilon n^h) \leq 2e^{\frac{-\varepsilon^2 n^{2h}}{2\binom{n}{2}\left(\binom{h}{2}n^{h-2}\right)^2}} = e^{-\Omega(n^2)}.$$

In particular, by a union bound with probability $1 - o(1)$ we have $X_U = 2^{-\binom{h}{2}}|U|^h + o(n^h)$ for every $U \subset V(T)$.

On the other hand the number of forward edges of T is distributed as $\text{Bin}(\binom{n}{2}, \alpha)$ so by standard Chernoff type estimates T has $\alpha\binom{n}{2} \pm o(n^2)$ forward edges with probability $1 - o(1)$ (see Appendix A of [4]). Combined with the previous paragraph, we have shown that there is an n-vertex tournament T which satisfies $\mathcal{P}_2^*(H)$ but not $\mathcal{P}_1$. Thus H cannot be locally forcing. $\qquad\square$

For the rest of this section it will be more convenient to work with a rescaled and recentered counting polynomial $q_H(x) := 2^{\binom{h}{2}} \times p_H\left(\frac{1+x}{2}\right)$. We now show several restrictions that the $\mathcal{T}_{cliq}$-forcing condition imposes on a tournament. The first one is that a $\mathcal{T}_{cliq}$-forcing tournament should be almost regular.

Lemma 3.3. *If H is a non-transitive h-vertex $\mathcal{T}_{cliq}$-forcing tournament then*

$$\sum_{u \in V(H)} \left(d_H^+(u) - d_H^-(u)\right)^2 \leq h(h-1). \tag{5}$$

Proof. We start by determining the coefficient of x^2 in q_H. We claim it is equal to

$$\frac{1}{6} \times \left(\sum_{u \in V(H)} \left(d_H^+(u) - d_H^-(u)\right)^2 - h(h-1) \right).$$

Given $e \in E(H)$ let I_e denote the function $I_e : S_h \to \{-1, 1\}$ with $I_e(\sigma) = 1$ if $e \in F_{\sigma, H}$ and $I_e(\sigma) = -1$ if $e \in B_{\sigma, H}$. Then we have

$$q_H(x) = \mathbb{E}_{\sigma \sim S_h}\left((1+x)^{|F_{\sigma,H}|}(1-x)^{|B_{\sigma,H}|}\right) = \mathbb{E}_{\sigma \sim S_h}\left(\prod_{e \in E(H)} (1 + I_e(\sigma)x) \right). \tag{6}$$

This implies that the coefficient of x^2 in $q_H(x)$ is

$$\sum_{\{e,e'\} \in \binom{E(H)}{2}} \left(\mathbb{E}_{\sigma \sim S_h} I_e(\sigma) I_{e'}(\sigma) \right). \tag{7}$$

The contribution of each pair $\{e, e'\} \in \binom{E(H)}{2}$ to this sum is determined by how these edges meet. If $e \cap e' = \emptyset$ then the contribution to the sum is 0. If e and e' are both out-edges of a single vertex, say $e = \overrightarrow{u_1 u_2}$ and $e' = \overrightarrow{u_1 u_3}$, the contribution to the sum is $1/3$. This is also true if e and e' are both in-edges of a single vertex. Lastly, if e is an out-edge of a vertex v and e' is an in-edge then the contribution to the sum is $-1/3$. By counting the non-zero contributions of pairs $\{e, e'\}$ according to the vertex of $V(H)$ which intersect in, by (7) the coefficient of x^2 in $q_H(x)$ is

$$\sum_{u \in V(H)} \left(\frac{1}{3}\binom{d_H^+(u)}{2} + \frac{1}{3}\binom{d_H^-(u)}{2} - \frac{1}{3}d_H^+(u)d_H^-(u) \right) = \sum_{u \in V(H)} \frac{(d_H^+(u) - d_H^-(u))^2}{6} - \frac{h(h-1)}{6}.$$

as claimed.

Finally, if (5) fails, $q_H(\varepsilon) = 1 + a\varepsilon^2 + O(\varepsilon^4)$ with $a > 0$, where we used Lemma 3.1 *(ii)* to get $q_H(0) = 1$ and *(iii)* to get that q_H is even. This implies that $q_H(\varepsilon) > 1$ for sufficiently small $\varepsilon > 0$. As H is not transitive, Lemma 3.1 *(ii)* gives $q_H(1) = 0$, and so by the intermediate value theorem $q_H(\varepsilon) = 1$ for some $\varepsilon \in (0, 1)$. But this gives $p_H\left(\frac{1+\varepsilon}{2}\right) = 2^{-\binom{h}{2}}$ and so H is not $\mathcal{T}_{cliq}$-forcing. $\qquad\square$

The following lemma shows that, in addition to being almost regular, any $\mathcal{T}_{cliq}$-forcing H must be strongly quasirandom.

Lemma 3.4. *If H is a non-transitive h-vertex $\mathcal{T}_{cliq}$-forcing tournament then for every ordering σ of H*

$$|F_{\sigma,H}| = \frac{1}{2}\binom{h}{2} \pm h^{3/2}\sqrt{\log h}.$$

Proof. Assume that there is an ordering σ such that $f = |F_{\sigma,H}| \geq \frac{1}{2}\binom{h}{2} + h^{3/2}\sqrt{\log h}$. If we let $b = |B_{\sigma,H}|$ since $f + b = \binom{h}{2}$ we get $b \leq \frac{1}{2}\binom{h}{2} - h^{3/2}\sqrt{\log h}$ and $f - b \geq 2h^{3/2}\sqrt{\log h}$.

We will show that a single term of $q_H(x)$, $(1+x)^f(1-x)^b > h!$ if we choose $x = \frac{3}{4}h^{-1/2}\sqrt{\log h}$. This shows that $q_H(x) > 1$ and we can complete the argument as in the previous lemma. To show the above inequality note that

$$(1+x)^f(1-x)^b > e^{(x-x^2)f + (-x-x^2)b}$$
$$> e^{x(f-b) - x^2h^2/2}$$
$$> e^{h \log h} = h^h > h!$$

where in the first inequality we used $1 + t > e^{t-t^2}$ for $|t| \leq 1/\sqrt{3}$. $\qquad\square$

8

By a result of Spencer [34] it is known that for any tournament H there is an ordering of its vertices σ for which $|F_{\sigma,H}| = \frac{1}{2}\binom{h}{2} + \Omega(h^{3/2})$. So the above result is best possible up to the $\sqrt{\log h}$ factor. Fernandez de la Vega [16] showed that for the random tournament $H \sim \mathcal{T}(h, 1/2)$ w.h.p. there is no ordering with $|F_{\sigma,H}| > \frac{1}{2}\binom{h}{2} + 2h^{3/2}$. So it seems quite likely that the $\sqrt{\log h}$ term in the above lemma is not necessary. In fact for a different definition of quasirandomness we do obtain a result which is best possible up to a constant factor.

Lemma 3.5. *If H is a non-transitive h-vertex $\mathcal{T}_{cliq}$-forcing tournament then for any disjoint subsets $U, W \subseteq V(H)$ we must have $e(U, W) \leq \frac{1}{2}|U||W| + 2h^{3/2}$.*

Proof. Assume towards a contradiction that there are disjoint sets $U, W \subseteq V(H)$ for which $e(U, W) > \frac{1}{2}|U||W| + 2h^{3/2}$. Using these sets we will find many orderings σ such that $|F_{\sigma,H}| \geq \frac{1}{2}\binom{h}{2} + 2h^{3/2}$. Let $S = V(H) \setminus (U \cup W)$. We will always place all vertices of U before all vertices of W in the orderings σ. Also if $e(S, U \cup W) \geq e(U \cup W, S)$ we place all the vertices of S before all vertices of $U \cup W$ and behind otherwise. Within the sets U, W, S we take the orderings which have more forward edges than backward edges. Note that for any subset of H exactly one out of each of its orderings and its reverse has at least half of the edges going forwards. Therefore, we get at least $\frac{|U|!}{2} \cdot \frac{|W|!}{2} \cdot \frac{|S|!}{2} = \frac{h!}{8\binom{h}{|U|,|W|,|S|}} > h!/(8 \cdot 3^h)$ such orderings (where we used the trinomial expansion). Note that each such ordering σ of H has $|F_{\sigma,H}| \geq \frac{1}{2}\binom{h}{2} + 2h^{3/2}$ since inside each set and between each pair of sets there are at least half of them going forwards and there is a gain of at least $2h^{3/2}$ between U and W.

Assume now σ is an ordering of H with $f = |F_{\sigma,H}| \geq \frac{1}{2}\binom{h}{2} + 2h^{3/2}$. If we let $b = |B_{\sigma,H}|$ similarly as in the previous lemma we obtain that if we choose $x = h^{-1/2}$ then

$$(1+x)^f(1-x)^b > e^{(x-x^2)f + (-x-x^2)b}$$
$$> e^{x(f-b) - x^2 h^2/2}$$
$$> e^{7h/2} > 8 \cdot 3^h.$$

From the previous two paragraphs we get $q_H(x) > 1$ and complete the proof as in Lemma 3.3. $\square$

Spencer showed in [34] that for some constant $c > 0$ in any h-vertex tournament there are disjoint sets of vertices U, W such that $e(U, W) \geq \frac{1}{2}|U||W| + \Omega(h^{3/2})$. So the above lemma is best possible up to the constant factor.

Proof of Theorem 1.3. Combining Theorem 3.2 with Lemmas 3.3, 3.4 and 3.5 proves Theorem 1.3. $\square$

Lemmas 3.4 and 3.5 show that in order for H to be $\mathcal{T}_{cliq}$-forcing it needs to be quasirandom in 2 different ways introduced by Chung and Graham in [6]. But it shows even more, namely that the error term should not only be qualitatively small (as in the definition of quasirandom properties in the introduction) but should in fact be close to their extremal value.

Since the previous two lemmas require H to be quasirandom and are both satisfied for the random tournament $\mathcal{T}(h, 1/2)$ w.h.p. a natural guess would be that it should be $\mathcal{T}_{cliq}$-forcing w.h.p.. This turns out to be false, due to Lemma 3.3.

Proof of Corollary 1.4. Let $H \sim \mathcal{T}(h, 1/2)$. We show that with positive probability $\sum_{v \in V(H)} (d_H^+(v) - d_H^-(v))^2 > h(h-1)$ which will show the result by Lemma 3.3 through Theorem 3.2.

Let us set $V(H) = [h]$ and define $I_{ij} = 1$ if $\overrightarrow{ij} \in E(H)$ and $I_{ij} = -1$ if $\overrightarrow{ji} \in E(H)$, so $\mathbb{P}(I_{ij} = 1) = \mathbb{P}(I_{ij} = -1) = 1/2$. Note that $I_{ij} = -I_{ji}$ and that otherwise indicators I_{ij} are mutually independent. We also set $I_{ii} = 0$ for convenience. We have

$$\sum_{i \in [h]} (d_H^+(i) - d_H^-(i))^2 = \sum_{i \in [h]} \left(\sum_{j \in [h]} I_{ij}\right)^2 = h(h-1) + 2\sum_{i,j,k \in [h], j \neq k} I_{ij}I_{ik}.$$

Let $X = \sum_{i,j,k \in [h], j \neq k} I_{ij}I_{ik}$. Note that $\mathbb{E}(X) = \sum_{i,j,k \in [h], j \neq k} \mathbb{E}(I_{ij})\mathbb{E}(I_{ik}) = 0$ since each I_{ij} is independent of I_{ik} when $j \neq k$ and $\mathbb{E}(I_{ij}) = 0$. Our goal is to show $\mathbb{P}(X > 0) > c$ for some $c > 0$ independent of h. To do this we need to determine some higher moments of X. Note first that

$$\mathbb{E}(X^2) = \sum_{i,j,k \in [h], j \neq k} \sum_{i',j',k' \in [h], j' \neq k'} \mathbb{E}(I_{ij} I_{ik} I_{i'j'} I_{i'k'}) = h \binom{h-1}{2}$$

where we used the fact that unless $i = i', j = j', k = k'$ at least one of the sets $\{i,j\}, \{i,k\}, \{i',j'\}$, $\{i',k'\}$ is distinct from the others, so its indicator is independent of the others and its contribution vanishes.

It is not hard to estimate $\mathbb{E}X^4$ directly but it requires some case analysis. Instead note that if we replace every occurrence of I_{ji} with $i < j$ with $-I_{ij}$ we get a degree 2 polynomial whose variables are independent indicators. Thus, by a special case of Bonami-Beckner's hypercontractive inequality (see [29] for a simple proof) we have that $\mathbb{E}(X^4) \leq (9\mathbb{E}(X^2))^2 = 81[\mathbb{E}(X^2)]^2$.

Now a simple lemma (see [2]) says that if a random variable Y has expectation 0, $\mathbb{E}Y^2 > 0$ and $\mathbb{E}Y^4/(\mathbb{E}Y^2)^2 \leq b$ then $\mathbb{P}(Y > 0) > 1/(2^{4/3}b)$. So in our case $b = 81$ and $\mathbb{P}(X > 0) \geq 1/205$. $\qquad \square$

4 Finding $\mathcal{T}_{cliq}$-forcing tournaments

In the previous section we saw that in order for a tournament to be $\mathcal{T}_{cliq}$-forcing it needs to be strongly quasirandom. We have also seen that the random tournament is not regular enough to be $\mathcal{T}_{cliq}$-forcing w.h.p. It is natural to try to avoid this obstruction by trying next the random regular tournaments. However, the probability space of random regular tournaments is not at all easy to work with so instead we consider a modification of a different probability space on regular tournaments first introduced by Adler, Alon and Ross in [1] to study the maximum number of Hamilton paths in tournaments. Using it we show that there are many $\mathcal{T}_{cliq}$-forcing h-vertex tournaments for any large enough h.

To describe the construction, suppose that triangles $\Delta_1, \ldots, \Delta_L$ and edges $e_1, \ldots, e_F$ form a partition $\mathcal{P}$ of the edge set of the complete graph K_h. We now generate a random tournament on the vertex set $[h]$ as follows. Orient the edges of each Δ_i as a cyclic triangle, each orientation appearing with probability $1/2$ and then orient each e_j uniformly at random as well, so that all the triangles and edges are oriented independently. Write $\mathcal{D}_{\mathcal{P}}$ for the resulting distribution on the set of h-vertex tournaments.

Lemma 4.1. *There exists $c > 0$ such that for $h \geq h_0$ and $F \leq ch^2$, the random tournament $H \sim \mathcal{D}_{\mathcal{P}}$ is $\mathcal{T}_{cliq}$-forcing w.h.p..*

Proof. Note that $3L + F = \binom{h}{2}$ so by taking c small enough we may assume $L \geq 10F$, implying $h^2/8 \leq L \leq h^2/6$.

To prove the lemma, we again work with $q_H(x)$, showing that with positive probability $q_H(x) \neq 1$ for all $x \in [-1,1] \setminus \{0\}$. Note that in any ordering σ of the vertices of H, the triangle corresponding to Δ_i either has two forward edges or two backward edges. Let $J_{i,\sigma}(H) = +1$ in the first case and $J_{i,\sigma}(H) = -1$ in the second case. Similarly let $J_{L+j,\sigma}(H) = \pm 1$ depending on whether edge e_j is forwards or backwards. We have

$$q_H(x) = \mathbb{E}_{\sigma \sim S_h} \left((1+x)^{|F_{\sigma,H}|} (1-x)^{|B_{\sigma,H}|} \right)$$

$$= \mathbb{E}_{\sigma \sim S_h} \left(\prod_{i \in [L]} (1+x)(1-x)(1 + J_{i,\sigma}(H)x) \prod_{i \in [L+1, L+F]} (1 + J_{i,\sigma}(H)x) \right)$$

$$= (1-x^2)^L \times \mathbb{E}_{\sigma \sim S_h} \left(\prod_{i \in [L+F]} (1 + J_{i,\sigma}(H)x) \right)$$

$$= (1-x^2)^L \times s_H(x).$$

As $q_H(x)$ is an even polynomial, so is the (random) polynomial $s_H(x)$. Let $s_H(x) = \sum_{\ell \in [(L+F)/2]} c_{2\ell} x^{2\ell}$, where $\{c_{2\ell}\}_{\ell \in [(L+F)/2]}$ are random variables depending on H.

10

To complete the lemma it will suffice to prove that with high probability $c_{2\ell} \leq \binom{L}{\ell}$ for all $\ell \in [(L+F)/2]$. Indeed, then

$$s_H(x) = \sum_{\ell \in [(L+F)/2]} c_{2\ell} x^{2\ell} \leq \sum_{\ell \in [(L+F)/2]} \binom{L}{\ell} x^{2\ell} \leq \sum_{\ell \in [L]} \binom{L}{\ell} x^{2\ell} = (1 + x^2)^L,$$

which gives $q_H(x) = (1 - x^2)^L \times s_H(x) \leq (1 - x^2)^L \times (1 + x^2)^L = (1 - x^4)^L < 1$ for $x \in [-1, 1] \setminus \{0\}$, i.e. H is $\mathcal{T}_{cliq}$-forcing. To obtain the required bound on $c_{2\ell}$, note that

$$c_{2\ell} = \sum_{A \in \binom{[L+F]}{2\ell}} \mathbb{E}_{\sigma \sim S_h} \Big(\prod_{i \in A} J_{i,\sigma}(H) \Big). \tag{8}$$

Notice that for any $A \in \binom{[L+F]}{2\ell}$ if there is $i \in A$ such that the triangle or edge corresponding to i is vertex disjoint from all other objects indexed by A then $\mathbb{E}_{\sigma \sim S_h} \Big(\prod_{i \in A} J_{i,\sigma}(H) \Big) = 0$. Indeed, in this case we can cancel the contribution of a permutation σ to the expectation with the permutation $\widetilde{\sigma}$ obtained from σ by reversing the orientation of Δ_i or e_{i-L}. Let $\mathcal{A} \subseteq \binom{[L+F]}{2\ell}$ denote the sets not of this form, i.e. if $A \in \mathcal{A}$ then every object indexed by A shares a vertex with another object indexed by A. We have shown that

$$c_{2\ell} = \sum_{A \in \mathcal{A}} \mathbb{E}_{\sigma \sim S_h} \Big(\prod_{i \in A} J_{i,\sigma}(H) \Big). \tag{9}$$

Note further that any $A \in \mathcal{A}$ must have a subset $A' \subset A$ with $|A'| = \ell$ such that any object indexed by A intersects an object indexed by A'. We can choose these A' in $\binom{L+F}{\ell}$ many ways and they span at most 3ℓ vertices. This leaves us with at most $3\ell h$ options for each of the remaining objects. In particular, we have shown that

$$|\mathcal{A}| \leq \binom{L+F}{\ell} \cdot \binom{3\ell h}{\ell}. \tag{10}$$

Turning back to upper bounding $c_{2\ell}$ note that (9) gives us

$$\mathbb{E}_{H \sim \mathcal{D}_{\mathcal{P}}} \big(c_{2\ell} \big)^2 = \mathbb{E}_{\sigma, \sigma' \sim S_h} \mathbb{E}_{H \sim \mathcal{D}_{\mathcal{P}}} \Big(\sum_{A, B \in \mathcal{A}} \prod_{i \in A} J_{i,\sigma}(H) \prod_{j \in B} J_{j,\sigma'}(H) \Big). \tag{11}$$

We claim that the contribution coming from distinct $A, B \in \mathcal{A}$ vanishes. Indeed let $A \neq B$ and let us fix an $i' \in A \setminus B$ (if $A \subset B$ we can use an analogous argument which takes $i' \in B \setminus A$ instead). For any fixed orderings σ, σ' we have that $J_{i',\sigma}(H)$ is independent from all $J_{i,\sigma}(H)$ for $i \in A \setminus \{i'\}$ and all $J_{j,\sigma'}(H)$ for $j \in B$. This implies

$$\mathbb{E}_{H \sim \mathcal{D}_{\mathcal{P}}} \Big(\prod_{i \in A} J_{i,\sigma}(H) \prod_{j \in B} J_{j,\sigma'}(H) \Big) = \mathbb{E}_{H \sim \mathcal{D}_{\mathcal{P}}} (J_{i',\sigma}) \cdot \mathbb{E}_{H \sim \mathcal{D}_{\mathcal{P}}} \Big(\prod_{i \in A \setminus \{i'\}} J_{i,\sigma}(H) \prod_{j \in B} J_{j,\sigma'}(H) \Big) = 0,$$

since $\mathbb{E}_{H \sim \mathcal{D}_{\mathcal{P}}} (J_{i',\sigma}) = 0$. Hence,

$$\mathbb{E}_{\sigma, \sigma' \sim S_h} \sum_{\substack{A, B \in \mathcal{A} \\ A \neq B}} \mathbb{E}_{H \sim \mathcal{D}_{\mathcal{P}}} \Big(\prod_{i \in A} J_{i,\sigma}(H) \prod_{j \in B} J_{j,\sigma'}(H) \Big) = 0.$$

Thus, $\mathbb{E}_{H \sim \mathcal{D}_{\mathcal{P}}} (c_{2\ell})^2 \leq |\mathcal{A}|$ by (11). By Markov's inequality the probability that $c_{2\ell} > \binom{L}{\ell}$ is at most $|\mathcal{A}|/\binom{L}{\ell}^2$. For $\ell \geq 2 \log L$ we get

$$|\mathcal{A}|/\binom{L}{\ell}^2 < \binom{L+F}{2\ell}/\binom{L}{\ell}^2 < \prod_{i=0}^{\ell-1} \Big(\frac{L+F-2i}{L-i} \Big)^2 \binom{2\ell}{\ell}^{-1} \leq \Big(\frac{L+F}{L} \Big)^{2\ell}/\binom{2\ell}{\ell} \leq 2^{-\ell} \leq \frac{1}{L^2} \tag{12}$$

where we used $L \geq F$ and in the second to last inequality we took c small enough and $\ell \geq \ell_0$ since $h \geq h_0$. When $\ell < 2 \log L$ using (10)

$$|\mathcal{A}|/\binom{L}{\ell}^2 \leq \binom{L+F}{\ell} \cdot \binom{3\ell h}{\ell}/\binom{L}{\ell}^2 \leq \Big(\frac{3\ell h(L+F)}{(L-\ell)^2} \Big)^\ell \leq \frac{16 \log L}{\sqrt{L}}.$$

11

Here we used $L \geq 10F$ and $h^2/8 \leq L \leq h^2/6$ and assumed h is large enough for $\frac{16 \log L}{\sqrt{L}} \leq 1$. Summing over all ℓ we have shown that $c_{2\ell} \leq \binom{L}{\ell}$ for all $\ell \in [L]$ with probability at least $1 - 2 \log L \cdot \frac{16 \log L}{\sqrt{L}} + \frac{L - 2 \log L}{L^2} = 1 - o(1)$. $\qquad\square$

Remark: It is well-known (see [24]) that for any $h \in \mathbb{N}$ with $h \equiv 1$ or $3 \mod 6$ the complete graph K_h on vertex set $[h]$ admits a Steiner triple decomposition. That is, there is a partition $\mathcal{P}$ of K_h as above for which $F = 0$. Note that in this case $\mathcal{D}_\mathcal{P}$ is always a regular tournament. So there is an infinite family of $\mathcal{T}_{cliq}$-forcing regular tournaments.

5 $\mathcal{T}_{bip}$-forcing tournaments

In the previous section we have shown that there are many $\mathcal{T}_{cliq}$-forcing tournaments. Our original goal however was to study locally forcing tournaments and Theorem 3.2 only says that locally forcing tournaments are necessarily $\mathcal{T}_{cliq}$-forcing. While we believe this to be a sufficient condition, our argument requires a weak additional assumption, which we will call $\mathcal{T}_{bip}$-forcing. We show that being $\mathcal{T}_{bip}$-forcing and $\mathcal{T}_{cliq}$-forcing is equivalent to being locally forcing, which allows us to show that many $\mathcal{T}_{cliq}$-forcing tournaments found in the previous section are in fact locally forcing. We now give the formal definition of $\mathcal{T}_{bip}$-forcing.

Definition. *We define the a-th order degree counting polynomial of an h-vertex tournament by:*

$$p_{H,a}(x) := \binom{h}{a}^{-1} 2^{-\binom{a}{2} - \binom{h-a}{2}} \sum_{A \in \binom{V(H)}{a}} x^{e(A, V \setminus A)} (1-x)^{e(V \setminus A, A)}.$$

Definition. *An h-vertex tournament is $\mathcal{T}_{bip}$-forcing if there is no $\alpha \in (1/2, 1]$ such that $p_{H,a}(\alpha) = 2^{-\binom{h}{2}}$ for all $1 \leq a \leq h - 1$ simultaneously.*

We have seen that if H fails to be $\mathcal{T}_{cliq}$-forcing then there is an $\alpha \neq 1/2$ such that $\mathcal{T}(n, \alpha)$ has the same local count of H as the random tournament $\mathcal{T}(n, 1/2)$. In the following lemma we will see that H not being $\mathcal{T}_{bip}$-forcing means that there is an $\alpha \neq 1/2$ such that $\mathcal{T}(n, n, \alpha)$ does have the same count of H as $\mathcal{T}(n, n, 1/2) = \mathcal{T}(2n, 1/2)$. In fact H not being $\mathcal{T}_{bip}$-forcing is a seemingly much stronger restriction on H in the sense that it means that there is an $\alpha > 1/2$ such that for *all* $1 \leq a \leq h - 1$, $\mathcal{T}(n, n, \alpha)$ has the correct count of H with a vertices embedded to the left and $h - a$ to the right side for all a simultaneously.

Lemma 5.1. *Every locally forcing tournament is $\mathcal{T}_{bip}$-forcing.*

Proof. Suppose towards a contradiction that there is an $\alpha \in (1/2, 1]$ such that $p_{H,a}(\alpha) = 2^{-\binom{h}{2}}$ for all $1 \leq a \leq h - 1$. Now select $T \sim \mathcal{T}(n, n, \alpha)$ with bipartition (L, R). For a set $U \subset L, W \subset R$ let $X_{U,W}$ denote the random variable $X_{U,W}(T) = N_T^*(H; U \cup W)$. The probability that a fixed subset of size a of U and a fixed subset of size $h - a$ of W span an embedding of H with $A \subset V(H)$ being embedded to the left is $2^{-\binom{a}{2} - \binom{h-a}{2}} \alpha^{e(A, V \setminus A)} (1 - \alpha)^{e(V \setminus A, A)}$. Summing over all possibilities we obtain

$$\mathbb{E}_{T \sim \mathcal{T}(n,n,\alpha)}(X_{U,W}) = \sum_{a=0}^{h} (|U|)_a (|W|)_{h-a} \times \sum_{A \in \binom{V(H)}{a}} 2^{-\binom{a}{2} - \binom{h-a}{2}} \alpha^{e(A, V \setminus A)} (1 - \alpha)^{e(V \setminus A, A)}$$

$$= (|W|)_h 2^{-\binom{h}{2}} + \sum_{a=1}^{h-1} (|U|)_a (|W|)_{h-a} \times \binom{h}{a} p_{H,a}(\alpha) + (|U|)_h 2^{-\binom{h}{2}}$$

$$= \sum_{a=0}^{h} \binom{h}{a} (|U|)_a (|W|)_{h-a} \times 2^{-\binom{h}{2}}$$

$$= (|U| + |W|)_h \times 2^{-\binom{h}{2}},$$

Where in the last equality we used the identity $\sum_{a=0}^{h}\binom{b}{a}\binom{c}{h-a}=\binom{b+c}{h}$ holding for any b,c.

The rest of the proof proceeds in the same way as the proof of Theorem 3.2 except that the number of forwards edges of $T(n,n,\alpha)$ is $(1/4+\alpha/2)\binom{2n}{2}\pm o(n^2)$ with probability $1-o(1)$. $\square$

This lemma together with Theorem 3.2 shows the 'only if' part of Theorem 1.6. We postpone the proof of the 'if' part to the next section, showing first that there are many tournaments which are $\mathcal{T}_{bip}$-forcing, in addition to being $\mathcal{T}_{cliq}$-forcing.

As already mentioned, we believe being $\mathcal{T}_{bip}$-forcing is a very weak additional condition which might be already implied by being $\mathcal{T}_{cliq}$-forcing. It is even possible that all tournaments are $\mathcal{T}_{bip}$-forcing. We now give certain simple conditions that make a tournament $\mathcal{T}_{bip}$-forcing. We say that an h-vertex tournament H is *nearly regular* if $|d_H^+(v)-d_H^-(v)|<\sqrt{h}/2$ for all $v\in V(H)$.

Lemma 5.2. *Any nearly regular tournament is $\mathcal{T}_{bip}$-forcing.*

Proof. We are going to show that $p_{H,1}(x)+p_{H,h-1}(x)<2\cdot 2^{-\binom{h}{2}}$ for all $x\in(1/2,1]$. It will be easier to work with the following polynomials

$$q_{H,a}(x)=2^{\binom{h}{2}}\cdot p_{H,a}\left(\frac{1+x}{2}\right)=\binom{h}{a}^{-1}\sum_{A\in\binom{V(H)}{a}}(1+x)^{e(A,V\setminus A)}(1-x)^{e(V\setminus A,A)}.$$

So that

$$q_{H,1}(x)+q_{H,h-1}(x)=\frac{1}{h}\sum_{v\in V(H)}(1+x)^{d^+(v)}(1-x)^{d^-(v)}+(1+x)^{d^-(v)}(1-x)^{d^+(v)}$$

$$=\frac{1}{h}\sum_{v\in V(H)}(1-x^2)^{\min(d^+(v),d^-(v))}\left((1+x)^{|d^+(v)-d^-(v)|}+(1-x)^{|d^-(v)-d^+(v)|}\right)$$

$$\le(1-x^2)^{h/6}\left((1+x)^{\lfloor\sqrt{h}/2\rfloor}+(1-x)^{\lfloor\sqrt{h}/2\rfloor}\right)$$

Where we used that $\min(d^+(v),d^-(v))\ge(h-1)/2-\sqrt{h}/2\ge h/6$, when $h\ge 5$ and that if $h<5$ near regularity implies regularity so $(h-1)/2\ge h/6$. Note that since $\binom{\sqrt{h}/2}{2\ell}\le\frac{(\sqrt{h}/2)^{2\ell}}{(2\ell)!}\le\left(\frac{h}{6\ell}\right)^{\ell}\le\binom{h/6}{\ell}$

$$(1+x)^{\lfloor\sqrt{h}/2\rfloor}+(1-x)^{\lfloor\sqrt{h}/2\rfloor}\le 2\sum_{\ell=0}^{\sqrt{h}/4}\binom{\sqrt{h}/2}{2\ell}x^{2\ell}\le 2\sum_{\ell=0}^{\sqrt{h}/4}\binom{h/6}{\ell}x^{2\ell}\le 2(1+x^2)^{h/6}.$$

Combining the above two inequalities we obtain $q_{H,1}(x)+q_{H,h-1}(x)\le 2(1-x^4)^{h/6}<2$ as desired. $\square$

Note that we know by Lemma 3.3 that any $\mathcal{T}_{cliq}$-forcing tournament must be almost regular but the restriction of the above lemma is slightly stronger. On the other hand our argument only uses a much weaker property than the one given to us by the definition of being $\mathcal{T}_{bip}$-forcing.

Remark: It is not hard to adapt our proof of Lemma 4.1 to show that the random tournament $\mathcal{D}_{\mathcal{P}}$ is $\mathcal{T}_{bip}$-forcing provided $\mathcal{P}$ consists of at most ch^2 edges (and the remaining objects are triangles) for sufficiently small c. So in some sense all our examples from the previous section are in fact locally forcing.

To conclude the section we deduce Theorem 1.5, assuming the 'if' statement from Theorem 1.6 (which will be proven in the next section).

Proof of Theorem 1.5. Let $\mathcal{P}$ be a partition of K_h consisting of triangles and edges with every vertex incident to less than $\sqrt{h}/2$ of the edges in $\mathcal{P}$. It is easy to see that such a partition exists for any large enough h (for example by a result of Gustavsson [21]).

By Lemma 4.1 $H\sim\mathcal{D}_{\mathcal{P}}$ is $\mathcal{T}_{cliq}$-forcing w.h.p.. Furthermore, any $H\sim\mathcal{D}_{\mathcal{P}}$ is nearly regular and so by Lemma 5.2 it is $\mathcal{T}_{bip}$-forcing. Putting these together, Theorem 1.6 shows that the tournament $H\sim\mathcal{D}_{\mathcal{P}}$ is w.h.p. locally forcing. $\square$

13

6 Proving local forcing

Before proceeding to the proof of the 'if' statement of Theorem 1.6 in subsection 6.2, we first recall some results on regularity lemmas for directed graphs in the next subsection.

6.1 Regularity and counting lemmas for directed graphs

A directed graph $D = (V, E)$ consists of a set V of vertices and a set of edges $E \subset V \times V$. Clearly any tournament is also a directed graph. Given disjoint sets $A, B \subset V$ we write $E(A, B)$ to denote the collection of edges $(a, b) \in E \cap (A \times B)$ and $e(A, B) = |E(A, B)|$. We will write $d(A, B)$ to denote the *density* of the pair (A, B), given by

$$d(A, B) = \frac{|E(A, B)|}{|A||B|}.$$

Note that if T is a tournament then $d(A, B) = 1 - d(B, A)$. Given disjoint sets $X, Y \subset V$ we say that (X, Y) is an ε-regular pair if all $X' \subset X$ and $Y' \subset Y$ with $|X'| \geq \varepsilon|X|$ and $|Y'| \geq \varepsilon|Y|$ satisfy

$$|d(X', Y') - d(X, Y)| \leq \varepsilon \qquad \text{and} \qquad |d(Y', X') - d(Y, X)| \leq \varepsilon.$$

A partition of $V = \{V_0, V_1, \ldots, V_K\}$ is said to be an ε-regular partition of D if:
   (i) $|V_0| \leq \varepsilon|V|$,
   (ii) $|V_1| = \cdots = |V_K|$,
   (iii) all but at most $\varepsilon\binom{K}{2}$ of the pairs (V_i, V_j) with $1 \leq i < j \leq K$ are ε-regular.
The set V_0 is called the exceptional set and the sets $V_1, \ldots, V_K$ are clusters. This partition is said to refine a partition $V = U_1 \cup \cdots \cup U_L$ if for all $k \in [K]$ we have $V_k \subset U_l$ for some $l \in [L]$.

The directed regularity lemma of Alon and Shapira from [3], which extends Szemerédi's graph regularity lemma [35], states the following:

Theorem 6.1. *Given $m, L \in \mathbb{N}$ and $\varepsilon > 0$, there is $M = M(m, L, \varepsilon)$ with the following property. Given a directed graph $D = (V, E)$ with $|V| \geq M$ and a partition $V = U_1 \cup \cdots \cup U_L$ there is an ε-regular partition $\{V_0, V_1, \ldots, V_K\}$ with $m \leq K \leq M$, which refines $U_1 \cup \cdots \cup U_L$.*

Remark: While the theorem in [3] does not mention refinements, it follows easily from the proof.

A convenient structure associated with a regular partition $\{V_0, V_1, \ldots, V_K\}$ is the reduced graph $\mathcal{R}$, which has vertex set $\{V_1, \ldots, V_K\}$ with the property that $V_i V_j$ is an edge of $\mathcal{R}$ if (V_i, V_j) is an ε-regular pair. Note that by definition $\mathcal{R}$ has at least $(1 - \varepsilon)\binom{K}{2}$ edges.

We will also require the following counting lemma.

Lemma 6.2. *Let $T = (V, E)$ be a tournament and $V_1, \ldots, V_h$ be disjoint subsets of V. Suppose that for each $1 \leq i < j \leq h$ the pair (V_i, V_j) is ε-regular with density d_{ij}, with $d_{ji} = 1 - d_{ij}$. Then given an h-vertex tournament H with $V(H) = \{u_1, \ldots, u_h\}$, the number of copies of H in T, with u_i sent to V_i for all $i \in [h]$ is*

$$\left(\prod_{\overrightarrow{u_i u_j} \in E(H)} d_{ij} \pm C_h \varepsilon \right) \prod_{l \in [h]} |V_l|,$$

where C_h is a constant depending only on h.

Proof. By deleting directed edges of T which are not of the form $\overrightarrow{v_i v_j}$ with $v_i \in V_i$, $v_j \in V_j$ and $\overrightarrow{u_i u_j} \in E(H)$ and ignoring the directions of the remaining edges, this follows immediately from the usual graph counting lemma; see [31]. □

An embedding ϕ of a h-vertex tournament H into a tournament T is said to be partite with respect to the disjoint sets $U_1, \ldots, U_h \subset V(T)$ if each set U_i receives one vertex of the embedding. Let $\text{Emb}_T(H; U_1, \ldots, U_h)$ denote the set of all such ϕ and let $N_T^*(H; U_1, \ldots, U_h) = |\text{Emb}_T(H; U_1, \ldots, U_h)|$.

The following proposition gives a 'partite version' of property $\mathcal{P}_2^*(H)$.

Proposition 6.3. *Let H be an h-vertex tournament. Suppose that for some constants ρ, C and an n-vertex tournament T we have $N_T^*(H; U) = \rho |U|^h \pm C$ for all $U \subset V(T)$. Then $N_T^*(H; U_1, \ldots, U_h) = h! \rho \prod_{i \in [h]} |U_i| \pm 2^h C$ for all disjoint sets $U_1, \ldots, U_h \subset V(T)$.*

Proof. By the inclusion-exclusion principle we have

$$N_T^*(H; U_1, \ldots, U_h) = \sum_{r=1}^{h} (-1)^{h-r} \Big(\sum_{I \in \binom{[h]}{r}} N_T^*\big(H; \bigcup_{i \in I} U_i\big) \Big).$$

Using $N_T^*(H; \bigcup_{i \in I} U_i) = \rho \big(\sum_{i \in I} |U_i| \big)^h \pm C$ for all $I \subset [h]$ gives

$$N_T^*(H; U_1, \ldots, U_h) = \rho \times \Big(\sum_{r=1}^{h} (-1)^{h-r} \Big(\sum_{I \in \binom{[h]}{r}} \big(\sum_{i \in I} |U_i| \big)^h \Big) \Big) \pm 2^h C = \rho \times \big(h! \prod_{i \in [h]} |U_i| \big) \pm 2^h C.$$

The final equality holds as the summed term in the penultimate equation can be viewed as counting, using inclusion-exclusion, the maps $g : [h] \to \bigcup_{i \in [h]} U_i$ sending each $i \in [h]$ to distinct U_j. Indeed there are $h! \prod_{i \in [h]} |U_i|$ such g which intersect each of the h sets U_i and for each $1 \le r \le h$ there are $\sum_{I \in \binom{[h]}{r}} \big(\sum_{i \in I} |U_i| \big)^h$ such maps which intersect at most r of the U_i's. $\qquad\square$

6.2 $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing imply local forcing

Before proceeding to the main result of this section, we note a simple consequence of Ramsey's theorem.

Lemma 6.4. *Given $\alpha > 0$ and $k, \ell \in \mathbb{N}$ there is $\gamma = \gamma(k, \ell, \alpha) > 0$ and $n_0 = n_0(k, \ell, \alpha) \in \mathbb{N}$ with the following property. Suppose that G is an n-vertex graph with $n \ge n_0$ and at least $(1 - \gamma)\binom{n}{2}$ edges. Then in any k-colouring of $E(G)$ there are vertex disjoint sets $U_1, \ldots, U_M \subset V(G)$ so that:*
  (i) $G[U_m]$ is a monochromatic clique of order ℓ for all $m \in [M]$;
  (ii) $|V(G) \setminus \big(\cup_{m \in [M]} U_m \big)| \le \alpha n$.

Proof. Set $R := R_k(\ell)$, the k-colour Ramsey number of an ℓ-vertex clique. Also set $\gamma = \alpha^2 / 2R$ and $n_0 = \lceil 2R / \alpha \rceil$. Suppose we are given a k-colouring of $E(G)$ as in the statement. To prove the lemma it suffices to show that every $W \subset V(G)$ with $|W| \ge \alpha n$ contains a set $U \subset W$ with $|U| = \ell$ so that $G[U]$ is a monochromatic clique. Indeed, using this property we can then greedily find sets $U_1, \ldots, U_M$ as in the lemma.

To see that this holds, first note that since $|W| \ge \alpha n \ge 2R$ and $|W|^2 \ge \alpha^2 n^2 = 2\gamma R n^2$, we have

$$e(G[W]) \ge \binom{|W|}{2} - \gamma \binom{n}{2} > \Big(1 - \frac{1}{R}\Big) \frac{|W|^2}{2} + \Big(\frac{|W|^2}{2R} - \frac{|W|}{2} - \frac{\gamma n^2}{2} \Big) \ge \Big(1 - \frac{1}{R}\Big) \frac{|W|^2}{2}.$$

By Turán's theorem there is $W' \subset W$ such that $|W'| = R$ and $G[W']$ is complete. As $G[W]$ is k-coloured, from Ramsey's theorem and the definition of R, there is $U \subset W'$ with $|U| = \ell$ and $G[U]$ is monochromatic. This completes the proof. $\qquad\square$

We now turn to the proof of the 'if' part of Theorem 1.6.

Theorem 6.5. *Any $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing tournament is locally forcing.*

Proof. Let H be a $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing tournament with $|H| = h$. We are required to show that given $\theta > 0$ there is $\delta > 0$ and n_0 such that the following holds. Suppose that T is an n-vertex tournament with $n \ge n_0$ which satisfies

$$N_T^*(H; U) = 2^{-\binom{h}{2}} |U|^h \pm \delta n^h \tag{13}$$

for all $U \subset V(T)$. Then any ordering of $V(T)$ has at most $\frac{1}{2}\binom{n}{2} + \theta n^2$ forward edges.

Let us now give a sketch of the proof before delving into the details. Let $v_1, \ldots, v_n$ be an ordering of $V(T)$. We begin by splitting the vertices into consecutive sets U_ℓ of almost equal size. Then we

apply the regularity lemma (Theorem 6.1) to refine this partition. Taking the reduced graph we define $\mathcal{R}_\ell$ to be its subgraph consisting of clusters contained in U_ℓ. We colour all edges of $\mathcal{R}_\ell$ which join clusters with density between them belonging to the same small interval in the colour indexed by this interval. If we split $[0,1]$ into finitely many such intervals we obtain a colouring of $\mathcal{R}_\ell$ to which we can apply Lemma 6.4 to group most of the clusters inside each $\mathcal{R}_\ell$ into monochromatic cliques. We now show that all edges inside these cliques must have density close to $1/2$. To see this assume the opposite, so that there is a clique C with all edges having density bounded away from $1/2$. Now using Lemma 6.2 and H being $\mathcal{T}_{cliq}$-forcing we conclude that there are too few copies of H between the clusters of C, compared to what is guaranteed by Proposition 6.3, which holds by (13).

We then proceed to upper bound the number of forwards edges of T. The main contribution comes from edges between ε-regular pairs of clusters between different U_ℓ's. To bound this number for a pair of cliques belonging to different U_ℓ's and a fixed $d > 1/2 + \theta$ we build an auxiliary bipartite graph with clusters of the two cliques making the sides of the bipartition and making an edge for any pair of ε-regular clusters which have density roughly d in the forwards direction. We show that this auxiliary graph can not contain $K_{a,h-a}$ for some a, as otherwise by using a similar argument as before we get h clusters between which we have a wrong count of the copies of H using the fact H is $\mathcal{T}_{bip}$-forcing. By grouping densities and applying the above reasoning for each group we show that there are few forwards edges between the two cliques. Trivially bounding the remaining contributions we show that there are fewer forwards edges than required, completing the proof.

Before beginning we will fix a number of parameters to be used in the proof. Let $\xi = \xi(H, \theta)$ be the minimum of the continuous function $f(x) = \max_{a \in [h-1]}(|2^{\binom{h}{2}} p_{H,a}(x) - 1|)$ on the interval $[1/2 + \theta, 1]$. Note that since H is $\mathcal{T}_{bip}$-forcing we have $f(x) > 0$ for each x in this range so since f is continuous we get $\xi > 0$. Set $\eta = \xi 2^{-\binom{h}{2}-2} h^{-2}$. Since H is $\mathcal{T}_{cliq}$-forcing and $p_H(x)$ is continuous there is $\zeta \in (0, \eta)$ with the property that if $x \in [0,1]$ and $p_H(x) \geq 2^{-\binom{h}{2}} - \zeta$ then $x = (1 \pm \eta)/2$. Take $\alpha = \theta/64$, $m = \max(\lceil 4h^2/\zeta \rceil, \lceil 2/\eta \rceil)$, $L = \lceil 16\theta^{-1} \rceil$ and $N_1 = \lceil (2(\theta\eta)^{-1}h)^h \rceil$. Also set $m_{min} = 4Ln_0(m, N_1, \alpha)$ and $\gamma = \gamma(m, N_1, \alpha)$ as in Lemma 6.4. With C_h as in Lemma 6.2, take $\varepsilon > 0$ so that

$$\varepsilon = \min\left(\frac{1}{4L}, \frac{\gamma}{8L^2}, \frac{\zeta}{4C_h}, \frac{\theta^3}{32}\right).$$

Lastly, set $n_0 = M = \max(M(m_{min}, L, \varepsilon), 8L)$ as in Theorem 6.1 and $\delta = \zeta(4M)^{-h}/2$.

To begin the proof set $U_\ell = \{v_i \in V(T) : i \in ((\ell-1)n/L, \ell n/L]\}$ for all $\ell \in [L]$. Note that $|U_\ell| \geq n/L - 2$. Provided $n \geq n_0$, we may apply the Theorem 6.1 to T to obtain an ε-regular partition $\{V_k\}_{k \in [K]} \cup \{V_0\}$ refining $\{U_\ell\}_{\ell \in [L]}$, with $m_{min} \leq K \leq M$. Let $\mathcal{R}$ denote the reduced graph of this partition and $\mathcal{R}_\ell$ denote the subgraph of $\mathcal{R}$ consisting of the clusters contained in U_ℓ. Setting $W_\ell = |\mathcal{R}_\ell|$, we have $W_\ell \geq \frac{K}{2L}$ for all $\ell \in [L]$. Indeed, since $|V_k| = |V_{k'}|$ for all $k, k' \in [K]$ we have

$$n \geq \sum_{k \in [K]} |V_k| = \frac{K}{W_\ell} \sum_{V_k \in V(\mathcal{R}_\ell)} |V_k| \geq \frac{K}{W_\ell}\left(|U_\ell| - |V_0|\right) \geq \frac{K}{W_\ell} \times \frac{n}{2L},$$

using $|U_\ell| - |V_0| \geq (n/L - 2) - \varepsilon n \geq \frac{n}{2L}$. Rearranging, we find $W_\ell \geq \frac{K}{2L}$ for all $\ell \in [L]$.

Claim: Each $\mathcal{R}_\ell$ contains a collection of vertex disjoint cliques $\mathcal{C}_\ell$ with the following properties:
  (i) Each clique $C \in \mathcal{C}_\ell$ has order N_1,
  (ii) $|\mathcal{R}_\ell \setminus (\cup_{C \in \mathcal{C}_\ell} C)| \leq \alpha W_\ell$,
  (iii) $d(V_k, V_{k'}) = 1/2 \pm \eta$ for each edge $V_k V_{k'}$ in a clique $C \in \mathcal{C}_\ell$.

To prove the claim, colour the edges of $\mathcal{R}_\ell$ with m colours, where each pair $V_k V_{k'}$ with $k < k'$ gets color $j \in [m-1]$ if $d(V_k, V_{k'}) \in j/m \pm 1/m$ (ties broken arbitrarily). The graph $\mathcal{R}_\ell$ contains at least $\binom{W_\ell}{2} - \varepsilon\binom{K}{2} > (1 - \gamma)\binom{W_\ell}{2}$ edges, since $W_\ell \geq \frac{K}{2L}$, $\frac{\gamma}{8L^2} \geq \varepsilon$ and $K \geq m_{min} \geq 4L$. Therefore, from Lemma 6.4, since $|\mathcal{R}_\ell| \geq \frac{K}{2L} \geq \frac{m_{min}}{2L} \geq n_0(m, N_1, \alpha)$, the graph $\mathcal{R}_\ell$ contains a collection $\mathcal{C}_\ell$ of vertex disjoint monochromatic cliques which satisfy parts (i) and (ii) from the claim.

It remains to show that part (iii) holds. Let $C \in \mathcal{C}_\ell$ be monochromatic with colour j and $V_{k_1}, \ldots, V_{k_h} \in C$ with $k_1 < \ldots < k_h$. As each pair $(V_k, V_{k'})$ in C is ε-regular with $d(V_{k_i}, V_{k_{i'}}) = (j \pm 1)/m$, by Lemma 6.2 we have

$$
\begin{aligned}
N_T^*(H; V_{k_1}, V_{k_2}, \ldots, V_{k_h}) &= \sum_{\sigma \in S_h} \left(\left(\frac{j \pm 1}{m}\right)^{|F_{\sigma,H}|} \left(\frac{m - j \pm 1}{m}\right)^{|B_{\sigma,H}|} \pm C_h \varepsilon \right) \prod_{i \in [h]} |V_{k_i}| \\
&= h! \big(p_H(j/m) \pm (C_h \varepsilon + h^2 m^{-1}) \big) \prod_{i \in [h]} |V_{k_i}| \\
&= h! \big(p_H(j/m) \pm \zeta/2 \big) \prod_{i \in [h]} |V_{k_i}|.
\end{aligned}
$$

In the second equality here we used repeatedly the fact that for any $x, y, t \in [0,1]$ such that $x \pm t \in [0,1]$ we have $(x \pm t)(y \pm t) = (x \pm t)y \pm t = xy \pm 2t$. On the other hand, from (13) and Proposition 6.3 we also have

$$
N_T^*(H; V_{k_1}, V_{k_2}, \ldots, V_{k_h}) = h! 2^{-\binom{h}{2}} \prod_{i \in [h]} |V_{k_i}| \pm 2^h \delta n^h = h! \big(2^{-\binom{h}{2}} \pm \zeta/2 \big) \prod_{i \in [h]} |V_{k_i}|. \tag{14}
$$

Here we have used that all clusters V_k satisfy $|V_k| \geq (n - |V_0|)/K \geq n/2M$ and $\delta(4M)^h = \zeta/2$. Combined, these bounds give $p_H(j/m) = 2^{-\binom{h}{2}} \pm \zeta$. By our choice of ζ this forces $j/m = (1 \pm \eta)/2$ and so $d(V_k, V_{k'}) = (1 \pm \eta)/2 \pm m^{-1} = 1/2 \pm \eta$ for any pair $V_k V_{k'}$ contained in a clique $C \in \mathcal{C}_\ell$ giving (iii). This completes the proof of the claim.

We can now proceed to prove an upper bound on the number of forward edges of T. To do this, first fix $1 \leq \ell < \ell' \leq L$ and $C \in \mathcal{C}_\ell$ and $C' \in \mathcal{C}_{\ell'}$. We will prove that

$$
\sum_{\substack{V_k \in C, V_{k'} \in C': \\ (V_k, V_{k'}) \ \varepsilon-\mathrm{reg}}} |E(V_k, V_{k'})| \leq \left(\frac{1}{2} + \frac{3\theta}{2}\right) \Big| \bigcup_{V_k \in C} V_k \Big| \Big| \bigcup_{V_{k'} \in C'} V_{k'} \Big|. \tag{15}
$$

To see this fix some $d > 1/2 + \theta$ and consider the auxilliary bipartite graph G with vertex set on one side being C and on the other side C'. We put an edge between clusters $V_k \in C$ and $V_{k'} \in C'$ if the pair $(V_k, V_{k'})$ is ε-regular and has density $d(V_k, V_{k'}) = d \pm \eta$. Let a be such that $|2^{\binom{h}{2}} p_{H,a}(d) - 1| \geq \xi$, which exists since $d > 1/2 + \theta$. Our goal is to show that G is $K_{a,h-a}$-free, which by Kővári-Sós-Turán theorem is going to tell us that G is sparse, allowing us to bound the number of forwards edges of T between C and C'.

To see this assume towards a contradiction that G contains a $K_{a,h-a}$. Let $V_{k_1}, \ldots, V_{k_a}$ make one side and $V_{k_{a+1}}, \ldots, V_{k_h}$ the other of this $K_{a,h-a}$. By part (iii) of the claim, all pairs $(V_{k_i}, V_{k_{i'}})$ are ε-regular with $d(V_{k_i}, V_{k_{i'}}) = 1/2 \pm \eta$ for distinct $i, i' \in [a]$ or distinct $i, i' \in [a+1, h]$. Any $\phi \in \mathrm{Emb}_T(H; V_{k_1}, \ldots, V_{k_h})$ embeds $\binom{a}{2} + \binom{h-a}{2}$ edges of H into pairs with density $1/2 \pm \eta$. If $A := \phi^{-1}(\{V_{k_1}, \ldots, V_{k_a}\})$ then $e(A, V \setminus A)$ edges of H get embedded into pairs with density $d \pm \eta$, and $e(V \setminus A, A)$ edges into pairs with density $1 - d \pm \eta$. So Lemma 6.2 gives

$$
\begin{aligned}
\frac{N_T^*(H; V_{k_1}, \ldots, V_{k_h})}{\prod_{i \in [h]} |V_{k_i}|} &= a!(h-a)! \sum_{A \in \binom{V(H)}{a}} \left(\left(\frac{1}{2} \pm \eta\right)^{\binom{a}{2} + \binom{h-a}{2}} (d \pm \eta)^{e(A, V \setminus A)} (1 - d \pm \eta)^{e(V \setminus A, A)} \pm C_h \varepsilon \right) \\
&= h! \binom{h}{a}^{-1} \sum_{A \in \binom{V(H)}{a}} \left(2^{-\binom{a}{2} - \binom{h-a}{2}} d^{e(A, V \setminus A)} (1 - d)^{e(V \setminus A, A)} \pm (h^2 \eta + C_h \varepsilon) \right) \\
&= h! (p_{H,a}(d) \pm (h^2 \eta + C_h \varepsilon)). \tag{16}
\end{aligned}
$$

As $(h^2 \eta + C_h \varepsilon) + \zeta \leq h^2 \eta + 2\zeta \leq 3h^2 \eta < \xi 2^{-\binom{h}{2}}$ this gives

$$
\left| \frac{N_T^*(H; V_{k_1}, \ldots, V_{k_h})}{\prod_{i \in [h]} |V_{k_i}| h!} - 2^{-\binom{h}{2}} \right| \geq \zeta,
$$

17

which contradicts (14). Thus, there is no $K_{a,h-a}$ in G. The Kővári-Sós-Turán theorem (see [15]) now tells us that G has at most $hN_1^{2-1/h}$ edges (recall that $|C| = |C'| = N_1$.) Since d was arbitrary (provided it is bigger than $1/2 + \theta$) by splitting the interval $[1/2 + \theta, 1]$ into at most η^{-1} subintervals of width at most 2η and building a graph as above for each of these subintervals with d being equal to the center of the interval we obtain that there are at most $\eta^{-1}hN_1^{2-1/h}$ pairs $V_k \in C, V_{k'} \in C'$ such that $(V_k, V_{k'})$ is ε-regular and $d(V_k, V_{k'}) > 1/2 + \theta$. As $|C| = N_1 \geq (2(\eta\theta)^{-1}h)^h$ this gives

$$\sum_{\substack{V_k \in C, V_{k'} \in C': \\ (V_k, V_{k'}) \ \varepsilon-reg}} |E(V_k, V_{k'})| \leq \left(\frac{1}{2}+\theta+\eta^{-1}hN_1^{-1/h}\right)\left|\bigcup_{V_k \in C} V_k\right|\left|\bigcup_{V_{k'} \in C'} V_{k'}\right| \leq \left(\frac{1}{2}+\frac{3\theta}{2}\right)\left|\bigcup_{V_k \in C} V_k\right|\left|\bigcup_{V_{k'} \in C'} V_{k'}\right|$$

i.e. (15) holds.

We can now complete the proof, upper bounding the number of forward edges of T. The ε-regular pairs $(V_k, V_{k'})$ with $V_k \in C \in \mathcal{C}_\ell$ and $V_{k'} \in C' \in \mathcal{C}_{\ell'}$ for some $\ell < \ell'$ contribute at most $(1/2 + 3\theta/2)\binom{n}{2}$ forward edges to T by (15). The remaining forward edges either (a) lie entirely in some set U_ℓ, (b) contain a vertex from cluster $V_k \notin \bigcup_\ell \bigcup_{C \in \mathcal{C}_\ell} C$, (c) contain a vertex in V_0, or (d) lie between pairs $(V_k, V_{k'})$ which are not ε-regular. The number of such edges is at most

$$\sum_{\ell \in [L]} \binom{|U_\ell|}{2} + \sum_{\ell \in [L]} |\mathcal{R}_\ell \setminus (\cup_{C \in \mathcal{C}_\ell} C)| \left(\frac{n^2}{K}\right) + |V_0|n + \varepsilon\binom{K}{2}\left(\frac{n}{K}\right)^2 \leq \left(\frac{2}{L} + 8\alpha + 3\varepsilon + \varepsilon\right)\binom{n}{2} \leq \frac{\theta}{2}\binom{n}{2}.$$

Here we have used that $|\mathcal{R}_\ell \setminus (\cup_{C \in \mathcal{C}_\ell} C)| \leq \alpha W_\ell \leq 2\alpha K/L$ by part (ii) of the claim, that $L \geq 16/\theta$, $64\alpha = \theta$ and $16\varepsilon \leq \theta$. Combined, these estimates show that T has at most $\frac{1}{2}\binom{n}{2} + \theta n^2$ forward edges, as required. $\qquad\square$

7 Concluding remarks and open problems

We have shown that a large tournament H is globally forcing if and only if H is transitive. Our main focus was on the stronger property of being locally forcing. We proved that while many tournaments do not satisfy this property (Lemma 3.1) it does hold for many tournaments which we draw from a certain random distribution on nearly regular tournaments. The most natural model of random regular tournaments is to take a uniform distribution over all regular tournaments. We believe that in fact this also gives w.h.p. a $\mathcal{T}_{cliq}$-forcing and hence (by Lemma 5.2 and Theorem 1.6) a locally forcing tournament.

Another result in this paper shows that a tournament is locally forcing if and only if it is $\mathcal{T}_{cliq}$-forcing and $\mathcal{T}_{bip}$-forcing. This in some sense says that in order to check whether a tournament H is locally forcing one only needs to check whether models $\mathcal{T}(n, \alpha)$ or $\mathcal{T}(n, n, \alpha)$ can have the same local counts of H as $\mathcal{T}(n, 1/2)$ for some $\alpha \neq 1/2$.

We actually believe that the $\mathcal{T}_{bip}$-forcing assumption may be dropped entirely in Theorem 1.6 either because it is implied by $\mathcal{T}_{cliq}$-forcing or because it is satisfied by every tournament. In other words it would be interesting to determine if every $\mathcal{T}_{cliq}$-forcing tournament is locally forcing. We reduced this question to the following problem about degree counting polynomials. Does there exist an h-vertex tournament H such that the rescaled and recentered a-th order degree counting polynomials defined as

$$q_{H,a}(x) := \binom{h}{a}^{-1} \sum_{A \in \binom{V(H)}{a}} (1+x)^{e(A,V\setminus A)}(1-x)^{e(V\setminus A,A)} - 1$$

have a common root in $(0, 1]$, for all $1 \leq a \leq h - 1$. We note that in a certain sense this is the correct tournament analogue of the Simonovits-Sós conjecture from the induced graph case, discussed in the introduction.

Our arguments rely on the regularity lemma. It is possible that one can find a nice class of tournaments for which one can show the locally forcing property directly, avoiding the use of the

18

regularity lemma. For example, the tournament Tr_3 is not globally forcing but is locally forcing and this is not hard to see directly. Indeed, any tournament T has at most $\frac{1}{2}\sum_{v \in T}\left(\binom{d^+(v)}{2} + \binom{d^-(v)}{2}\right) = \frac{1}{4}\sum_{v \in T}\frac{1}{2}\left((d^+(v)+d^-(v))^2 + (d^+(v)-d^-(v))^2\right) - (n-1) = \frac{1}{8}n^3(1+o(1)) + \frac{1}{8}\sum_{v \in T}(d^+(v)-d^-(v))^2$ copies of Tr_3. So in order for the local counts to match that of the random tournament, by Cauchy-Schwarz, we must have $n^{-1}\left(\sum_{v \in U}|d^+_{T[U]}(v) - d^-_{T[U]}(v)|\right)^2 \leq \sum_{v \in U}(d^+_{T[U]}(v) - d^-_{T[U]}(v))^2 = o(n^3)$ for any $U \subseteq V(T)$. This gives condition $\mathcal{P}_5$ from Chung and Graham [6] and implies T must be quasirandom. Thus Tr_3 is locally forcing. Since any three vertices of a tournament either induce Tr_3 or C_3, if any subset of $V(T)$ has the correct count of Tr_3 then it has the correct count of C_3 as well. Thus C_3 is also locally forcing.

We note that deciding whether a fixed tournament is $\mathcal{T}_{cliq}$-forcing is a matter of counting how many zeros of the counting polynomial (minus a constant) one can find in the interval $[1/2, 1]$. Since counting polynomial is of degree at most $\binom{h}{2}$ this can be done by Sturm's algorithm in polynomial time. Deciding whether a fixed tournament is $\mathcal{T}_{bip}$-forcing can be done in a similar fashion, we first find the greatest common divisor of our degree counting polynomials and then find the number of roots of this greatest common divisor in $[1/2, 1]$. While all of this can be done in polynomial time it is not clear how to compute the coefficients of our polynomials in polynomial time since they are defined in terms of vertex orderings of which there are $h!$ or subsets, in which case there are potentially as many as $\binom{h}{h/2}$. It could be interesting to determine if these terms could be calculated in polynomial time, as this would give a polynomial time algorithm for deciding whether a given tournament is locally forcing or not.

Finally, it was brought to our attention that Fox, Himwich and Mani [17] independently studied related problems in general directed graphs including certain forcing problems in this setting.

Acknowledgements

References

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

[2] N. Alon, G. Gutin and M. Krivelevich, Algorithms with large domination ratio, *Journal of Algorithms*, **50**(1) (2004), 118–131.

[3] N. Alon and A. Shapira, Testing subgraphs of directed graphs, *J. Comput. System Sci.*, **69** (2004), 353-382.

[4] N. Alon and J.H. Spencer, **The probabilistic method**, Wiley, 4th ed. (2016).

[5] S. Basu, R. Pollack and M.-F. Roy, **Algorithms in real algebraic geometry**, Springer, 2nd ed. (2006).

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

[7] F. Chung and R. L. Graham, Quasi-random set systems, *J. Amer. Math. Soc.* **4** (1991), 151–196.

[8] F. Chung, R. L. Graham and R. M. Wilson, Quasi-random graphs, *Combinatorica* **9** (1989), 345–362.

[9] D. Conlon, J. Fox and B. Sudakov, Hereditary quasirandomness without regularity. *Math. Proc. Cambridge Philos. Soc.* 164 (2018), 385-399.

[10] D. Conlon, H. Hàn, Y. Person and M. Schacht, Weak quasi-randomness for uniform hypergraphs, *Random Structures Algorithms* **40** (2012), 1-38.

[11] J. N. Cooper, Quasirandom permutations, *J. Combin. Theory Ser. A* **106** (2004), 123–143.

[12] L.N. Coregliano, R.F. Parente and C.M. Sato, On the maximum density of fixed strongly connected subtournaments, *Electron. J. Combin.* **26** (2019), no. 1, Paper 1.44.

[13] L. N. Coregliano and A. A. Razborov, On the density of transitive tournaments, *J. Graph Theory* **85** (1), (2017) 12–21.

[14] D. Dellamonica Jr. and V. Rödl, Hereditary quasirandom properties of hypergraphs, *Combinatorica* **31** (2), (2011) 165–182.

[15] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math. 2* (1964), 183–190.

[16] W. Fernandez de la Vega, On the maximum cardinality of a consistent set of arcs in a random tournament. *J. Combin. Theory Ser. B*, **35** (3), (1983), 328–332.

[17] J. Fox, Z. Himwich and N. Mani, personal communication.

[18] W. T. Gowers, Hypergraph regularity and the multidimensional Szemerédi theorem, *Annal. Math.* **166** (2007), 897–946.

[19] W. T. Gowers, Quasirandom groups, *Combin., Probab. Comput.*, **17**, (2008), 363–387.

[20] S. Griffiths, Quasi-random oriented graphs, *J. Graph Theory*, **74** (2), (2013), 198–209.

[21] T. Gustavsson, Decompositions of large graphs and digraphs with high minimum degree, PhD thesis, Univ. of Stockholm, (1991).

[22] R. Hancock, A. Kabela, D. Kral, T. Martins, R. Parente, F. Skerman, J. Volec. *No additional tournaments are quasirandom-forcing.* preprint arXiv:1912.04243.

[23] S. Kalyanasundaram and A. Shapira, A Note on Even Cycles and Quasi-Random Tournaments, *J. Graph Theory* **73** (2013), 260–266.

[24] T.P. Kirkman, On a problem in combinations, *The Cambridge and Dublin Mathematical Journal (Macmillan, Barclay, and Macmillan) II* (1847), 191–204.

[25] D. Král' and O. Pikhurko, Quasirandom permutations are characterized by 4-point densities, *GAFA*, **23** (2013), 570–579.

[26] M. Krivelevich and B. Sudakov, Pseudo-random graphs. In *More Sets, Graphs and Numbers, Bolyai Society Mathematical Studies 15*, 199-262. Springer, 2006.

[27] L. Lovász, **Combinatorial problems and exercises**, North-Holland Publishing Co., 2nd ed. (1993).

[28] B. Nagle, V. Rödl and M. Schacht, The counting lemma for regular k-uniform hypergraphs, *Random Structures Algorithms* **28** (2006), 113–179.

[29] R. O'Donnell, **Analysis of Boolean functions**, Cambridge University Press (2014).

[30] V. Rödl and J. Skokan, Regularity lemma for k-uniform hypergraphs, *Random Structures Algorithms* **25** (2004), 1–42.

[31] M. Schacht, Regularity lemma and its applications, lecture notes, `https://www.math.uni-hamburg.de/home/schacht/lnotes/GT/SzRL.pdf`

[32] M. Simonovits and V.T. Sós, Hereditary extended properties, quasi-random graphs and induced subgraphs, *Combin. Probab. Comput.*, **12** (2003), 319–344.

[33] M. Simonovits and V.T. Sós, Hereditarily extended properties, quasi-random graphs and not necessarily induced subgraphs *Combinatorica*, **17** (1997), 577–596.

[34] J. Spencer, Optimal ranking of tournaments, *Networks*, **1** (1971), 135–138.

[35] E. Szemerédi, Regular partitions of graphs, *Problmes combinatoires et thorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, Colloq. Internat. CNRS **260** (1978), 399–401.

[36] A. Thomason, Pseudo-random graphs, Proceedings of Random Graphs, Poznan 1985 (M. Karonski, ed.) *Annals of Discrete Math.* **33** (1987), 307–331.

[37] A. Thomason, Random graphs, strongly regular graphs and pseudo-random graphs, in Surveys in Combinatorics 1987 (C. Whitehead, ed.) LMS Lecture Notes Series 123, Cambridge Univ. Press, Cambridge, 173–196.