Random Tournaments

Tejas Iyer

October 2016

A thesis submitted in partial fulfilment of the degree
of Bachelor of Philosophy (Honours).

Australian
National
University

For my grandmothers.

Acknowledgements

I'd like to thank my supervisor Professor Brendan McKay and Dr Mikhail Isaev. Both of them dedicated hours to helping me with this thesis, and without them none of the work would have been possible. I am very grateful for their support.

Thanks also to Associate Professor Scott Morrison for acting as my co-supervisor within the MSI, and providing me with a lot of feedback on earlier drafts. His advice was often very helpful, and I really appreciate the time he spent proofreading.

I'd also like to thank Dr Pierre Portal, who provided me with helpful comments on the final draft of my thesis and Dr Joan Licata for her tireless efforts as Honours coordinator in making the Honours program run as smoothly as possible.

Finally, I'd like to thank my family and friends for all of their support over the years.

Contents

Introduction

Tournaments are an important class of directed graph. They were first introduced by Landau in 1953 [Lan53] in order to model dominance relations in flocks of chickens and nowadays are used widely in comparison based ranking, with applications to biology, chemistry, networks and sports.

This thesis concerns itself with results related to *random tournaments,* which are probability distributions over tournaments. In Chapter 1, we provide a brief introduction to the theory of tournaments and prove many famous results in the area. Then, in Chapter 2, we introduce random tournaments and provides a brief survey of results in the field.

In Chapter 3, we introduce the combinatorial objects of *multi-tournaments, generalised multi-tournaments* and *generalised tournaments* which are closely related to random tournaments. By using the tools of convex analysis, we prove some interesting results related to these objects which, for example, have interesting implications for voting theory.

Chapter 4 uses the theory developed in Chapter 3 to derive results related to a special type of random tournament called the β-model. This has many interesting applications to the study of comparison based ranking. By making use of results from this chapter, in Chapter 5 we derive an asymptotic formula for the number of tournaments with a particular *score vector* (see Definition 1.1.21) that works for a wide range of 'tame' score vectors. This provides a means for one to calculate the probability of a tournament with a particular 'tame' score vector with respect to any β-model random tournament.

The methodologies in Chapters 3, 4 and 5 are mostly original, inspired by similar results in the field of random graphs and the direction of the author's supervisor Brendan McKay and Mikhail Isaev. The results are also new, unless otherwise stated.

Notation

This thesis will use the following notation:

Notation 0.0.1. Suppose n is a positive integer. We denote by $[n]$ the set $\{1, 2, \cdots, n\}$.

Notation 0.0.2. Let $f(\boldsymbol{x})$ be a complex valued function and $g(\boldsymbol{x})$ be a real valued function with both functions taking inputs in $\mathbb{R}^n$. We say $f(\boldsymbol{x}) = O(g(\boldsymbol{x}))$ as $\boldsymbol{x} \to \infty$ if there exists some positive constant M such that

$$|\Re f(\boldsymbol{x})| \leq M|g(\boldsymbol{x})|$$

and

$$|\Im f(\boldsymbol{x})| \leq M|g(\boldsymbol{x})|,$$

for $|\boldsymbol{x}|$ sufficiently large.

Notation 0.0.3. Let $f(x)$ and $g(x)$ be real valued functions taking inputs in $\mathbb{R}$. We say $f(x) = \Theta(g(x))$ if for x sufficiently large there exist positive constants c_1 and c_2 such that

$$c_1|g(x)| \leq |f(x)| \leq c_2|g(x)|.$$

Also note that for convenience we will often write vectors in $\boldsymbol{x} \in \mathbb{R}^n$ as a list indicating the components of the vector, so that if $\boldsymbol{x}$ has components $x_1, \cdots, x_n$, we write $\boldsymbol{x} = (x_1, \cdots, x_n)$.

Chapter 1

Tournaments

In this chapter, we will assume that the reader has some basic familiarity with graph theory.

1.1 Basic Definitions

1.1.1 Preliminaries

Intuitively we may think of a tournament on n vertices as recording the results of a round-robin competition with n players, where each game can only result in a win or a loss. In order to make this precise, we first recall some definitions about directed graphs.

Directed Graphs

Definition 1.1.1. A *directed graph,* or *digraph* D is an ordered pair (V, E), where V is a set of *vertices,* and E is a set of *directed edges* consisting of ordered pairs of vertices in V.

To avoid confusion between vertex and edge sets of different directed graphs, if D is a directed graph, we use the notation $V(D)$ and $E(D)$ to denote the vertex set of D and edge set of D respectively. In general, we will only consider directed graphs whose vertex set is finite. Moreover, we will assume that there are no loops, that is, edges of the form (v, v).

Definition 1.1.2. Two directed graphs D and D' are said to be *vertex disjoint* if $V(D) \cap V(D') = \emptyset$ and *edge disjoint* if $E(D) \cap E(D') = \emptyset$.

A directed graph may be represented as a diagram where each directed edge (u, v) between vertices u and v is represented as an arrow from u to v. In fact, a lot of the intuition for this subject comes from the pictures, the formalism is just a means for making these ideas rigorous.

Example 1.1.3. Figure 1.1 is an example of a directed graph on 5 vertices represented as a diagram.

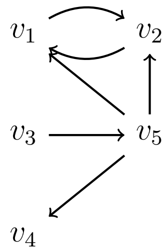


Figure 1.1: A directed graph on 5 vertices

Definition 1.1.4. Let D be a directed graph. For any vertex $v \in V(D)$ the *out-degree* of v, denoted $\deg^+(v)$, is the number of directed edges in $E(D)$ directed from v (i.e edges of the form (v, u), where $u \in V(D)$). The *in-degree* of v, denoted $\deg^-(v)$ is the number of directed edges in $E(D)$ directed to v (i.e edges of the form (u, v), where $u \in V(D)$). An *isolated* vertex v has $\deg^+(v) = \deg^-(v) = 0$.

Example 1.1.5. In Figure 1.1, $\deg^+(v_5) = 3$ and $\deg^-(v_5) = 1$.

Often it is useful to consider digraphs contained in larger digraphs.

Definition 1.1.6. Let D be a digraph. A *subdigraph* D' of D is a digraph such that $V(D') \subseteq V(D)$ and $E(D') \subseteq E(D)$. Given a subset $U \subseteq V(D)$, the subdigraph of D *induced by* U is the subdigraph $D_U = (U, E(D_U))$ such that each edge $(u, v) \in E(D)$ with $u, v \in U$ is contained in $E(D_U)$.

Example 1.1.7. Figure 1.2 shows a digraph D on vertex set $V(D) = \{v_1, v_2, v_3, v_4, v_5\}$, and the subdigraph D_U induced by $U = \{v_2, v_3, v_5\}$ coloured in red.

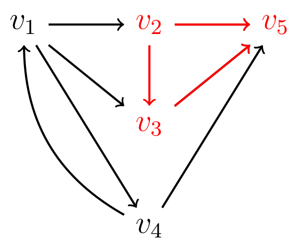


Figure 1.2: A subdigraph induced by a set of vertices.

It is also useful to consider *directed paths* and *directed cycles* in directed graphs.

Definition 1.1.8. A *directed trail* from vertices v_0 to v_n in a directed graph D is an alternating sequence $v_0 e_1 v_1 e_2 \ldots v_{n-1} e_n v_n$ of vertices $v_i \in V(D)$ and directed edges $e_i \in E(D)$, where each $e_i = (v_i, v_{i+1})$. Moreover, every edge in the sequence is distinct. The *length* of a trail is the number of edges in it. A directed trail in which all the vertices are distinct, except possibly the first and the last vertex, is called a *directed path.* A directed path which starts and ends at the same vertex is called a *directed cycle.*

Definition 1.1.9. A *Hamiltonian path* in a directed graph D is a spanning directed path in it, that is a directed path that hits every vertex in D. A *Hamiltonian cycle* is a spanning directed cycle. A digraph that consists of exactly 1 Hamiltonian cycle is called a *cycle digraph* or a *cycle.*

Example 1.1.10. Figure 1.3 below illustrates a directed path (red) and a directed cycle (purple) in a digraph.

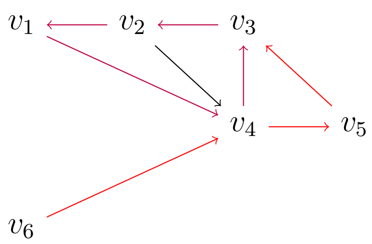


Figure 1.3: A directed path in a digraph from vertex v_6 to vertex v_3 (red), and a directed cycle (purple).

Definition 1.1.11. Let D be a digraph. If, given vertices u and v in $V(D)$, there is a directed path from u to v, we say that $\{u \to v\}$. This induces a binary relation on $V(D)$, called the *reachability* relation. In this case we say v is *reachable* from u.

A special type of digraph has every vertex reachable from each other:

Definition 1.1.12. A digraph D is called *strongly connected* or *strong* if, given any pair of vertices $u, v \in V(D)$, $\{u \to v\}$ and $\{v \to u\}$. A *strongly connected component* is a subdigraph of D which is strongly connected.

Since a single vertex digraph is strongly connected, evidently the vertices in the strongly connected components of a digraph D partition $V(D)$.

Example 1.1.13. Figure 1.4 shows the digraph of Figure 1.3 partitioned into strongly connected components.

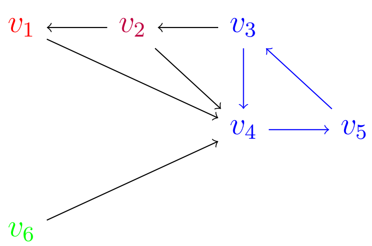


Figure 1.4: The strongly connected components of Figure 1.1, in blue, green, red and purple.

Oriented Graphs

The main type of directed graph we will encounter in this thesis, is called an *oriented graph.*

Definition 1.1.14. An *oriented graph* is a directed graph $D = (V, E)$ such that between any pair of vertices $u, v \in V$, there can only be one directed edge: (u, v) or (v, u).

The terminology 'oriented graph' comes from the fact that these directed graphs may be thought of as graphs where each edge is assigned an orientation, or direction. In particular, given a graph $G = (V, E)$, an *orientation* of G is a map $\psi : E \to V \times V$, such that each *unordered* pair of vertices $\{u, v\} \in E$ maps to exactly one of the *ordered* pairs $(u, v) \in V \times V$ or $(v, u) \in V \times V$. Such an orientation of G yields a directed graph $(V, \psi(E))$ in the natural way. Conversely, the condition that there can be exactly one directed edge between any two vertices implies that any oriented graph D is the orientation of some simple graph.

Definition 1.1.15. Given an oriented graph $D = (V, E)$, the *underlying simple graph* $S(D)$ is the graph obtained by removing the orientation of each edge. Precisely, $S(D) = (V, \psi^{-1}(E))$, where ψ^{-1} maps any ordered pair $(u, v) \in E$ to the unordered pair $\{u, v\}$.

Example 1.1.16. Figure 1.5 shows an oriented graph D on three vertices $\{v_1, v_2, v_3\}$, and the underlying simple graph $S(D)$.

$$v_1 \longrightarrow v_2 \longrightarrow v_3$$

(a) An oriented graph $D = (V, E)$ on three vertices. Here $V = \{v_1, v_2, v_3\}$ and $E = \{(v_1, v_2), (v_2, v_3)\}$.

$$v_1 \;\text{———}\; v_2 \;\text{———}\; v_3$$

(b) The underlying simple graph $S(D)$

Figure 1.5: An oriented graph and its underlying simple graph.

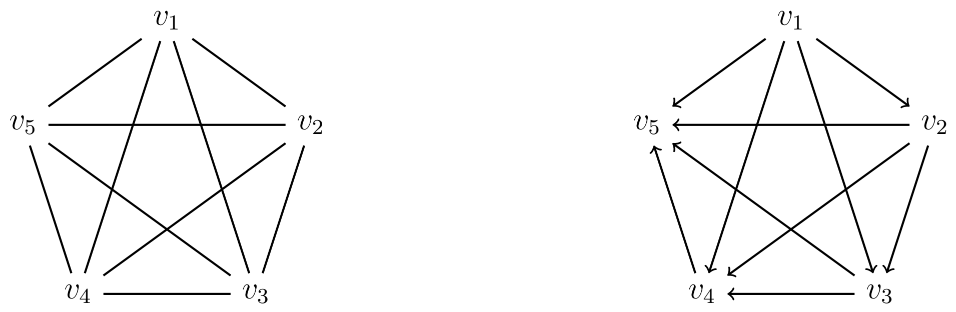
We now have the terminology we need to define *tournaments.*

1.1.2 Tournaments

Definition 1.1.17. A *tournament* on n vertices is a directed graph obtained by orienting every edge of the complete graph K_n. If, given vertices u and v, an edge is oriented from u to v, we say that u *dominates* v. Given $u, v \in V(T)$, a *game* between u and v is the unordered pair $\{u, v\}$ in the underlying simple graph K_n.

We may thus this of a tournament as the results of the $\binom{n}{2}$ possible different games between n players, where the result is recorded by orienting the edge.

Example 1.1.18. Figure 1.6 shows a complete graph on five vertices $\{v_1, v_2, v_3, v_4, v_5\}$ (left), and a tournament on the same set of vertices obtained by orienting each edge (right).



(a) The complete graph K_5 (b) A tournament on 5 vertices.

Figure 1.6: A tournament on 5 vertices obtained by orienting each edge of K_5

We denote by $\mathbb{T}_n$ the set of all tournaments on a set of vertices $\{v_1, \cdots, v_n\}$. Such a tournament contains $\binom{n}{2}$ edges, each of which can be oriented in exactly 2 ways,

$$|\mathbb{T}_n| = 2^{\binom{n}{2}}.$$

Continuing with the analogy of a round-robin competition, the out-degree of the vertex u then records the total number of wins by u, or the total *score* of u in the competition.

Definition 1.1.19. Given a tournament T, the *score* of a vertex $v \in V(T)$ is the out-degree of v in T.

It is often useful to record the results of each player in a round-robin tournament in a list. This leads us to the following definitions:

Definition 1.1.20. Let T be a tournament on n vertices. An *ordering* of the vertices of $V(T)$ is a bijection $\pi : V(T) \to [n]$.

Definition 1.1.21. Let T be a tournament on n vertices, with an ordering π. A *score vector corresponding to* π is a vector $\boldsymbol{s} \in \mathbb{R}^n$ such that each entry s_i is the score of the vertex $\pi^{-1}(i)$ in T.

If a tournament T has an indexed vertex set $\{v_1, \ldots, v_n\}$, we will generally use the term *score vector* to mean the score vector corresponding to the natural map π such that for each $i \in [n]$, $v_i \mapsto i$.

Often it is convenient to arrange the entries in the score vector in non-decreasing order, so that the score vector conveys information about the performance of the players. We give this type of score vector a name:

Definition 1.1.22. A *score sequence* of a tournament T on n vertices is a score vector $\boldsymbol{s}$ with an ordering such that the entries of the score vector are in non-decreasing order, i.e

$$s_1 \leq s_2 \leq \cdots \leq s_n.$$

Example 1.1.23. The following tournament on 4 vertices, has score vector $\boldsymbol{s} = (2, 1, 2, 1)$. The score sequence of this tournament is $(1, 1, 2, 2)$.

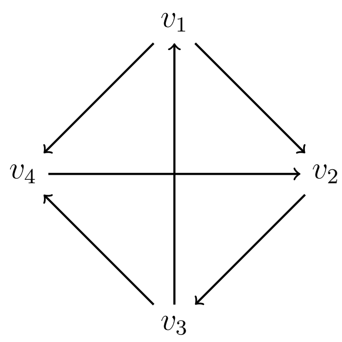


Figure 1.7

Remark 1.1.24. Note that summing the scores of each of the vertices counts the total number of possible edges between vertices (one may think of this as the total number of games played), so that, for a tournament on n vertices, the sum of the scores will be $\binom{n}{2}$. This is evident in Example 1.1.23, as we see that that $2 + 1 + 2 + 1 = 6 = \binom{4}{2}$.

Notation 1.1.25. We denote by $\mathbb{T}(\boldsymbol{s})$ the set of tournaments with score vector $\boldsymbol{s}$.

If T is a tournament, then any subset $U \subseteq V(T)$ induces a subdigraph which is also a tournament. We call this a *subtournament induced by U.*

Notation 1.1.26. If U is a subset of $V(T)$ for a tournament T, then we denote by T_U the subtournament induced by vertex set U.

Example 1.1.27. Figure 1.8 shows a tournament and a subtournament induced by a vertex set in red.

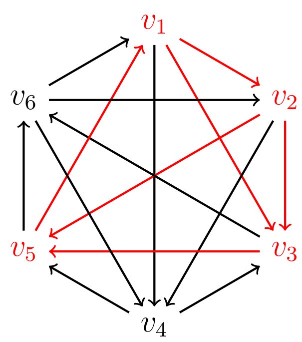


Figure 1.8: A subtournament of a tournament on 6 vertices induced by the vertex set $\{v_1, v_2, v_3, v_5\}$.

Definition 1.1.28. If T is a tournament, and $u \in V(T)$, we denote by $W(u)$ the set of vertices in $V(T)$ that are dominated by u, and $L(u)$ the set of vertices that dominate u. Thus, $\{u\}$, $W(u)$ and $L(u)$ partition $V(T)$.

Example 1.1.29. In Figure 1.8, $W(v_4) = \{v_3, v_5\}$ and $L(v_4) = \{v_1, v_2, v_6\}$.

1.2 Landau's Theorem

An interesting question to ask is which sequences of length n correspond to score sequences of a tournament. Suppose $\boldsymbol{s} = (s_1, s_2, \cdots, s_n)$ is the score sequence of a tournament T. For any subset $A \subseteq V(T)$ of size k let $(s_1^A, \cdots, s_k^A)$ be the score sequence of the subtournament T_A. We then find that, if A is the subset of vertices corresponding to the first k terms of the score sequence,

$$\sum_{i=1}^{k} s_i \geq \sum_{i=1}^{k} s_i^A = \binom{k}{2}.$$

Thus,

$$\sum_{i=1}^{k} s_i \geq \binom{k}{2}$$

for all positive integers $k \leq n$, with equality when $k = n$. In 1953, Landau showed that these conditions are also sufficient [Lan53]:

Theorem 1.2.1 (Landau's Theorem)**.** *A sequence of non-negative integers $s_1 \leq s_2 \leq \cdots \leq s_n$ is a score sequence for some tournament $T \in \mathbb{T}_n$ if and only if*

$$\sum_{i=1}^{k} s_i \geq \binom{k}{2} \tag{1.1}$$

for all positive integers $k < n$, and

$$\sum_{i=1}^{n} s_i = \binom{n}{2}. \tag{1.2}$$

The following clever proof is due to Carsten Thomassen, from ([Cha81], pages 589-591). We first require the following Definition and Lemma:

Definition 1.2.2. If v is a vertex in a tournament T such that v can reach any other vertex via a path of length 1 or 2, v is called a *king.*

Lemma 1.2.3. *In any tournament T the vertex with the highest score is a king.*

Proof. Suppose the vertex with the highest score is v. If $u \neq v$ is a vertex such that u is not reachable from v via a path of length 1, then u dominates v. If u is also not reachable from v via a path of length 2, then for all vertices w such that v dominates w, it is also the case that u dominates w. This implies that the score of u is strictly larger than the score of v, a contradiction. □

Proof of Theorem 1.2.1. We have already seen the necessity of these conditions. For sufficiency, assume the contrary. Then, let N be the smallest integer such that some ordered non-decreasing sequence $\boldsymbol{s} = (s_1, \cdots, s_N)$ satisfies (1.1) and (1.2), yet is not the score sequence for some tournament on N vertices. Suppose further that s_1 is minimal for this value of N.

If there exists $M < N$ such that $\sum_{i=1}^{M} s_i = \binom{M}{2}$, the sequence $\boldsymbol{x} = (s_1, \cdots, s_m)$ satisfies (1.1) and (1.2). Therefore, by the minimality of N, $\boldsymbol{x}$ is the score sequence of a tournament T_1 on M vertices. Moreover,

$$\sum_{i=1}^{w}(s_{M+i} - M) = \sum_{i=M+1}^{M+w}(s_i - M) \geq \binom{M+w}{2} - \binom{M}{2} - Mw = \binom{w}{2}$$

for each w such that $1 \leq w \leq N - M$ and with equality when $w = N - M$. Thus, again by minimality of N, the sequence $\boldsymbol{y} = (s_{M+1} - M, ..., s_N - M)$ is the score sequence of some tournament T_2 on $N - M$ vertices. Now, if we form the tournament on N vertices as a disjoint union of T_1 and T_2, with edges directed from T_2 to T_1, then this is a tournament on N vertices with score sequence $\boldsymbol{s}$, contradicting our assumption.

Now suppose that each inequality in (3.6) is strict when $k < N$. Then, in particular, $s_1 > 0$, and $s_N < (N-1)$. One may easily check that the score sequence $(s_1 - 1, \cdots, s_{N-1}, s_N + 1)$ satisfies (1.1) and (1.2), and thus by minimality of s_1, has a corresponding tournament. By Lemma 1.2.3 v_N is a king, so there is a directed path of length 1 or 2 from v_N to v_1. By reversing the edges of this directed path we obtain a tournament with score sequence $(s_1, \cdots, s_N)$, a contradiction. □

Since any score sequence can be obtained by arranging the scores of a tournament in non-decreasing order, Landau's Theorem provides conditions for $\mathbb{T}(\boldsymbol{s})$ to be non empty. In particular, if $I \subseteq [n]$ and $\boldsymbol{s} = (s_1, \cdots, s_n)$ is any score vector for a tournament T, then the Landau conditions are evidently equivalent to the condition that

$$\sum_{i \in I} s_i \geq \binom{|I|}{2},$$

with equality if $I = [n]$.

Landau's Theorem has an interesting application to ranking in sports. Suppose we have a cricket competition consisting of 10 teams, where each team plays against each other once, and every result is either a win or a loss. Suppose the teams are ranked by their number of wins. How many possible winners are there in this competition? The answer is 9, since the sequence of integers $(0, 5, 5, 5, 5, 5, 5, 5, 5, 5)$ satisfies the conditions of Landau's Theorem (1.1) and (1.2), but no sequence containing ten integers all greater than 4 does.

1.3 Dominance Relations

1.3.1 Relations

Other than reachability (Definition 1.1.11), another relation on the vertices of a tournament is that of *dominance.* This is a total, reflexive anti-symmetric relation corresponding to the orientation of the edge between any pair of vertices.

Many results in the field are related to the problem of ranking vertices based on the structure of a tournament. Under the dominance relation, an unambiguous 'winner' of a tournament would be a vertex that dominates every other vertex. This type of vertex has a special name:

Definition 1.3.1. If v is a vertex in a tournament T such that v dominates every other vertex, we call v a *emperor.*

Evidently, a tournament can have exactly one emperor, but it is also the case that many tournaments have no emperors at all. If we weaken the condition on a 'winner' v so that v can reach every vertex via a path of length at most 2, we get the definition of a king of a tournament (Definition 1.2.2). The following Proposition can be proven in a similar way to Lemma 1.2.3.

Proposition 1.3.2 (Adapted from [FH66], Theorem 3)**.** *If a tournament T has no emperor, then it contains at least three kings.*

Proof. If T has less that 3 vertices, T always has an emperor. Thus, we may assume T has at least 3 vertices. Let v be a vertex of T with maximum score. Then by Lemma 1.2.3, v is a king. Also, since T has no emperor, there is at least one point which dominates v. Among all such points, let u be the one with highest score. A similar argument to the proof of Lemma 1.2.3 shows that this is a king. Now, choose a vertex w with maximum score among those vertices which dominate u. The same argument again shows that this vertex is also a king. Since the dominance relation is asymmetric, and we have that w dominates u and u dominates v, it follows that these vertices are distinct points, and the result follows. □

Examples 1.3.3. Figure 1.9a illustrates an emperor in a tournament, whilst Figure 1.9b shows that Proposition 1.3.2 is tight.

(a) A tournament on 5 vertices with an emperor and its directed edges in purple.

(b) A tournament on 5 vertices with exactly 3 kings (coloured vertices).

Figure 1.9: Two tournaments on 5 vertices: one having an emperor, and one having exactly three kings.

1.4 Directed Paths

A directed path has a natural meaning with regards to tournaments: it represents a sequence of vertices where each vertex dominates the succeeding vertex. The following famous result due to Rédei in 1934 ([Ŕ34]) shows that it is always possible to arrange every vertex in a tournament in such a sequence.

Theorem 1.4.1. *Every tournament contains a Hamiltonian path.*

Proof. We use induction on the number of vertices in the tournament. If T denotes a tournament, the base case $V(T) = 1$ is trivial. If T has n vertices, fix a vertex $v_n \in V(T)$. Then, by induction hypothesis, the subtournament T_U induced by the $U = V(T) - \{v_n\}$ has a Hamiltonian path. Suppose this Hamiltonian path is $P = v_1 e_1 v_2 \cdots v_{n-2} e_{n-2} v_{n-1}$. If v_n dominates v_1 then by adjoining v_n to P it is evident that T has a Hamiltonian path. Otherwise, let i be the minimal value of the index such that v_n is dominated by v_i but dominates v_{i+1}. Then, if e_{out} denotes the edge (v_i, v_n) and e_{in} denoted the edge (v_n, v_{i+1}). Then the path $v_1 e_1 v_2 \cdots v_i e_{\text{out}} v_n e_{\text{in}} v_{i+1} \cdots v_{n-1}$ is a Hamiltonian path in T. □

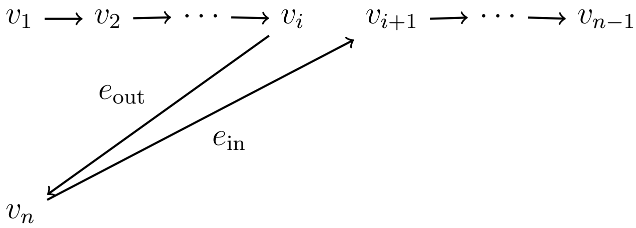


Figure 1.10: The above diagram illustrates the inductive step of the proof of Theorem 1.4.1.

Example 1.4.2. The following picture illustrates a Hamiltonian path in a tournament on 6 vertices.

Figure 1.11: A tournament on 6 vertices (left), and a Hamiltonian path coloured in red (right).

1.5 Transitive Tournaments

Whilst every tournament does contain a Hamiltonian path, it is not always possible to find a total ordering of vertices in a tournament based on the dominance relation, since the tournament may have cycles. However, if the tournament is such that the dominance relation is transitive, it is possible, and such tournaments have a special name:

Definition 1.5.1. A tournament is *transitive* if its dominance relation is transitive.

The following theorem characterises transitive tournaments:

Theorem 1.5.2. *The following are equivalent for any tournament $T \in \mathbb{T}_n$:*

1. *T is transitive.*
2. *The score sequence of T is $(0, 1, \cdots, n-1)$.*
3. *T is acyclic (does not contain a directed cycle).*
4. *T contains a unique Hamiltonian path.*

Proof.

$1 \Rightarrow 2$ Any transitive tournament must contain an emperor. If this emperor is removed, the resulting tournament is also transitive, and the result follows by induction.

$2 \Rightarrow 3$ Any emperor cannot be part of a directed cycle, so any such cycle must be part of the subtournament with the emperor removed and again the result follows by induction.

$3 \Rightarrow 4$ Suppose T is acyclic. If $V(T) = 2$ the result is clear. If $V(T) = n$, then the subtournament induced by the vertex set with $v_n \in V(T)$ removed is also acyclic, and the resulting tournament contains a unique Hamiltonian path. Denote the Hamiltonian path by $v_1 e_1 v_2 \cdots v_{n-2} e_{n-2} v_{n-1}$. Now, reintroduce v_n and the incident edges. If v_n is an emperor, then the result follows immediately. Otherwise, if there are two indices $i < j$ such that v_i dominates v_n which dominates v_{i+1} (and similarly for v_j), then the resulting tournament contains a cycle, as is illustrated in blue in the Figure below.

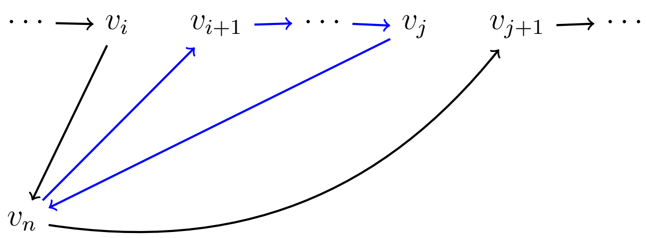


Figure 1.12

It follows that there is exactly one index i satisfying this property, and moreover, v_n dominates every vertex v_j with $j > i$ and $j \leq n-1$.

$4 \Rightarrow 1$ Applying a similar argument to the previous proof, if T contains a Hamiltonian path, denoted by $v_1 e_1 v_2 \cdots v_{n-1} e_{n-1} v_n$, if each v_i does not dominate every succeeding vertex v_j for $j > i$, then this Hamiltonian path is not unique. It follows that T must be transitive.

$\square$

1.6 Irreducible Tournaments and Strong Score Sequences

Another special type of tournament is a *reducible tournament:*

Definition 1.6.1. A tournament is *reducible* if it is possible to partition its vertices into sets A and B such that all the vertices in set A dominate all of those in set B. Otherwise a tournament is *irreducible.*

If T is a tournament with score sequence $(s_1, s_2, \cdots, s_n)$, evidently T is reducible if and only if for some $k < n$

$$\sum_{i=1}^{k} s_i = \binom{k}{2}.$$

Therefore, tournament is irreducible if and only if each inequality in (1.1) is strict for $k < n$.

Theorem 1.6.2. *A tournament is irreducible if and only if it is strong.*

Proof. If a tournament T is strong, evidently it must be irreducible. For the converse, suppose T is not strong. Then, for some $v \in V(T)$, the set A containing v along with all vertices reachable from v is such that $V(T) - A \neq \emptyset$. By partitioning $V(T)$ into A and $V(T) - A$, it follows that T is reducible. □

Theorem 1.6.3 (Adapted from [FH66], Theorem 7)**.** *If a tournament $T \in \mathbb{T}_n$ is strong, it contains a cycle of length $3, 4, \cdots, n$. In particular, it contains a Hamiltonian cycle.*

Proof. A strong tournament $T \in \mathbb{T}_n$ is not transitive, hence must contain a directed triangle. Now, applying induction, suppose T contains a cycle of length $k < n$, which we denote as $v_0 e_0 \cdots v_{k-1} e_{k-1} v_0$. Suppose for some vertex v outside this cycle, the directed edges (v_i, v) and (v, v_{i+1}) are in $E(T)$, where the index values are taken modulo k. Then, by removing the edge (v_i, v_{i+1}) and incorporating these edges, we can form a cycle of length $k + 1$.

Otherwise, every vertex outside the cycle either dominates, or is dominated by the vertices inside the cycle. If we partition the vertices outside the cycle into disjoint sets A and B corresponding to either case, then since T is strong, by Theorem 1.6.2 both A and B are nonempty. Moreover, there must be a vertex $b \in B$ which dominates a vertex $a \in A$. If $e_a = (a, v_0)$, $e_b = (v_{k-2}, b)$ and $e_p = (b, a)$, the cycle $a e_a v_0 e_0 \cdots v_{k-2} e_b \, b e_p a$ is a cycle of length $k + 1$, and this completes the proof. □

1.7 Regular Tournaments

One special type of tournament, which we call a *regular tournament,* has the scores of each of the vertices as evenly distributed as possible. One may think of this as a highly 'competitive' tournament.

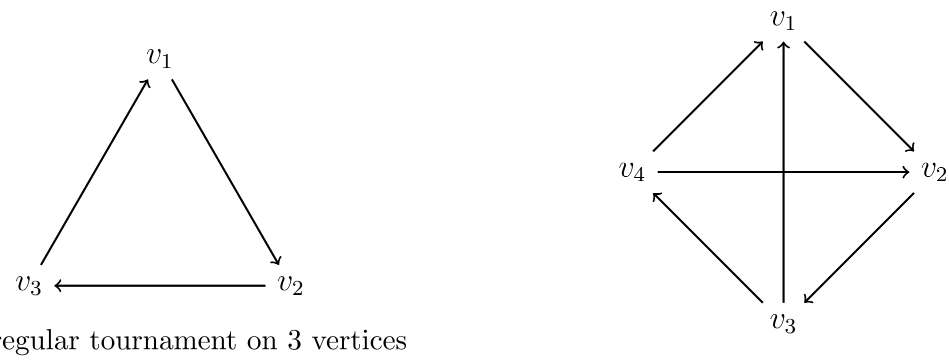
Definition 1.7.1. A *regular tournament* on n vertices, with $n \geq 3$ is a tournament with score sequence

$$\boldsymbol{s} = \begin{cases} \left(\frac{n-1}{2}, \frac{n-1}{2}, \cdots, \frac{n-1}{2}\right), & \text{if } n \text{ is odd,} \\ \left(\frac{n-2}{2}, \frac{n-2}{2}, \cdots, \frac{n-2}{2}, \frac{n}{2}, \cdots, \frac{n}{2}, \frac{n}{2}\right), & \text{if } n \text{ is even.} \end{cases}$$

Note that in the case where n is even exactly $\frac{n}{2}$ vertices have score $\frac{n-2}{2}$.

Remark 1.7.2. If n is even, it is not possible for all of the vertices to have the same score, because this would imply that each vertex has score $\frac{n-1}{2}$, which is not an integer.

Examples 1.7.3. Here are some examples of regular tournaments on even and odd numbers of vertices.



(a) A regular tournament on 3 vertices

(b) A regular tournament on 4 vertices.

Figure 1.13: Regular tournaments on 3 and 4 vertices.

We see that in the case of 3 vertices, each vertex has score 1, whilst in the case of 4 vertices, v_1 and v_2 have score 1 and the other vertices have score 2. This shows that the tournaments are regular.

1.7.1 Kelly's Conjecture

One interesting thing about odd regular tournaments is that it appears that they always can be decomposed into edge disjoint Hamiltonian cycles. This result is known as *Kelly's Conjecture* ([KO13]).

Conjecture 1.7.4. *Any regular tournament on an odd number of vertices can be expressed as the edge disjoint union of Hamiltonian cycles.*

This result, while easy to state, appears to be very difficult to prove. Recently, in 2013, a partial result was obtained by Kühn and Osthus, who proved the result to be true when the number of vertices in the tournament is sufficiently large [KO13].

Theorem 1.7.5 (Kühn and Osthus, 2013 [KO13])**.** *There exists some constant K such that any odd regular tournament on $n \geq K$ vertices has a decomposition into Hamiltonian cycles.*

□

Example 1.7.6. Figure 1.14 shows an example of a decomposition implied by Kelly's Conjecture:

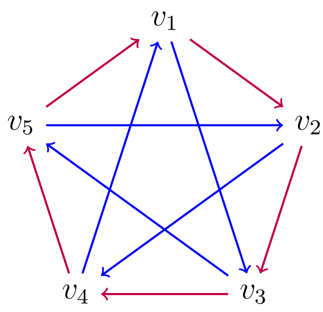


Figure 1.14: A decomposition of a tournament on 5 vertices into edge disjoint Hamiltonian cycles (blue and red).

Kühn and Osthus also proved that when n is even a tournament on n vertices has a decomposition into edge disjoint Hamiltonian paths [KO13].

1.7.2 Excess Sequences

We will show in Chapter 2, that in some sense 'most' tournaments are 'close' to regular (see Theorem 2.4.7). In order to quantify how 'close' a tournament is to regular we need *excesses.*

Definition 1.7.7. Let T be a tournament on n vertices with vertex set V. The *excess* of a vertex $v \in V(T)$ is the out-degree minus the in-degree of v in T.

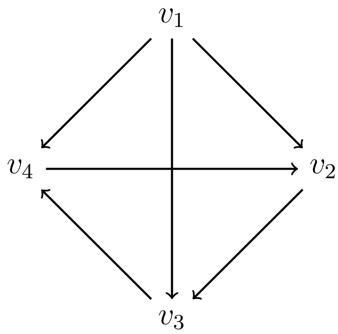
We can then define *excess vectors* and *excess sequences* in a similar way to score vectors and score sequences. Often we will use the symbol $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)$ to represent an excess vector.

Example 1.7.8. If T is a regular tournament, the excess sequence of T is such that

$$\boldsymbol{\delta} = \begin{cases} (0, \cdots, 0), & \textit{if } n \textit{ is odd,} \\ (-1, \cdots, -1, 1, \cdots, 1), & \textit{if } n \textit{ is even.} \end{cases}$$

In some sense, $\|\boldsymbol{\delta}\|_\infty$ gives an indication of how close the tournament is to regular. If $\|\boldsymbol{\delta}\|_\infty$ is large, then at least one vertex has a much larger out-degree than in-degree or much larger in-degree than out-degree, and at least $\|\boldsymbol{\delta}\|_\infty - 1$ directed edges need to be reversed to get a regular tournament.

Example 1.7.9. The following tournament on 4 vertices, has excess vector $\boldsymbol{\delta} = (3, -1, -1, -1)$ under the 'natural' ordering such that $v_i \mapsto i$. The excess sequence of this tournament is $(-1, -1, -1, 3)$.



Remark 1.7.10. Note that summing the scores of the vertices and summing the in-degrees both count the number of possible edges between vertices. Thus, if we have an excess sequence $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)$ of a tournament T

$$\sum_{i=1}^{n} \delta_i = \sum_{v \in V(T)} \deg^+(v) - \sum_{v \in V(T)} \deg^-(v) = \binom{n}{2} - \binom{n}{2} = 0.$$

This is in evident Example 1.7.9 where the sum of the excesses is $3 - 1 - 1 - 1 = 0$.

We will denote by $\mathbb{T}^d(\boldsymbol{\delta})$ the set of all tournaments with excess sequence $\boldsymbol{\delta}$.

1.8 Further References

In this chapter, we have provided an introduction to the theory behind tournaments, and proved some interesting results about them. For a more thorough introduction to the subject we recommend reading the papers 'The Theory of Round Robin Tournaments' by Frank Harary and Leo Moser ([FH66]), and 'The King Chicken Theorems' by Stephen Maurer ([Mau80]).

Chapter 2

Random Tournaments

2.1 Introduction

The purpose of this chapter is to give an introduction to the topic of *random tournaments.* We will then provide a brief survey of the area. Most of the more technical proofs will be omitted, since the approaches tend to vary widely, and are not immediately relevant to the approach taken in this thesis. Results related to paired comparison models and asymptotic enumeration will be considered separately in Chapters 4 and 5. In this chapter, we content ourselves with an overview of the other main results in the field, to provide the reader a broader understanding of the area, and a context for the results proved in Chapters 3 - 5.

We will assume the reader is familiar with basic probability theory. The content of ([Fel68]) should suffice.

2.2 Preliminaries

We define a *random tournament* in a similar way to the way a random graph is defined ([Bol01], page xiii):

Definition 2.2.1. A *random tournament* $\mathscr{T}_n$ on n vertices is a probability distribution over $\mathbb{T}_n$. We will use the notation $T \in \mathscr{T}_n$ to denote a tournament valued random variable T drawn from this distribution.

As this is a probability distribution over a finite set, we can generate a random tournament by specifying the probability $P(T)$ of each tournament $T \in \mathbb{T}_n$, so that the sum over all probabilities is 1.

Examples 2.2.2.

1. The *uniform random tournament* $\mathscr{U}_n$ is the uniform distribution over all tournaments defined on a set of n vertices. Each tournament has probability $\frac{1}{2^{\binom{n}{2}}}$.

2. Given a tournament T, the *Dirac random tournament on T* gives the probability $P(T) = 1$, and all other tournaments probability 0.

A more general random tournament may be generated by having each directed edge occur independently of the others, and with a specified probability:

Definition 2.2.3. The *edge model random tournament* is the random tournament obtained by having each directed edge from vertex v_i to vertex v_j occur independently of the others, with specified probability p_{ij}.

Clearly we must have $p_{ij} + p_{ji} = 1$, since one of the directed edges must occur. We may think of an edge model random tournament as a type of *generalised tournament,* which will be defined and studied in Chapter 3.

Examples 2.2.4.

1. The uniform random tournament is an edge model random tournament where the directed edge from any vertex v_i to any other vertex v_j occurs with probability $\frac{1}{2}$.

2. The Dirac random tournament is such that, if p_{ij} is the probability of an edge directed from vertex v_i to v_j, with $i \neq j$,

$$p_{ij} = \begin{cases} 1, & (v_i, v_j) \in T, \\ 0, & \text{otherwise.} \end{cases}$$

3. The *β-model* random tournament on n vertices, given *strength parameters* $\beta_1, \cdots, \beta_n$ has each directed edge from vertex v_i to vertex v_j, with $i \neq j$, occurring with probability

$$p_{ij} = \frac{e^{\beta_i}}{e^{\beta_i} + e^{\beta_j}}.$$

We will discuss this model in depth in Chapter 4. It is used widely in the *method of paired comparisons* in statistics, under the name *Bradley-Terry model.*

We can define *expected scores* of vertices in a random tournament $\mathscr{T}$ as follows:

Definition 2.2.5. The *expected score* s_i of a vertex v_i in a random tournament $\mathscr{T}_n$ on n vertices is the expected value of the score of a tournament $T \in \mathscr{T}_n$. The *expected score vector* of a random tournament $\mathscr{T}_n$ is the vector $\boldsymbol{s}$ such that the ith component of $\boldsymbol{s}$ is s_i. The *expected score sequence of* $\mathscr{T}_n$ is the sequence obtained by arranging the components of the *expected score vector* in non-decreasing order.

If p_{ij} denotes the probability of a directed edge from vertices v_i to v_j in a tournament $T \in \mathscr{T}_n$, then by linearity of expectation, we have

$$s_i = \sum_{j \neq i} p_{ij}.$$

The results of interest in this chapter are related to the properties a tournament $T \in \mathscr{T}_n$ is 'expected' to have, or those properties that occur with 'large' probability. Often the probability of a tournament $T \in \mathscr{T}_n$ having a particular property P tends to 1 as n goes to infinity. The convention in a lot of literature related to random graphs is to say that property P occurs *almost surely* ([Bol01], pages xiii), but as this is a slight abuse of terminology, we use a different expression:

Definition 2.2.6. Let P be a property of a tournament $T \in \mathscr{T}_n$. We say T has property P *asymptotically almost surely (a.a.s.)* if the probability of P tends to 1 as n tends to infinity. Equivalently, we may say *asymptotically almost all (a. almost all)* tournaments $T \in \mathscr{T}_n$ have property P. If the random tournament is the uniform random tournament $\mathscr{U}_n$, we will omit $\mathscr{U}_n$, and simply say that a. almost all tournaments have property P.

Remark 2.2.7. Another term often used in discrete mathematics and computer science that means exactly the same thing is *with high probability (w.h.p.).*

Examples 2.2.8.

1. Let $\mathscr{T}_n$ be the uniform random tournament, and fix a tournament K. Then a. almost surely, a tournament $T \in \mathscr{T}$ is not equal to K, since

$$P(T = K) = \frac{1}{2^{\binom{n}{2}}} = O\left(2^{-\frac{n(n-1)}{2}}\right).$$

2. Let W denote the event that a tournament $T \in \mathscr{T}$ has a vertex v that dominates every other vertex. Then W does not occur a.a.s, since

$$P(W) = \frac{n}{2^{n-1}} = O\left(2^{-cn}\right)$$

for a constant $c > 0$.

2.2.1 A Note on the 'Probabilistic Method'

One of the most important, and most beautiful, contributions of the theory of random tournaments is the birth of the *probabilistic method.* The probabilistic method is a non-constructive proof technique used widely in combinatorics, number theory and other areas. This technique is attributed to Szele, who, in 1943 [Sze43], used it to prove that there are tournaments which contain a large number of Hamilton paths.

Theorem 2.2.9 (Szele, 1943)**.** *Let $H(T)$ denote the number of Hamiltonian paths in a tournament T on n vertices. Then there exists T such that*

$$H(T) \geq \frac{n!}{2^{n-1}}.$$

Proof. Let $\mathscr{U}_n$ denote the uniform random tournament. In any Hamiltonian path in a tournament $X \in \mathscr{U}_n$, each of the $n-1$ directed edges in the path occurs with probability $\frac{1}{2}$. Thus, the probability of a particular path is $\frac{1}{2^{n-1}}$. As there are $n!$ possible Hamiltonian paths (corresponding to the $n!$ arrangements of the vertices), by linearity of expectation

$$\mathbb{E}\, H(X) = \frac{n!}{2^{n-1}}.$$

But as

$$\mathbb{E}\, H(X) = \sum_{K \in \mathbb{T}_n} P(K) H(K) \leq \max_{K \in \mathbb{T}_n} H(K),$$

it follows that some tournament T has $H(T) \geq \frac{n!}{2^{n-1}}$. □

In this way, the probabilistic method can be used to convert results about random tournaments into existence results for tournaments. For more about the probabilistic method, the author strongly recommends reading [**?**].

2.3 The Regular Random Tournament

In 1974, in [Spe74], Joel Spencer proved a number of results related to the properties a ‘random’ regular tournament is likely to have. The motivation behind this paper has been to develop an approach to attack problems specifically related to regular tournaments, for example Kelly’s Conjecture (Conjecture 1.7.4).

To make this more precise, we define the *regular random tournament* to be the probability distribution which is uniform on all regular tournaments and zero otherwise. Let RT_n denote the set of regular tournaments on a set of n vertices $\{v_1, \cdots, v_n\}$.

Definition 2.3.1. The *regular random tournament* $\mathscr{R}_n$ is the random tournament on n vertices such that, if $T \in \mathscr{R}_n$ then

$$P(T) = \begin{cases} 0, & T \notin RT_n, \\ \frac{1}{|RT_n|}, & \text{otherwise.} \end{cases}$$

Remark 2.3.2. We will find an explicit asymptotic formula for $|RT_n|$ in Chapter 5.

The following Theorem provides upper and lower bounds for the probability that the set of games played by vertices in a subset (with each score equal to a fixed value m) is a particular set. For $T \in \mathscr{R}_n$ let S_m be the set of all subsets A such that the score of each $v_i \in A$ in T is exactly m. Also given such a set A, denote by G_A the set of games played by vertices in A, and set

$$\mathscr{S}(A) = \{G_A : A \in S_m\}.$$

Theorem 2.3.3 (Spencer, 1974 [Spe74]). *Suppose* $|A| = k$, *with* $k \leq n^{0.6}$ *and* $g_1 \in \mathscr{S}(A)$. *Then for sufficiently small* $\varepsilon > 0$ *and* n *sufficiently large,*

$$|\mathscr{S}(A)|^{-1} e^{-\frac{k^2}{2-\varepsilon}} \leq P_{\mathscr{R}_n}(G_A = g_1 | T \in RT_n) \leq |\mathscr{S}(A)|^{-1} e^{\frac{k^2}{2-\varepsilon}}.$$

□

The next Theorem, which is similar to the first, provides upper and lower bounds on the probability that the subtournament induced by a vertex set A, where each vertex has the same score m, is a particular tournament. With A as defined previously, denote by $\mathbb{T}_A$ the set of all tournaments on vertex set A.

Theorem 2.3.4 (Spencer, 1974 [Spe74]). *With* A *as defined previously, suppose* $|A| = k$, *with* $k \leq n^{0.6}$. *Let* $T_1 \in \mathbb{T}_A$. *Then for sufficiently small* $\varepsilon > 0$ *and* n *sufficiently large*

$$|\mathbb{T}_A|^{-1} e^{-(2+\varepsilon)\frac{k^3}{n}} \leq P_{\mathscr{R}_n}(T_A = T_1 | T \in RT_n) \leq |\mathbb{T}_A|^{-1} e^{(2+\varepsilon)\frac{k^3}{n}},$$

where T_A *is the subtournament induced by* A.

□

Spencer also proved the following:

Theorem 2.3.5 (Spencer, 1974 [Spe74]). *Set* $t = [3\log_2 n]$. *Then, for a. almost all* $T \in \mathscr{T}_n$, *there do not exist subsets* A, $B \subseteq \{v_1, \cdots, v_n\}$ *with* $A \cap B = \emptyset$, $|A| = |B| = t$ *such that all of the vertices in* A *dominates all of those in* B *(or vice-versa).*

□

In some sense this result makes precise how 'competitive' a regular tournament is likely to be: for n large it is very unlikely that that you can split a regular tournament into disjoint subsets of size $3\log_2 n$ such that the vertices of one set dominate the other.

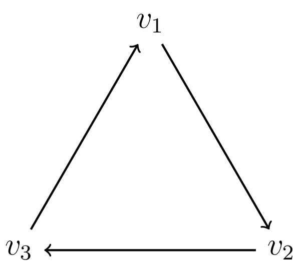
2.4 The Uniform Random Tournament

More generally, results related to the uniform random tournament allow one to make conclusions about the properties any 'random' tournament is likely to have. Thus, a number of results in the literature on random tournaments concern the uniform random tournament. In this section, we provide an overview of the area.

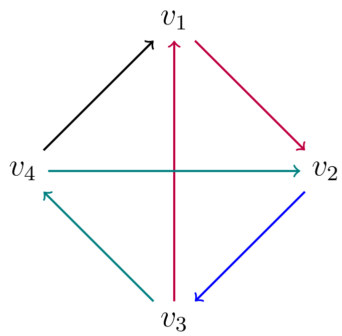
2.4.1 Subgraph Counts

If a tournament is drawn at random from a uniform distribution, what is the probability of it containing a particular subdigraph? In how many ways can one fit such a subdigraph on the tournament so that the directed edges overlap?

Example 2.4.1. If the tournament in Figure 3.5 was drawn from $\mathscr{U}_4$, it would contain two copies of the directed cycle in Figure 3.3 (coloured).



(a) A directed triangle



(b) A regular tournament on 4 vertices.

Figure 2.1: The tournament on the right contains two copies of the directed cycle on the left. One triangle is coloured in teal and the other in purple, with the edge common to both coloured in blue (the edge (v_2, v_3)).

Evidently, since a tournament is an oriented graph, it can only contain copies of oriented graphs. In this subsection we review some general results related to subdigraphs in a uniform random tournament.

Three Cycles and Four Cycles

Denote by $NC_3(T)$ and $NC_4(T)$ the number of labeled directed 3 and 4 cycles respectively in a tournament T. For example, if T_1 is the tournament defined by Figure 3.5, then $NC_3(T_1) = 2$. If we have a uniform random tournament $T \in \mathscr{U}_n$, then given three vertices v_1, v_2, v_3, the probability of a cycle containing these three vertices is $\frac{1}{4}$. Indeed, there are 2 different cycles containing these vertices (illustrated below), and each of these cycles has probability $\frac{1}{8}$.

Figure 2.2: The two possible cycles on vertex set $\{v_1, v_2, v_3\}$.

By linearity of expectation, if $T \in \mathscr{U}_n$ it follows that

$$\mathbb{E}\, NC_3(T) = \frac{1}{4}\binom{n}{3}.$$

Similarly,

$$\mathbb{E}\, NC_4(T) = \frac{3}{8}\binom{n}{4}.$$

A result by (Moran, 1947 [Mor47]) shows that the distribution of both of these are asymptotically normal.

Theorem 2.4.2 (Moran, 1947 [Mor47]). *Let $T \in \mathscr{U}_n$. Then, as n tends to infinity, the distributions of*

$$\frac{NC_3(T) - \frac{1}{4}\binom{n}{3}}{\sqrt{\frac{3}{16}\binom{n}{3}}},$$

and

$$\frac{NC_4(T) - \frac{3}{8}\binom{n}{4}}{\sqrt{\frac{3}{64}\binom{n}{4}(4n-11)}},$$

converge to the standard normal distribution.

□

2.4.2 Hamiltonian Cycles

Recall that a Hamiltonian Cycle in a directed graph is a spanning cycle. Let $H(T)$ denote the number of Hamiltonian cycles in a tournament T. Then, a similar argument to the case of finding $\mathbb{E}\, NC_3(T)$ shows that

$$\mathbb{E}\, H(T) = \frac{(n-1)!}{2^n}.$$

In 1995, Svante Janson [Jan95] showed that $H(T)$ is also asymptotically normally distributed:

Theorem 2.4.3 (Janson, 1995 [Jan95]). *Let $T \in \mathscr{U}_n$. Then as*

$$\operatorname{Var} H(T) = \left(\frac{2}{n} + O(n^{-2})\right)(\mathbb{E}\, H(T))^2,$$

and as $n \to \infty$

$$\frac{H(T) - \mathbb{E}\, H(T)}{\sqrt{\operatorname{Var} H(T)}}$$

converges in distribution to the standard normal distribution.

□

By computing the conditional expectation of $H(T)$ in $\mathscr{U}_n$ with respect to various types of tournament, Wormald managed to find new asymptotic lower bounds on the maximum number of Hamiltonian cycles in a tournament [Wornt].

Paths of Length Two

Let r_{ij} denote the number of directed paths of length 2 from vertices v_i to v_j in a tournament $T \in \mathscr{U}_n$. Let R_λ denote the expected number of ordered pairs of distinct vertices v_i and v_j such that

$$\left| r_{ij} - \frac{n-2}{4} \right| > \lambda.$$

The following theorem comes from J W Moon ([Moo68], page 42-43):

Theorem 2.4.4 (Moon, 1968 [Moo68]). *If $w(n)$ is such that $\log(n-2) + w(n) \to \infty$ and $w(n)n^{-\frac{1}{3}}$ as $n \to \infty$, and*

$$\lambda = \left(\frac{3}{4}(n-2)\left(\log n - 2 + w(n)\right) \right)^{\frac{1}{2}},$$

then asymptotically,

$$R_\lambda \sim \left(2\pi \left(\log n - 2 + w(n)\right)\right)^{-\frac{1}{2}} e^{-2w(n)}.$$

□

Tournaments with a Given Excess Sequence

If we know that a tournament drawn from $\mathscr{U}$ has a given score vector $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)$, what is the probability that it contains a copy of a specified oriented graph D? In an enumerative result by Gao, Mckay and Wang in 1998 [GMW00], under certain circumstances, the asymptotic probability has been computed.

Let $\deg_i S(G)$ denote the degree of vertex v_i in the underlying simple graph $S(D)$. Moreover, define $\delta_i(D)$ to be the outdegree of vertex v_i in D minus the indegree, and fix an excess sequence $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)$.

Also, given an oriented graph D and an excess sequence $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)$, set

$$\beta_1 = \frac{1}{2n} \sum_{1 \le j \le n} \left(2\delta_j \delta_j(D) - \delta_j^2(D)\right) + \frac{1}{3n^3} \sum_{1 \le j \le n} \delta_j^3 \delta_j(D) + \frac{1}{n^4} \sum_{1 \le j \le n} \delta_j^2 \sum_{1 \le j \le n} \delta_j \delta_j(D),$$

and

$$\beta_2 = -\frac{1}{2n^2} \sum_{(j,k) \in D} \left(\delta_j - \delta_k - \delta_j(D) + \delta_k(D)\right)^2.$$

Then, they proved the following result:

Theorem 2.4.5 (Gao, McKay, Wang, 1998 [GMW00]). *Let D be an oriented graph containing m directed edges such that for each $i \in [n]$, $\deg_i S(D) = O\left(n^{-\frac{1}{2}-\varepsilon'}\right)$ where ε' is any positive constant. Moreover, let $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)$ be an excess vector with each $\delta_i = o\left(n^{\frac{2}{3}}\right)$, and such that $\delta_i \deg_j S(D) = o(n)$ for all $i, j \in [n]$. Then, for $T \in \mathscr{U}_n$ and β_1 and β_2 as defined above*

$$P\left(D \subseteq T | \, \boldsymbol{\delta}\right) \sim 2^{-m} \exp\left(\frac{m}{n} + \beta_1 + \beta_2\right)$$

uniformly as $n \to \infty$.

□

2.4.3 Discrepancy Results

In the previous subsection we reviewed a number of results regarding random tournaments containing oriented graphs. Informally, many of these results related to drawing a fixed oriented graph on the vertex set of a tournament so that the edges overlapped. In this subsection we review results of a similar flavour: regarding *discrepancy*.

If T_1 and T_2 are tournaments defined on sets of n vertices, roughly speaking, the *discrepancy* between T_1 and T_2 measures how much their edges can be made to overlap if T_1 is drawn on the vertex set $V(T_2)$. More precisely, the *positive discrepancy* of these tournaments T_1 and T_2 is defined by

$$\operatorname{disc}^+(T_1, T_2) := \max_{\phi} |E(\phi(T_1)) \cap E(T_2)| - \frac{1}{2}\binom{n}{2}$$

and the *negative discrepancy* is defined by

$$\operatorname{disc}^-(T_1, T_2) := \frac{1}{2}\binom{n}{2} - \min_{\phi} |E(\phi(T_1)) \cap E(T_2)|$$

where the maximum and minimum are taken over all bijections ϕ from the vertex set of T_1 to the vertex set of T_2. The reason for the $\frac{1}{2}\binom{n}{2}$ term in these definitions is that any two tournaments can be drawn on the same vertex set so that at least $\frac{1}{2}\binom{n}{2}$ edges overlap.

Indeed, if T_1 is drawn randomly on the same set of vertices as T_2, then the probability of a particular edge between a pair of vertices v_i and v_j overlapping is $\frac{1}{2}$. By linearity of expectation, the expected number of common edges is $\frac{1}{2}\binom{n}{2}$, so that there exists at least one drawing of T_1 on the vertex set of T_2 satisfying the required property.

The (unsigned) *discrepancy* is defined as

$$\operatorname{disc}(T_1, T_2) = \max\{\operatorname{disc}^+(T_1, T_2), \operatorname{disc}^-(T_1, T_2)\}.$$

Let TR_n denote any transitive tournament on n vertices. Then, in ([Spe71], 1971), and ([Spe80], 1980), Spencer showed that a. almost all tournaments $T \in \mathscr{U}_n$ are such that

$$\operatorname{disc}^+(T, TR_n) = \Theta\left(n^{\frac{3}{2}}\right).$$

The upper bound in this result was sharpened by Vega in 1983, who showed that a. almost surely

$$\operatorname{disc}^+(T, TR_n) \leq 1.73n^{\frac{3}{2}}.$$

In 2015, Bollabás and Scott [BS15] considered the (unsigned) discrepancy between two tournaments drawn from $\mathscr{U}_n$. They showed that if $T_1, T_2 \in \mathscr{U}_n$, then a. almost surely

$$\operatorname{disc}(T_1, T_2) = \Theta\left(n^{\frac{3}{2}}\sqrt{\log n}\right).$$

This provided an asymptotic answer to the question "how much more can two random tournaments be made to agree or disagree?"

2.4.4 A Surprising Correlation

Let $\mathscr{U}_n$ be the uniform random tournament on vertices $V = \{v_1, \cdots, v_n\}$. Let a, s and b be three vertices in V, and denote by $\{a \to s\}$ the event that there is a directed path from a to s, and similarly define $\{s \to b\}$.

Recall that two events X_1 and X_2 are said to be negatively correlated if

$$P(X_1 \wedge X_2) < P(X_1)P(X_2),$$

and positively correlated if

$$P(X_1 \wedge X_2) > P(X_1)P(X_2).$$

One might expect the events $\{a \to s\}$ and $\{s \to b\}$ to be negatively correlated. Indeed, increasing the probability of the event $\{a \to s\}$ might involve increasing the number of edges directed towards s which might decrease the probability of the event $\{s \to b\}$. However, a result by Sven Erick Alm and Svante Linussen in 2009 [EL09] shows that, somewhat counter-intuitively this is not the case if the number of vertices $n \geq 4$:

Theorem 2.4.6 (Alm and Linussen, 2009 [EL09]). *Let $\mathscr{U}_n$ be the uniform random tournament on n vertices, with $n \geq 3$, and the events $\{a \to s\}$ and $\{s \to b\}$ as defined above. Then, these two events are negatively correlated if $n = 3$, independent if $n = 4$, and positively correlated otherwise.*

□

This rather interesting result has also been studied in the case of random orientations of arbitrary graphs G, where the orientation of each edge is independent, and each orientation has equal probability.

2.4.5 Dominating Sets of Vertices

Since the universal random tournament can be expressed as an edge model random tournament with each edge having probability $\frac{1}{2}$, one might expect any tournament drawn from this distribution to be 'competitive' in the sense that few vertices dominate a large number of other vertices. This is indeed the case.

In 1962, Moon and Moser showed that a. almost all tournaments $T \in \mathscr{U}_n$ are irreducible. In fact, they showed ([Moo62], [Moo68] pages 3-5) that if I denotes the event that a tournament $T \in \mathscr{U}_n$ is irreducible, then

$$\left| P(I) - \frac{n}{2^{n-2}} \right| < \frac{1}{2} \left(\frac{n}{2^{n-2}} \right)^2.$$

More generally, if $\boldsymbol{\delta}$ denotes the excess sequence of a tournament T, a. almost all tournaments T are such that $\boldsymbol{\delta} \leq O\left(n^{\frac{1}{2}+\varepsilon}\right)$ for any $\varepsilon > 0$. This result has been alluded to in [Spe74], but does not appear to have been explicitly proven, hence we provide a proof here.

Theorem 2.4.7. *Suppose $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)$ is the excess sequence of a tournament $T \in \mathscr{U}_n$. Then, a. almost surely, $\|\boldsymbol{\delta}\|_\infty \leq O\left(n^{\frac{1}{2}+\varepsilon}\right)$.*

Proof. In $\mathscr{U}_n$, we may think of the score of any vertex v_i as the sum of $n-1$ i.i.d Bernoulli random variables, with success probability $\frac{1}{2}$. Now, for any vertex v_i and constant $c > 0$, the probability that

$$P\left(|\delta_i| > cn^{\frac{1}{2}+\varepsilon}\right)) = P\left(|2s_i - n - 1| > cn^{\frac{1}{2}+\varepsilon}\right).$$

By Hoeffding's inequality applied to s_i, since $\mathbb{E}\, s_i = \frac{n-1}{2}$, we have

$$P\left(\left|s_i - \frac{n-1}{2}\right| > \frac{c}{2} n^{\frac{1}{2}+\varepsilon}\right) \leq 2\exp\left(-\frac{c}{2} n^{\varepsilon}\right),$$

so that the probability of any vertex having excess larger that $cn^{\frac{1}{2}+\varepsilon}$ is bounded above by

$$2n \exp\left(-\frac{c}{2}\varepsilon\right)$$

which tends to 0 as n goes to ∞. □

Unbeaten k-sets

If a set $A \subseteq V$ is such that $|A| = k$, and at least one vertex in A dominates a vertex outside A, A is said to be an *unbeaten k-set* . In (Barbour et al. 1993 [BGQ97]), Barbour, Godbole and Qian have managed to show that the number of $k-sets$ in a random tournament $T \in \mathscr{U}_n$ is well approximated (with respect to a probability metric) by a Poisson Distribution with mean

$$\binom{n}{k} \left(1 - 2^{-k}\right)^{n-k}.$$

2.4.6 Quasi-Random Results

A class of properties of tournaments is said to be *quasi-random* if $T \in \mathscr{U}_n$ (or any random tournament $\mathscr{T}_n$) satisfies one of them, then a.a.s T satisfies all of them. These properties are particularly significant in relation to the uniform random tournament because it means that if n is large, then in order to verify that a tournament T satisfies a variety of properties, with a high probability it is enough to verify that one property holds.

A quasi-random class of properties have been studied in a paper by Fan Chung and Ronald Graham [Chu91]. In order to state their result we will need to introduce some terminology.

Suppose T is a tournament on vertex set V with $|V|= n$. Let $\binom{T}{T'}$ denote the number of copies of a digraph T' in T.

For $u, v \in V$, we denote by

$$\text{same}(u, v) := |(W(u) \cap W(v)) \cup (L(u) \cap L(v))| \,.$$

Intuitively, this denotes the set of vertices w whose encounter with u and v is the same (ie w dominates u and w dominates v or w is dominated by u and w is dominated by v).

Finally, if T has excess vector $\boldsymbol{\delta} = (\delta_1, ..., \delta_n)$, we say T is *almost balanced* if

$$\sum_{i=1}^{n} |\delta_i|= o(n^2).$$

We are now able to state the main result of (Chung and Graham, 1991 [Chu91]):

Theorem 2.4.8 (Chung and Graham, 1991 [Chu91]). *Let $\mathscr{U}_n$ be the uniform random tournament on vertex set V, and let $T \in \mathscr{U}_n$. The following form a quasi-random class of properties of T:*

1. *If T' is a tournament on s vertices, $s \leq n$*
$$\binom{T}{T'} = (1 + o(1))\, n^s 2^{-\binom{s}{2}}.$$
2. $NC_4(T) = (1 + o(1))\, \frac{n^4}{2}$.
3. $\sum_{u,v \in V} \left|\text{same}\,(u, v) - \frac{n}{2}\right| = o(n^3)$.
4. $\sum_{u,v \in V} \left||W(u) \cap W(v)| - \frac{n}{4}\right| = o(n^3)$.
5. *For all subsets $X \subset V$, the subtournament T_X is almost-balanced.*
6. *Every subtournament of T on $\lfloor \frac{n}{2} \rfloor$ vertices is almost balanced.*

7. *For every partition of* $V = A \cup B$ *with* $|A| = \lfloor \frac{n}{2} \rfloor$ *and* $|B| = \lceil \frac{n}{2} \rceil$*, we have*

$$\sum_{v \in A} ||W(v) \cap B| - |L(v) \cap B|| = o(n^2).$$

8. *For all* $A, B \subset V$*,*

$$\sum_{v \in A} ||W(v) \cap B| - |L(v) \cap B|| = o(n^2).$$

9. *For every ordering* π *of* T*,*

$$|\{(u,v) \in E(T) : \pi(u) < \pi(v)\}| = (1 + o(1))\,\frac{n^2}{4}.$$

□

2.5 The p Model Random Tournament

A few results have been proved for a more general type of random tournament, which we call the *p-model* random tournament.

Definition 2.5.1. The *p-model* random tournament $\mathscr{T}_n(p)$ on the ordered vertex set $[n]$ is a type of edge model random tournament where the probability of a directed edge from vertex i to vertex j is p if $i < j$, and $1 - p$ otherwise.

To the author's knowledge, this type of random tournament was first studied by (Frank, 1968 [Fra68]), who showed that for $T \in \mathscr{T}_n(p)$,

$$\mathbb{E}\, NC_3(T) = \binom{n}{3} p(1-p)$$

and

$$\operatorname{Var} NC_3(T) = \binom{n}{3} p(1-p)\left(1 - p(1-p) + \frac{n-3}{2}\right).$$

Note that this is consistent with the case $p = \frac{1}{2}$ in (2.4.1).

In 1996, (Luczaks et al. [RG96]) proved a number of asymptotic results related to this random tournament, when p can vary as a function of n.

For a graph $G = (V(G), E(G))$, define $d(G)$ to be the ratio

$$d(G) = \frac{|E(G)|}{|V(G)|},$$

and set $m(G) = \max_{H \subseteq G} d(H)$. Also, recall that if D is an oriented graph, then $S(D)$ denotes the underlying simple graph on the same vertex set. Then we have the following threshold result:

Theorem 2.5.2 (Luczacs et al. 1996 [RG96])**.** *Let $T \in \mathscr{T}_n(p)$ and let D be a oriented graph with at least one cycle. Then,*

$$\lim_{n\to\infty} P(D \subseteq T) = \begin{cases} 0, & \textit{if } np^{m(S(D))} \to 0, \\ 1, & \textit{if } n\left(\min\{p, 1-p\}\right)^{m(S(D))} \to \infty, \\ 0, & \textit{if } n(1-p)^{m(S(D))} \to 0. \end{cases}$$

□

Now, if $T \in \mathscr{T}_n(p)$, define

$$\alpha := \min\left\{i : (i,j) \in T \text{ for some } j > i\right\},$$

and

$$\beta := \max\left\{j : (i,j) \in T \text{ for some } i < j\right\}.$$

In the same paper, (Luczacs et al.) showed the following:

Theorem 2.5.3 (Luczacs et al. 1996 [RG96])**.** *Let $T \in \mathscr{T}_n(p)$. If $n^2p \to \infty$ and $p \leq \frac{1}{2}$, a. almost surely the set $\{\alpha, \alpha+1, \cdots, \beta\}$ is the vertex set of a strong component of T.*

□

Note that this result can be extended to random tournaments with $p \geq \frac{1}{2}$ by the duality between $\mathscr{T}_n(p)$ and $\mathscr{T}_n(1-p)$.

Chapter 3

More General Types of Tournament

3.1 Introduction and Literature

In this chapter we will briefly describe some more general types of tournaments. There are three types of classes we will consider: *generalised tournaments, multi-tournaments* and *generalised multi-tournaments.*

It is difficult to find literature on these more general classes of tournaments. The class of *generalised tournaments* was, to the best of our knowledge, first introduced by Moon [Moo63]. In [Jec83], Jech introduced the types of tournament we will call *generalised multi-tournaments,* motivated by applications to comparison based ranking.

The primary motivation behind this chapter is to study these types of tournaments in the context of polytopes. As a result much of this chapter will assume basic knowledge of convex analysis, which may be found in any introductory text on on the subject (for example [HU01]). As generalised tournaments are closely related to random tournaments, the results proved in this chapter will be useful in Chapters 4 and 5. On the way we provide proofs of some Landau-like theorems for these types of tournaments. In particular, Theorem 3.3.12 shows that for any score vector $\boldsymbol{s}$ there exists a *generalised tournament* with this score vector that can be expressed as the convex combination of transitive tournaments. We will explain this more detail later on, but this has an interesting interpretation in terms of voting theory (Subsection 3.3.4).

3.1.1 A Word on Notation and Terminology

In this chapter we will use terminology and notation related to convex analysis. We will assume the reader has some familiarity with these terms.

Given a set $S \subseteq \mathbb{R}^n$ recall that the *affine hull* of S is the set of affine combinations of vectors in S.

Notation 3.1.1 (Affine Hull)**.** We denote by $\text{aff}(S)$ the *affine hull* of S.

Also recall that convex hull of a set is defined similarly, in terms of convex combinations rather than affine combinations.

We use the standard notation for open balls:

Notation 3.1.2 (Ball)**.** Given a point $p \in \mathbb{R}^n$ and $r \in \mathbb{R}_+$, we denote by $\mathrm{B}_r(p)$ the open ball of radius r centred at p.

Finally, recall that the *relative interior* of a set in $\mathbb{R}^n$ is the interior of the set when viewed as a subset of its affine hull. Specifically, the relative interior of a set $S \in \mathbb{R}^n$ is the set

$$\{\boldsymbol{x} \in S : \exists \varepsilon > 0 \text{ such that } \mathrm{B}_\varepsilon(x) \cap \text{aff}(S) \subseteq S\}$$

We then have the following notation:

Notation 3.1.3 (Relative Interior)**.** Given a set $S \in \mathbb{R}^n$, we denote by $\text{rint}\,(S)$ the *relative interior* of S.

3.2 Multi-Tournaments

A tournament can be thought of as encoding the results of a round-robin competition, where every pair of individuals competes against each other once. A multi-tournament allows the possibility of multiple encounters, as is the case with many sporting leagues around the world.

Let M be an $n \times n$ symmetric matrix with each entry $m_{ij} \in \mathbb{Z}$, and

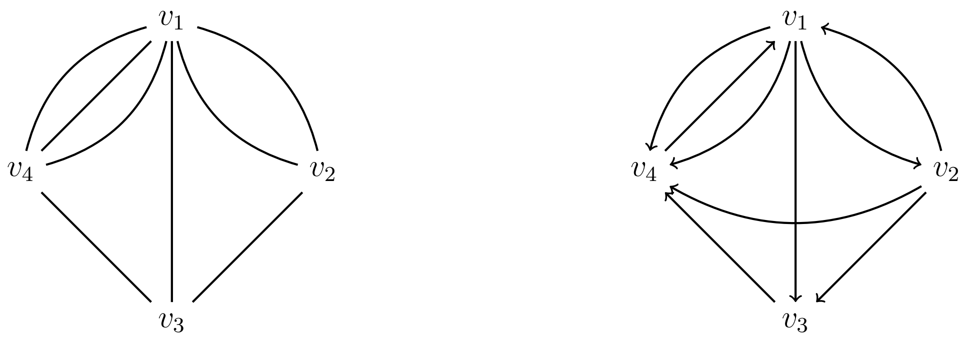
$$m_{ij} \begin{cases} = 0, & i = j, \\ > 0, & i \neq j. \end{cases} \tag{3.1}$$

Definition 3.2.1. A *multi-tournament* T on a labelled vertex set $V = \{v_1, \cdots, v_n\}$ with *schedule matrix* M is an orientation of the multi-graph on the vertex set V having m_{ij} edges between vertices v_i and v_j.

Remark 3.2.2. A multi-tournament is hence a type of *multidigraph,* or a *quiver.*

Example 3.2.3. Figure 3.1b is an example of a multi-tournament with schedule matrix

$$M = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 3 & 1 & 1 & 0 \end{bmatrix}.$$



(a) A multi-graph corresponding to the schedule matrix M.

(b) A multi-tournament with schedule matrix M.

Figure 3.1

We define the *score* of each vertex, the *score vector,* and *score sequence* in a similar way to usual tournaments so that, if $\boldsymbol{s}$ is the score sequence of a multi-tournament, then the following Landau-like conditions are satisfied:

$$\sum_{i=1}^{k} s_i \geq \sum_{i<j,\, i,j\in[k]} m_{ij}, \tag{3.2}$$

for each $k \in [n]$, with equality if $k = n$.

Example 3.2.4. The score vector $\boldsymbol{s}$ of the multi-tournament in Figure 3.1b is $(4, 3, 1, 1)$. The score sequence is $(1, 1, 3, 4)$ and one may readily check that conditions (3.2) are satisfied.

One may immediately wonder whether these Landau-like conditions are also sufficient. The general case appears to be difficult, but a very similar proof to Theorem 1.2.1 in Chapter 1 shows that there is a Landau-like theorem in the particular case that each m_{ij} is equal to some constant p.

Theorem 3.2.5. *Let M be the schedule matrix for a multi-tournament T where each $m_{ij} = p$ for $i \neq j$. A sequence of non-negative integers $s_1 \leq s_2 \leq ... \leq s_n$ is a score sequence for T if and only if*

$$\sum_{i=1}^{k} s_i \geq p\binom{k}{2} \tag{3.3}$$

for all positive integers $k < n$, and

$$\sum_{i=1}^{n} s_i = p\binom{n}{2}. \tag{3.4}$$

Proof. The proof follows a similar argument to Theorem 1.2.1 in Chapter 1. □

3.2.1 Generalised Multi-Tournaments

In *generalised multi-tournaments* every game between participants i and j results in each player given a rating corresponding to their performance. The sum of the ratings of both players is always 1. For example, if the result was a draw, both players i and j might get $\frac{1}{2}$.

Definition 3.2.6. A generalised multi-tournament G with schedule matrix M on n vertices is an ordered pair (V, T). Here V is a finite set of vertices $\{v_1, ..., v_n\}$ and T is a collection of non-negative real number *edges* $\{t_{ij}^q : i, j \in [n], q \in [m_{ij}]\}$ such that $t_{ij}^q + t_{ji}^q = 1$, $t_{jj}^q = 0$ and $0 \leq t_{ij}^q \leq 1$. If we set $K = 2\sum_{1\leq i<j\leq n} m_{ij} + n$, we may think of the edge set as a vector $\boldsymbol{t} \in \mathbb{R}^K$ obtained by adjoining the edge values.

Example 3.2.7. We may think of a normal multi-tournament as a generalised multi-tournament, where each t_{ij}^q is either 0 or 1, corresponding to whether i dominates j or j dominates i.

Definition 3.2.8. Let $G = (V, T)$ be a generalised multi-tournament. Given a *vertex* $v_i \in V$, the *score* s_i of v_i is the sum

$$\sum_{j=1}^{n} \sum_{q=1}^{m_{ij}} t_{ij}^q.$$

We then define *score sequences* and *score vectors* in a similar way to other tournaments. It is possible to prove the following generalisation of Landau's Theorem for generalised multi-tournaments:

Theorem 3.2.9. *Let M be the schedule matrix for a generalised multi-tournament G. A sequence of non-negative real numbers $s_1 \leq s_2 \leq ... \leq s_n$ is a score sequence for G if and only if*

$$\sum_{i=1}^{k} s_i \geq \sum_{i<j,\, i,j\in[k]} m_{ij},$$

for all positive integers $k \leq n$, with equality if $k = n$.

Proof. The proof of necessity of these conditions is straightforward. For ease of notation, we will only prove sufficiency in the specific case where each $m_{ij} = 1$, in Theorem 3.3.12, however the main idea is as follows:

1. Construct polytopes P_M (depending on the schedule matrix M) corresponding to the necessary conditions a score vector must satisfy (similar to Definition 3.3.3).

2. Compute the vertices of this polytope using a similar argument to Proposition 3.3.7.

3. Observe that these vertices correspond to the score vector of a specific type of 'transitive' generalised multi-tournament. Then, using the fact that a bounded polytope is the convex hull of its vertices, contruct a generalised multi-tournament with the required score vector as a convex combination of these 'transitive' generalised multi-tournaments.

□

3.3 Some Polytopes

In this section we outline an approach to study generalised multi-tournaments using polytopes. For brevity, we only study the particular case of a generalised multi-tournament where each $m_{ij} = 1$ for $i \neq j$, but the results for the other cases are similar.

Let Q be the symmetric $n \times n$ matrix such that $Q_{ij} = 0$ if $i = j$ and 1 otherwise.

Definition 3.3.1. A *generalised tournament* is a generalised multi-tournament with schedule matrix Q.

We may think of a generalised tournament as the combinatorial structure underlying an edge model random tournament. This type of tournament G may be represented as a diagram of with directed edges from vertices v_i to v_j labelled with the value of the edge t_{ij}. By convention, we will only draw edges having values greater than or equal to $\frac{1}{2}$, and assume any unlabelled directed edge has value 1.

Example 3.3.2. The following diagram illustrates a generalised tournament on 4 vertices:

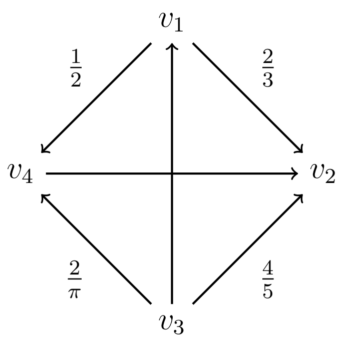


Figure 3.2: A generalised tournament on 4 vertices. This tournament has score vector $\left(\frac{7}{6}, \frac{8}{15}, \frac{9}{5} + \frac{2}{\pi}, \frac{5}{2} - \frac{2}{\pi}\right)$.

The Landau-like conditions for generalised tournaments imply that the set of 'permissible' score vectors of a generalised tournament are those satisying a set of linear inequality and equality constraints. Indeed, if $\boldsymbol{s} \in \mathbb{R}^n$ is a score vector then

1. $s_i \geq 0 \quad \forall i \in [n]$,
2. $\forall I \subsetneq [n] \quad \sum_{i \in I} s_i \geq \binom{|I|}{2}$,
3. $\sum_{i=1}^{n} s_i = \binom{n}{2}$.

We may thus think of the set of all score vectors of generalised tournaments on n vertices as a convex polytope in $\mathbb{R}^n$. We call this the *score vector polytope* .

Definition 3.3.3. The *score vector polytope* is the convex polytope defined by:

$$P_Q^n := \left\{ \boldsymbol{s} \in \mathbb{R}^n : s_i \geq 0 \quad \forall i \in [n],\ \forall I \subsetneq [n] \ \sum_{i \in I} s_i \geq \binom{|I|}{2},\ \sum_{i=1}^{n} s_i = \binom{n}{2} \right\}.$$

Now if G is a generalised tournament on n vertices with score vector $\boldsymbol{s}$, the edges $\{t_{ij} : i, j \in [n]\}$ that define G are also vectors that satisfy a set of linear equality and inequality constraints and thus also define a convex polytope.

Definition 3.3.4. We call the set

$$GT_n := \left\{ \boldsymbol{t} \in \mathbb{R}^{n^2} : t_{jk} \geq 0,\ t_{jj} = 0,\ \forall k \neq j\ t_{jk} + t_{kj} = 1 \right\}$$

the polytope of *generalised tournaments on n vertices* .

Definition 3.3.5. We call the convex polytope

$$GT_n(\boldsymbol{s}) := \left\{ \boldsymbol{t} \in \mathbb{R}^{n^2} : t_{jk} \geq 0,\ t_{jj} = 0,\ \forall k \neq j\ t_{jk} + t_{kj} = 1,\ \sum_{k=1}^{n} t_{jk} = s_j \right\}$$

the polytope of *generalised tournaments with score vector* $\boldsymbol{s}$.

Note that by the necessity of the Landau-like conditions, we have $GT_n = \bigcup_{\boldsymbol{s} \in P_Q^n} GT_n(\boldsymbol{s})$.

We now proceed to work out some details about these polytopes.

3.3.1 Vertices of the Polytopes

In this subsection we find the vertices of some of these polytopes. This is significant, because since these polytopes are bounded, they are equal to the convex hull of their vertices ([Zie95], Theorem 1.1).

Proposition 3.3.6 (Vertices of GT_n)**.** *The vertices of GT_n correspond to usual tournaments $T \in \mathbb{T}_n$.*

Proof. This proof is rather straightforward, and follows from the definition of vertices of a polytope. □

Proposition 3.3.7 (Vertices of P_Q^n)**.** *The vertices of P_Q^n are the vectors corresponding to the $n!$ permutations of $(0, 1, 2, \cdot, n-1)$.*

Proof. Suppose $\boldsymbol{s} \in P_Q^n$ has its ith component $s_i = n-1$. Then, the conditions on P_Q^n imply that this is the maximum possible value of s_i, so that if $\boldsymbol{s}$ is the convex combination of vectors in P_Q^n, all of these vectors have their ith components equal to $n-1$. The maximum value of any of the other components of these vectors is then $n-2$, so that if the jth component of $\boldsymbol{s}$ is $n-2$, then the jth component of all of these vectors is $n-2$. By iterating in this manner, it follows that $\boldsymbol{s}$ can only be expressed as the convex combination of itself in P_Q^n and therefore must be a vertex.

Suppose that a vector $\boldsymbol{s} \in P_Q^n$ is not a permutation of $(0, 1, 2, \cdots, n-1)$. We will show that $\boldsymbol{s}$ is not a vertex by expressing it as the convex combination of two other vectors in

P_Q^n. Re-order the components of $\boldsymbol{s}$ so that they are non-decreasing. Then the condition for inclusion in P_Q^n is equivalent to

$$\sum_{i=1}^{k} s_i \geq \binom{k}{2},$$

with equality if $k = n$. If $\boldsymbol{s}$ has every component distinct, then if j is the first index such that $s_j > j - 1$, for ε sufficiently small the vectors obtained from $\boldsymbol{s}$ such that $s_j \mapsto s_j + \varepsilon$, $s_{j+1} \mapsto s_{j+1} - \varepsilon$, and $s_j \mapsto s_j - \varepsilon$, $s_{j+1} \mapsto s_{j+1} + \varepsilon$ are contained in P_Q^n, and evidently this allows $\boldsymbol{s}$ to be expressed as a convex combination.

Otherwise, $\boldsymbol{s}$ has some components, say $s_j, \cdots, s_{j+l}$ having the same value. It must be the case that

$$\sum_{i=1}^{j} s_i > \binom{j}{2},$$

so we may apply a similar argument to show that the vector such that $s_j \mapsto s_j - \varepsilon$ and $s_{j+l} \mapsto s_{j+l} + \varepsilon$ is in P_Q^n. Noting the symmetry in the fact that $s_j = \cdots = s_{j+l}$, the vector with $s_j \mapsto s_j + \varepsilon$ and $s_{j+l} \mapsto s_{j+l} - \varepsilon$ also belongs to P_Q^n, and this again allows us to express $\boldsymbol{s}$ as a convex combination. $\square$

The interesting thing about the Proposition 3.3.7, is that it shows that the polytope P_Q^n is actually a known polytope called the *permutation polytope* ([Bow72]).

Examples 3.3.8. The following pictures illustrate the polytopes P_Q^n when $n = 2, 3$ and 4. In the cases $n = 2, 3$, the polytope is illustrated as a shaded region within the ambient space. For $n = 4$, the region occupied by the polytope has been projected onto three-dimensions, and the net of the polytope has been drawn. These figures have been produced using the open source software system SageMath.

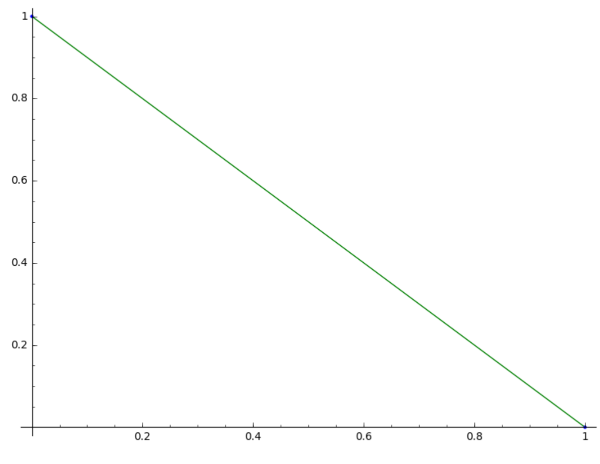


Figure 3.3: The green line illustrates the polytope P_Q^2 in $\mathbb{R}^2$.

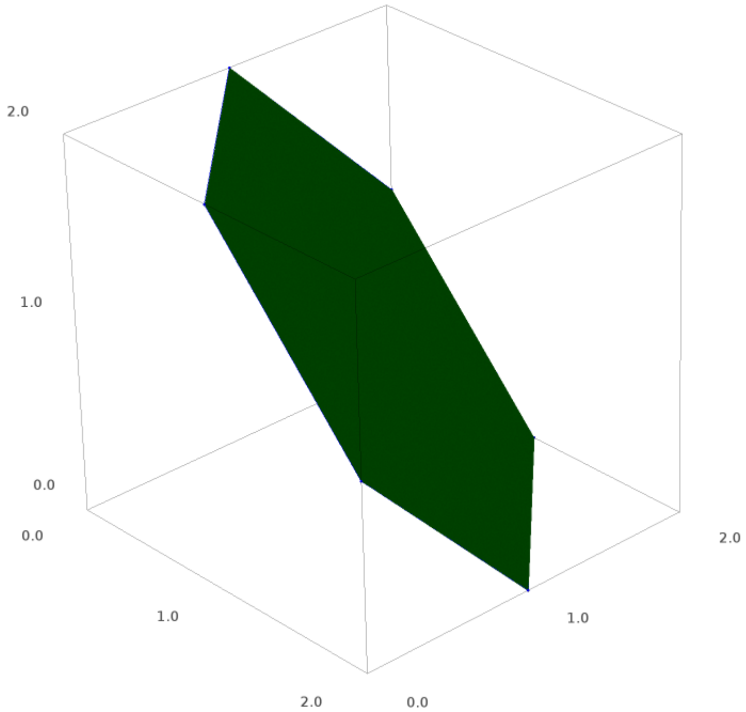


Figure 3.4: The green region illustrates the polytope P_Q^3 in $\mathbb{R}^3$.

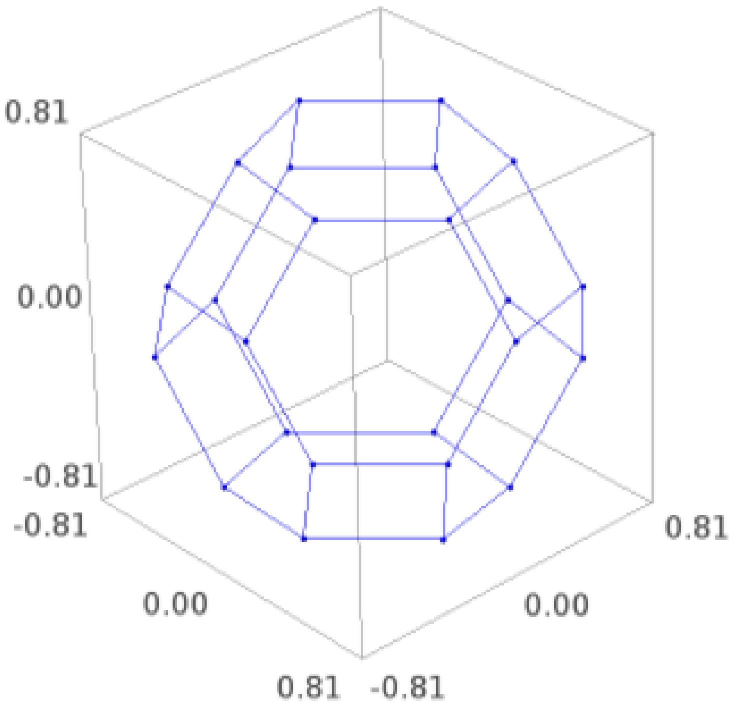


Figure 3.5: The net of the polytope P_Q^4, after the local region has been projected onto three dimensions.

3.3.2 Relative Interiors

In this subsection we compute the relative interiors of P_Q^n and $GT_n(\boldsymbol{s})$. We will require the following fundamental result from convex analysis on the relative interior of the intersection of two convex sets:

Lemma 3.3.9. *Let C_1 and C_2 be convex sets such that $\operatorname{rint}(C_1) \cap \operatorname{rint}(C_2) \neq \emptyset$. Then*

$$\operatorname{rint}(C_1 \cap C_2) = \operatorname{rint}(C_1) \cap \operatorname{rint}(C_2).$$

Proof. See [HU01], Proposition 2.1.10. □

This allows us to compute the relative interiors of P_Q^n and $GT_n(\boldsymbol{s})$:

Proposition 3.3.10. *The relative interior of P_Q^n is such that*

$$\operatorname{rint}(P_Q^n) = \left\{ \boldsymbol{s} \in P_Q^n : \forall\, \emptyset \subsetneq I \subsetneq [n] \sum_{i \in I} s_i > \binom{|I|}{2} \right\}.$$

Proof. Define

$$C_1 := \left\{ \boldsymbol{s} \in \mathbb{R}^n : s_i \geq 0,\ \forall\, \emptyset \subsetneq I \subsetneq [n] \sum_{i \in I} s_i \geq \binom{|I|}{2} \right\}$$

and

$$C_2 := \left\{ \boldsymbol{s} \in \mathbb{R}^n : \sum_{i=1}^{n} s_i = \binom{n}{2} \right\}.$$

Then C_1 and C_2 are convex sets, and $P_Q^n = C_1 \cap C_2$. Now, since $\operatorname{aff}(C_2) = C_2$ we have $\operatorname{rint}(C_2) = C_2$. We claim that

$$\operatorname{rint}(C_1) = \left\{ \boldsymbol{s} \in \mathbb{R}^n : s_i \geq 0,\ \forall\ \emptyset \subsetneq I \subsetneq [n]\ \sum_{i \in I} s_i > \binom{|I|}{2} \right\} := H.$$

For one inclusion, if $\boldsymbol{s} \in H$ then for $\varepsilon > 0$ sufficiently small $\mathrm{B}_\varepsilon(\boldsymbol{s}) \subseteq C_1$, thus $\boldsymbol{s} \in \operatorname{rint}(C_1)$. For the other inclusion, suppose for some $\boldsymbol{s} \in \operatorname{rint}(C_1)$ there is some set $\emptyset \subsetneq I \subsetneq [n]$ such that $\sum_{i \in I} s_i = \binom{|I|}{2}$. Then for $w > 0$ sufficiently small $\mathrm{B}_w(\boldsymbol{s}) \cap \operatorname{aff}(C_1) \subset C_1$. Since the vector $\boldsymbol{q}$ such that $q_i = s_i + 1$ is in C_1, for all $\lambda > 0$ we have that $(1+\lambda)\boldsymbol{s} - \lambda \boldsymbol{q} \in \operatorname{aff}(C_1)$. Thus, the vector $\boldsymbol{r}$ such that $r_i = s_i - \frac{w}{n}$ satisfies $\boldsymbol{r} \in \mathrm{B}_w(\boldsymbol{s}) \cap \operatorname{aff}(C_1)$, however, $\sum_{i \in I} r_i < \binom{|I|}{2}$ so $\boldsymbol{r} \notin C_1$. This contradiction proves the claim.

Now, evidently $\operatorname{rint}(C_1) \cap \operatorname{rint}(C_2) \neq \emptyset$ since, for example, the vector $\left(\frac{n-1}{2}, \cdots, \frac{n-1}{2}\right)$ belongs to the set. Thus, by Lemma 3.3.9,

$$\operatorname{rint}(P_Q^n) = \operatorname{rint}(C_1) \cap \operatorname{rint}(C_2) = \left\{ \boldsymbol{s} \in P_Q^n : \forall\ \emptyset \subsetneq I \subsetneq [n]\ \sum_{i \in I} s_i > \binom{|I|}{2} \right\}.$$ □

Proposition 3.3.11. *Suppose there exists $\boldsymbol{t} \in GT_n(\boldsymbol{s})$ such that, for all $i \neq j$, $t_{ij} > 0$. Then* $\operatorname{rint}(GT_n(\boldsymbol{s})) = \{\boldsymbol{t} \in GT_n(\boldsymbol{s}) : t_{ij} > 0\}$.

Proof. The proof is similar to Proposition 3.3.10, by considering $GT_n(\boldsymbol{s})$ as the intersection of two convex sets K_1 and K_2 where

$$K_1 := \left\{\boldsymbol{t} \in \mathbb{R}^{n^2} : t_{jj} = 0,\ t_{ij} \geq 0\right\}$$

and

$$K_2 := \left\{\boldsymbol{t} \in \mathbb{R}^{n^2} : \forall k \neq j\ \ t_{ij} + t_{ji} = 1,\ \sum_{k=1}^{n} t_{jk} = s_j\right\}. \qquad \square$$

3.3.3 Connections between Polytopes

The results in this subsection relate the polytopes P_Q^n and $GT_n(\boldsymbol{s})$.

Theorem 3.3.12. *A vector $\boldsymbol{s}$ belongs to P_Q^n if and only if the polytope $GT_n(\boldsymbol{s})$ is non-empty.*

Proof. Evidently if $GT_n(\boldsymbol{s}) \neq \emptyset$ then $P_Q^n \neq \emptyset$, by the necessity of the Landau-like conditions.

For the other direction, since P_Q^n is a bounded polytope, it is equal to the convex hull of its vertices. Thus, given $\boldsymbol{s} \in P_Q^n$, we have

$$\boldsymbol{s} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_{n!}, \quad \sum_i c_i = 1,\ c_i \geq 0. \tag{3.5}$$

In this equation $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{n!}$ are the vertices of P_Q^n. Now, as the vertices $\boldsymbol{v}_i \in P_Q^n$ correspond uniquely to transitive tournaments $\boldsymbol{t}_{v_i}$ in GT_n, if we set

$$\boldsymbol{p} = \sum_i c_i \boldsymbol{t}_{v_i},$$

it is not too difficult to check that $\boldsymbol{p} \in GT_n(\boldsymbol{s})$ so it follows that $GT_n(\boldsymbol{s})$ is non-empty. $\square$

Corollary 3.3.13. *Given any generalised tournament G with score vector $\boldsymbol{s}$, there exists a generalised tournament with the same score vector that can be expressed as the convex combination of transitive tournaments.*

Our final result of this subsection relates the relative interiors of the two polytopes.

Proposition 3.3.14. *A vector $\boldsymbol{s}$ belongs to the relative interior of P_Q^n if and only if there exists a vector $\boldsymbol{p} \in GT_n(\boldsymbol{s})$ such that for all $i \neq j$ $p_{ij} > 0$.*

Proof. The necessity of these conditions is straightforward. For the other direction, suppose $\boldsymbol{s} = (s_1, \cdots, s_n) \in \operatorname{rint}(P_Q^n)$. Then, for sufficiently small $\varepsilon > 0$, define $\boldsymbol{s}^{\boldsymbol{\varepsilon}}$ to be the vector such that the ith component $s_i^{\varepsilon} = \frac{s_i - (n-1)\varepsilon}{1 - 2\varepsilon}$ for each $i \in [n]$. A routine calculation shows that for sufficiently small $\varepsilon > 0$, $\boldsymbol{s}^{\boldsymbol{\varepsilon}} \in P_Q^n$. Therefore, by Theorem 3.3.12, the polytope $GT_n(\boldsymbol{s}^{\boldsymbol{\varepsilon}})$ is non-empty. Suppose $\boldsymbol{\lambda}$ is a vector such that $\boldsymbol{\lambda} \in GT_n(\boldsymbol{s}^{\varepsilon})$. Then a simple calculation shows that the vector $\boldsymbol{p} \in GT_n(\boldsymbol{s})$ defined by

$$p_{ij} = \begin{cases} (1-2\varepsilon)\lambda_{ij} + \varepsilon, & i \neq j, \\ 0, & \text{otherwise.} \end{cases}$$

satisfies the required properties. $\square$

3.3.4 Applications to Voting Theory

In any preferential election with candidates $\{c_1, \cdots, c_n\}$, the preferences of a single voter d_1 may be expressed as a transitive tournament T_1. If all of the preferences of each voter are taken into account, the result of an election with k voters may then be represented as a generalised tournament, corresponding to the convex combination

$$\sum_{i=1}^{k} \frac{1}{k} T_i.$$

The famous voters paradox [Sil92] shows that the individual preferences may be in conflict with each other in such an election.

Example 3.3.15. Suppose an election has three candidates: Hillary Clinton (C), Donald Trump (T) and Scott Morrison (M). Suppose there are three voters, and their preferences are represented by the following transitive tournaments:

$$\begin{tikzcd} & \mathrm{T} \arrow[dl] \arrow[dr] & & & \mathrm{C} \arrow[dl] \arrow[dr] & & & \mathrm{M} \arrow[dl] \arrow[dr] & \\ \mathrm{M} & & \mathrm{C} \arrow[ll] & \mathrm{T} & & \mathrm{M} \arrow[ll] & \mathrm{C} & & \mathrm{T} \arrow[ll] \end{tikzcd}$$

Figure 3.6

Then, the result of the election may be represented by the following generalised tournament:

$$\begin{tikzcd} & \mathrm{T} \arrow[dr, "\frac{2}{3}"] & \\ \mathrm{M} \arrow[ur, "\frac{2}{3}"] & & \mathrm{C} \arrow[ll, "\frac{2}{3}"] \end{tikzcd}$$

Figure 3.7

We see that the majority of the population preference Trump over Hillary, Hillary over Morrison, and Morrison over Trump, so there can be no fair way to call this election. All of the candidates have equal score.

More generally, if we think of the score vector of the resulting generalised tournament as a means of ranking the participants, it is interesting to see which score vectors are possible outcomes of such an election. Corollary 3.3.13 shows that any score vector is possible if the coefficients of the convex combination are real numbers. This result roughly means that such an election, with a fixed number of candidates, can have any possible result as the number of voters tends to infinity.

3.4 Moon's Theorem for Generalised Tournaments

Theorem 3.3.12 was actually first proved by Moon in 1963 [Moo63]. In this section we provide another proof of this theorem, by applying Theorem 3.2.5 along with a convergence argument adapted from the proof by Bang and Sharp in 1976 [BjS77].

It will first be beneficial to express Theorem 3.2.5 in a different form, by using the following definition.

Definition 3.4.1. A *p-weighted tournament* P_n on n vertices is an ordered pair (V, T) where, V is a finite set of *vertices* $\{v_1, ..., v_n\}$, and T is a collection of non-negative integer *edges* $\{t_{ij} : i, j \in [n]\}$ such that $t_{ij} + t_{ji} = p$, $t_{jj} = 0$ and $0 \leq t_{ij} \leq p$ for all $i, j \in [n]$. We define *scores, score sequences* and *score vectors* in a similar way to other tournaments we have studied.

It is then not too difficult to check the following reformulation of Theorem 3.2.5:

Theorem 3.4.2. *A sequence of non-negative integers $s_1 \leq s_2 \leq \cdots \leq s_n$ is a score-sequence for some p-weighted tournament $P_n = (V, T)$ on n vertices if and only if*

$$\sum_{i=1}^{k} s_i \geq p\binom{k}{2} \tag{3.6}$$

for all positive integers $k < n$, and

$$\sum_{i=1}^{n} s_i = p\binom{n}{2}. \tag{3.7}$$

□

Theorem 3.4.3 (Moon)**.** *A sequence $\boldsymbol{s}$ of non-negative real numbers $s_1 \leq s_2 \leq \cdots \leq s_n$ is a score sequence for some generalised tournament G if and only if*

$$\sum_{i=1}^{k} s_i \geq \binom{k}{2} \tag{3.8}$$

for all positive integers $k \leq n$, with equality if $k = n$.

Proof. The necessity of these conditions are straightforward.

For the other direction, assume first that the sequence $\boldsymbol{s}$ is rational. Then, for some integer r sufficiently large, $(rs_1, ..., rs_n)$ is a sequence satisfying equations (3.6) and (3.7), so by Theorem 3.2.5 there exists a r-weighted tournament R_n on n vertices. Suppose the edge set of this r-weighted tournament is given by $\{p_{ij} : i, j \in [n]\}$. Then, setting $t_{ij} := \frac{p_{ij}}{r}$, $t_{jj} = 0$, we obtain a generalised tournament $\boldsymbol{t}$ with this score sequence.

Otherwise, since the set of possible generalised tournaments on n vertices

$$W := \left\{ \boldsymbol{t} \in \mathbb{R}^{n^2} : t_{jk} \geq 0,\ t_{jj} = 0,\ \forall k \neq j\ t_{jk} + t_{kj} = 1 \right\}$$

is a closed and bounded subset of $\mathbb{R}^{n^2}$, it is compact. Thus, if we can approximate any real sequence $\boldsymbol{s}$ by rational sequences converging to $\boldsymbol{s}$, by a compactness argument the corresponding generalised tournaments $\boldsymbol{t} \in W$ will contain a subsequence converging to a generalised tournament $\boldsymbol{t}' \in W$ with score sequence $\boldsymbol{s}$.

In this way, suppose we are given a sequence $\boldsymbol{s} = (s_1, \cdots, s_n)$ with at least one irrational component. There exists a largest integer $m < n$ such that $s_m < s_{m+1}$. Now for each integer $K > 0$, choose rational numbers $q_1, \cdots, q_m$ such that $q_1 \leq q_2 \leq \cdots \leq q_m$ and for $1 \leq i \leq m$

$$s_i < q_i \leq s_i + \frac{s_{m+1} - s_m}{2Km},$$

whilst for $i > m$ we define

$$q_i = q_{i+1} = ... = q_n = \frac{\binom{n}{2} - \sum_{k=1}^{m} q_k}{n - m}.$$

It can easily be checked that $\boldsymbol{q} = (q_1, ..., q_n)$ satisfies equation (3.8) and converges to $\boldsymbol{s}$ as K goes to infinity, thus the argument is complete. □

3.5 Connection to Random Tournaments

If $\mathscr{T}_n$ is a random tournament on n vertices, we can construct a generalised tournament G by setting each edge t_{ij} equal to the probability of an edge directed from vertex v_i to v_j.

More precisely, let $\mathbb{P}_n$ be the space of all random tournaments on n vertices $\{v_1, \cdots, v_n\}$ and define the map

$$F : \mathbb{P}_n \to GT_n$$

such that

$$\mathscr{T}_n \mapsto \boldsymbol{p} = (p_{ij} :\ i, j \in [n])$$

where each p_{ij} is the probability of an edge from vertex v_i to v_j.

Definition 3.5.1. Following the above notation, if $\mathscr{T}_n \in \bigcup_{\boldsymbol{t}\in GT_n(\boldsymbol{s})} F^{-1}(\boldsymbol{t})$ we say R lies *above* $GT_n(\boldsymbol{s})$.

Remark 3.5.2. Clearly the set $F^{-1}(\boldsymbol{t})$ is non-empty because we can use the edges of G to construct an edge model random tournament.

Remark 3.5.3. By linearity of expectation, if R is a random tournament above the generalised tournament G_n, the score of a vertex v_i corresponds to the expected score of v_i in R.

Proposition 3.5.4. *A vector $\boldsymbol{s} \in P_Q^n$ if and only if there exists a random tournament R above $GT_n(\boldsymbol{s})$.*

Proof. By definition, $F(R) \in GT_n(\boldsymbol{s})$, so $\boldsymbol{s} \in P_Q^n$ by Theorem 3.3.12. For the other direction, by Theorem 3.3.12, choose a vector $\boldsymbol{p}$ in $GT_n(\boldsymbol{s})$. Then, picking a random tournament R above $\boldsymbol{p}$ we have the result. $\square$

Chapter 4

The Beta Model

4.1 Introduction

4.1.1 Motivation

There are many situations where it is important to rank objects based on a number of pairwise comparisons. But what does such a ranking mean? To answer this, we first look at some examples:

Examples 4.1.1.

1. Before a big cricket competition, a bookmaker would like to know the probabilities of the result of each game. He has, at his disposal, the history of all previous games between teams. How should he proceed?

2. A board wishes to rank the candidates participating in an election, based on the way a random voter is expected to preference the candidates (similar to Example 3.3.15).

We can see that in these examples, the objective of the 'ranking' system is to predict the outcome of future paired comparisons. We make the following assumptions:

1. The result of any paired comparison is binary — that is — it can have exactly one of two states.

2. The result of every pairing is independent of the others, and not predetermined (either state has non-zero probability).

3. We have, in our 'history', results from every possible pairing of objects we are comparing.

If the outcome of every pairing is represented by a directed edge and the objects are vertices, the goal is then to construct an edge model random tournament, with each edge having non-zero probability, that best describes the history. We will see that if we can use the history to form a generalised tournament with score vector $\boldsymbol{s}$ such that $\boldsymbol{s} \in \operatorname{rint} P_Q^n$, then such a construction is possible.

4.1.2 Literature

This problem is not new and has been widely studied in statistics, in what is known as the 'method of paired comparisons'. The model we will describe is a special case of what is known in the statistics literature as the 'Bradley-Terry' model, due to the paper by Bradley and Terry in 1952 [RAB52]. However, the approach was first discovered by Zermelo in 1929 [Zer29] and independently rediscovered again by Ford in 1957 and Jech in 1983 [Jec83]. These examples actually consider the possibility of multiple comparisons between the same pairs of objects, (corresponding to a multi-tournament). Similar to the approach outlined here, the main idea in these papers has been to use maximum likelihood estimation.

Simons and Yao, in 1999, have studied the asymptotics of the model as the number of comparisons tends to infinity, see [SY99]. A good review of this model, including many generalisations, may be found in [Hun04].

It is important to mention that the approach here only takes into account the overall score of each player obtained from a generalised tournament that represents the 'history' of previous encounters. However, other models, such as the Kendall-Wei method [Wei52] [Ken55] also take into account the quality of players defeated in previous games.

The literature on methods of paired comparisons is large, and could be the topic of a thesis in its own right. Whilst the idea of using maximum likelihood estimation is the same, the approach taken here to derive the model is quite different to a lot of other approaches in that in does not assume that the expected scores are integer valued. This has been inspired by literature on the so called β-model for random graphs. In particular, many results in this chapter are based on similar results in [CHK$^+$12] and [Bar12] for the random graph case. We will hence call this model the β-model for random tournaments, to emphasise the connections between these models. We also remark that a similar analysis could be carried out for multi-tournaments, however, we restrict ourselves to the tournament case for brevity.

4.2 The Beta Model

4.2.1 Preliminaries

If we have a history of paired comparisons between objects, we may use this data to construct a generalised tournament, where the weight of the edge λ_{jk} represents the relative strength of vertex j over vertex k. If this generalised tournament has score vector $\boldsymbol{s}$, the problem of predicting the results of future paired comparisons can be phrased as finding the most 'appropriate' random tournament having its expected score vector equal to $\boldsymbol{s}$. According to the Principle of Maximum Entropy, the probability distribution that best describes the current state is the probability function with maximum *entropy* (see [SJ80]). If X is a discrete random variable, taking values in $\{x_1, \cdots, x_n\}$, the entropy function is given by

$$H(X) = -\sum_{i=1}^{n} P(X = x_i) \log \left(P(X = x_i) \right),$$

with the convention that the values of the terms corresponding to $0 \log 0$ are set to 0. Note that the entropy function is additive with respect to independent events.

In order to find such any random tournament with fixed expected scores $\boldsymbol{s}$ this requires solving the system of equations

$$\sum_{j \neq i}^{n} p_{ij} = s_i \tag{4.1}$$

subject to the constraint

$$p_{jj} = 0 \quad \forall j \in [n] \tag{4.2}$$

and

$$p_{ij} + p_{ji} = 1 \quad \forall i \neq j \in [n]. \tag{4.3}$$

Here there are n^2 different variables, and $2n + \binom{n}{2}$ independent equations, so for $n \geq 4$ there will be infinitely many solutions to this system. If we assume, however, that there are 'strength' parameters β_i associated with each vertex v_i, then we may expect that the probability of a future encounter p_{ij} should depend on the difference $\beta_j - \beta_i$ (or vice versa). This motivates the following definition:

Definition 4.2.1. Let $\boldsymbol{p}$ be an edge model random tournament on n vertices such that each edge has non-zero probability. We say that $\boldsymbol{p}$ belongs to the *β-model* if there exists real parameters $\beta_1, \cdots, \beta_n$ such that for every pair of vertices v_i, v_j with $i \neq j$ we have

$$p_{ij} = \frac{e^{\beta_i}}{e^{\beta_i} + e^{\beta_j}} = \frac{e^{\beta_i - \beta_j}}{1 + e^{\beta_i - \beta_j}}. \tag{4.4}$$

where p_{ij} denotes the probability that i dominates j.

Remark 4.2.2. Note that as this equation depends only on the differences $\beta_i - \beta_j$ (representing relative strengths of the objects being compared), translating each β_i by a constant will result in the same edge model random tournament.

Example 4.2.3. Since every edge in the uniform random tournament has probability $\frac{1}{2}$, the uniform random tournament belongs to the β-model, since we may fix some constant c and set every $\beta_i = c$.

4.2.2 Existence

The following Theorem shows that the random tournament with maximum entropy satisfying equation (4.1) belongs to the β-model. Note that the entropy function of an arbitrary edge model random tournament is given by

$$H(\boldsymbol{p}) = -\sum_{1 \leq i < j \leq n} (p_{ij} \log p_{ij} + (1 - p_{ij}) \log (1 - p_{ij})). \tag{4.5}$$

Theorem 4.2.4 (Existence of β Parameters)**.** *Suppose* $\boldsymbol{s} \in \operatorname{rint} P_Q^n$*. Then the maximiser of* (4.5) *subject to the constraint* (4.1) *is a* β*-model random tournament.*

Proof. By Proposition 3.3.14, since $\boldsymbol{s} \in \operatorname{rint} P_Q^n$, there exists a random tournament $\boldsymbol{y}$ satisfying equation (4.1) such that all of the probabilities $y_{ij} > 0$. Now, suppose $\boldsymbol{b}$ is a random tournament that maximises H, with at least one of the probabilities $b_{ij} = 0$ or 1. At $b_{ij} = 0$ the right handed partial derivative satisfies

$$\frac{\partial}{\partial b_{ij}} = \infty$$

and at $b_{ij} = 1$ the left handed partial derivative satisfies

$$\frac{\partial}{\partial b_{ij}} = -\infty,$$

whilst the partial derivative is finite if $0 < b_{ij} < 1$. It follows, that for some $\varepsilon > 0$ the random tournament $\boldsymbol{z} := (1-\varepsilon)\,\boldsymbol{b} + \varepsilon\boldsymbol{y}$ has all of its edge probabilities $z_{ij} > 0$, satisfies equation (4.1) and is such that $H(\boldsymbol{z}) > H(\boldsymbol{b})$, a contradiction. It follows that the random tournament $\boldsymbol{p}$ that maximises H is such that $\boldsymbol{p} \in \operatorname{rint} GT_n(\boldsymbol{s})$.

If we denote Lagrangian multipliers by β_i and set

$$H^*(\boldsymbol{p}) = H(\boldsymbol{p}) + \sum_{i=1}^{n} \beta_i \left(\sum_{j \neq i} p_{ij} - s_i \right),$$

at the maximum, the partial derivatives satisfy

$$\frac{\partial H^*(\boldsymbol{p})}{\partial p_{ij}} = -\log \frac{p_{ij}}{1-p_{ij}} + \beta_i - \beta_j = 0$$

This implies that the p_{ij} satisfy equation (4.4), and, by Definition 4.2.1, $\boldsymbol{p}$ belongs to the β-model. □

4.2.3 Uniqueness

Remark 4.2.2 shows that the β_i corresponding to a β-model random tournament will in general not be unique. However, since the edge probabilities are only dependent on the differences $\beta_i - \beta_j$, we may expect the parameters to be unique up to translation by a constant.

Suppose therefore that $\boldsymbol{s} \in \operatorname{rint} P_Q^n$, and we have a solution for the β_i, such that, by (4.4), for each $i \in [n]$

$$\sum_{j \neq i} \frac{e^{\beta_i}}{e^{\beta_j} + e^{\beta_i}} = s_i. \tag{4.6}$$

By relabelling vertices if necessary, we may assume that $s_1 \leq s_2 \leq \cdots \leq s_n$, so that, by Proposition 4.2.8 $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n$. Moreover, by the translation invariance property (Remark 4.2.2), we may assume $\beta_1 = 0$.

In order to simplify notation, we define the vector $\boldsymbol{\alpha}$ so that $\alpha_i = e^{\beta_i}$ for each $i \in [n]$. Then, if $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is the function defined by

$$\phi_i(\boldsymbol{y}) = s_i \left(\sum_{j \neq i} \frac{1}{y_j + y_i} \right)^{-1},$$

$\boldsymbol{\alpha}$ is a fixed point of this function. Moreover, by (4.6), any fixed point of this function corresponds to a solution for the β-parameters. Note that if $\mathbb{R}^n_+$ denotes the set of all vectors in $\mathbb{R}^n$ with positive components, then ϕ fixes this space.

Define ρ on $\mathbb{R}^n_+$ such that for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n_+$,

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = \max\left(\max_{i \in [n]} \frac{x_i}{y_i}, \max_{i \in [n]} \frac{y_i}{x_i} \right).$$

Lemma 4.2.5. *For any integer $n \geq 1$ and positive numbers $p_1, \cdots, p_n$ and $q_1, \cdots, q_n$,*

$$\frac{p_1 + \cdots + p_n}{q_1 + \cdots + q_n} \leq \max_{i \in [n]} \frac{p_i}{q_i}, \tag{4.7}$$

with equality exactly when all of the ratios $\frac{p_i}{q_i}$ are the same.

Proof. Suppose each $p_i \leq k q_i$ for some $k > 0$. Then the left hand side of equation (4.7) is bounded above by k, with equality exactly when each $\frac{p_i}{q_i} = k$. ☐

Theorem 4.2.6. *For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n_+$ with $\boldsymbol{x} \neq c\boldsymbol{y}$ for any constant $c > 0$*

$$\rho\left(\phi(\boldsymbol{x}), \phi(\boldsymbol{y})\right) < \rho(\boldsymbol{x}, \boldsymbol{y}).$$

Proof. We have

$$\frac{\phi_i(\boldsymbol{x})}{\phi_i(\boldsymbol{y})} = \frac{\sum_{j \neq i} (y_j + y_i)^{-1}}{\sum_{j \neq i} (x_j + x_i)^{-1}} \leq \max_{j \neq i} \frac{x_j + x_i}{y_j + y_i} \leq \max_{j \neq i} \max\left(\frac{x_j}{y_j}, \frac{x_i}{y_i} \right),$$

where both inequalities follow from Lemma 4.2.5. We also see that equality holds precisely when $\boldsymbol{x} = c\boldsymbol{y}$ for some constant $c > 0$, thus the inequality is strict. The same argument holds when we reverse $\boldsymbol{x}$ and $\boldsymbol{y}$, and the result follows. ☐

Corollary 4.2.7 (Uniqueness of β-Parameters)**.** *The β_i corresponding to the random tournament $\boldsymbol{p}$ in Theorem 4.2.4 can be computed by iterating ϕ. Moreover, these β_i will be unique up to the addition of a constant.*

Proof. Define the metric $d(\boldsymbol{x},\boldsymbol{y}) = \log \rho(\boldsymbol{x},\boldsymbol{y})$ on $\mathbb{R}^n_+$. If $\boldsymbol{\alpha}$ denotes a fixed point of ϕ, then every point on the line $\{c\boldsymbol{\alpha} : c > 0\}$ is also a fixed point, by Remark 4.2.2. If we start with a point $\boldsymbol{x}_0$, and construct the sequence of iterates $\boldsymbol{x}_{i+1} = \phi(\boldsymbol{x}_i)$, then by Theorem 4.2.6, the distance from each $\boldsymbol{x}_i$ and this line is strictly decreasing and bounded below by 0. Moreover, any limit point $\lim_{n\to\infty} \boldsymbol{x}_n$ must be a fixed point, and by Theorem 4.2.6, must lie on this line.

Since the fixed point $\boldsymbol{\alpha}$ is unique up to multiplication by a positive constant, the β_i are unique up to addition by a constant. □

Note that whilst this Corollary provides an algorithm for computing the β_i, the main importance of this Corollary is to prove the uniqueness up to translation by a constant. Interior point methods can be applied to solve the constrained optimisation problem in (4.2.4) for the random tournament $\boldsymbol{p}$ directly, from which the β_i can be calculated.

4.2.4 Equivalence to Ranking Based on Scores

If our motivation for studying the β-model is purely to rank objects based on previous information encoded in some generalised tournament, how do the β_i compare to naively ranking based on the score vector? The following Proposition shows, rather unsurprisingly, that the two are equivalent:

Proposition 4.2.8. *Let $\boldsymbol{p}$ be a random tournament on n vertices belonging to the β-model $\{1, 2, \cdots, n\}$. If $\boldsymbol{s}$ denotes the expected score vector of $\boldsymbol{p}$, then*

$$s_1 \leq s_2 \leq \cdots \leq s_n$$

if and only if

$$\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n.$$

Proof. Note that for each i

$$s_i = \sum_{k\neq i\in[n]} \frac{e^{\beta_i-\beta_k}}{1+e^{\beta_i-\beta_k}}.$$

Now, if we define a function $f : \mathbb{R} \to \mathbb{R}$ by

$$-\frac{1}{2} + \sum_{k=1}^{n} \frac{e^{x-\beta_k}}{1+e^{x-\beta_k}},$$

this function is monotonically increasing for $x \in \mathbb{R}$. Noting that $f(\beta_i) = s_i$ for $i \in [n]$, the result follows. □

In summary, an approach for applications to comparison based ranking could be as follows:

1. Form a generalised tournament where each edge between objects corresponds to information from previous known encounters between the objects. For example, in sporting applications, the relative number of wins over all previous encounters could be used.

2. Compute the generalised score vector $\boldsymbol{s}$ and check that $\boldsymbol{s} \in \operatorname{rint} P_Q^n$. This will involve sorting the scores in non-decreasing order, and then checking a series of inequality constraints.

3. Apply either Corollary 4.2.7 or an interior point method to solve for the β_i.

4.2.5 Maximum Likelihood Estimation in the β-Model

Suppose we view a tournament T with score vector $\boldsymbol{s} \in \operatorname{rint} P_Q^n$, and wish to find the β-model random tournament that maximises the probability of T. The probability of a tournament T with score vector $\boldsymbol{s} = (s_1, s_2, \cdots, s_n)$ in the β-model is given by

$$\prod_{(i,j)\in T} p_{ij} = \prod_{(i,j)\in T} \frac{e^{\beta_i}}{e^{\beta_i} + e^{\beta_j}} = \frac{\prod_{i=1}^n e^{s_i \beta_i}}{\prod_{1\le i<j\le n} (e^{\beta_i} + e^{\beta_j})}, \tag{4.8}$$

where the last equality follows from the fact that the denominator is same for the probability of the edge (i,j) and the probability of the edge (j,i). The log-likelihood equations are given by

$$\frac{\partial}{\partial \beta_i} \log P(T) = s_i - \sum_{j\neq i} \frac{e^{\beta_i}}{e^{\beta_i} + e^{\beta_j}},$$

which implies that the maximum likelihood estimator is the β-model that also maximises the entropy in Theorem 4.2.4. We thus find that a β-model random tournament with fixed integral expected scores maximises the likelihood of any tournament with expected score vector $\boldsymbol{s}$.

This is the context in which the β-model is usually presented in the literature. Assuming the expected score vector has only integral values, the condition $\boldsymbol{s} \in \operatorname{rint} P_Q^n$ is equivalent to this score vector being the score vector of a strong tournament, which, by Theorem 1.6.2, is equivalent to the following condition (which implies a tournament is irreducible), seen in [For57], [SY99],[Hun04] and other papers on the topic:

Condition 1. For every partition of the objects into two non-empty sets, an object in the second set has dominated an object in the first at least once.

Thus, this is equivalent to the approach taken here when the underlying score vector of the generalised tournament corresponding to the 'history' takes only integer values.

Note that these papers also show that the same condition can be applied for ranking in multi-tournaments. The details in this case are quite similar.

4.2.6 Conditional Distribution of a Tournament given a Score Vector

According to equation (4.8), the probability of a particular tournament T in the β-model, conditioned on the fact that it has score vector $\boldsymbol{s}$, is uniform. The following Theorem shows that the reverse implication also holds. This is unsurprising, since, in Theorem 4.2.4 the β_i were derived by considering an entropy maximisation problem.

Theorem 4.2.9. *Let T be an edge model random tournament on n vertices with independent edges, such that each edge has a non-zero probability. Then T belongs to the β-model if and only if for any score vector $\boldsymbol{s} = (s_1, \cdots, s_n)$, the conditional distribution of a tournament given that score vector is uniform.*

Proof. Necessity follows immediately from equation (4.8). For sufficiency, first note that if a tournament contains a directed triangle, reversing the edges of the triangle preserves the score vector of the tournament. Since any two tournaments with the same score vector have equal probability, it follows that the probability of any directed triangle occurring is the same as the probability of the reverse triangle occurring.

Now, for any three distinct vertices $v_i, v_j, v_k \in V(T)$ there obviously exists a tournament containing a directed triangle with these three points. Thus, if for every pair of vertices v_i, v_j we have

$$\frac{p_{ij}}{1-p_{ij}} := \mu_{ij},$$

then for any $v_i, v_j, v_k \in V(T)$

$$\mu_{ij}\mu_{jk}\mu_{ki} = 1.$$

If we fix two vertices v_i and v_j, for $p \neq i, j \in [n]$ we get

$$\prod_{p\neq i,j} \mu_{ij}\mu_{jp}\mu_{pi} = \mu_{ij}^{n-2} \prod_{p\neq i,j} \mu_{jp}\mu_{pi} = 1.$$

Also, since $\mu_{ij}\mu_{ji} = 1$, we have

$$\mu_{ij}^{n} \prod_{p\neq j} \mu_{jp} \prod_{p\neq i} \mu_{pi} = 1. \tag{4.9}$$

Since the product

$$\prod_{p\neq i} \mu_{ip}$$

only depends on i, if for every vertex v_m in T we define

$$e^{\beta_m} := \sqrt[n]{\prod_{p\neq m} \mu_{mp}}, \tag{4.10}$$

we have $\mu_{ij} = e^{\beta_i - \beta_j}$, and this implies that the tournament belongs to the β-model. $\square$

The expected score vector $\boldsymbol{s}$ of a β-model random tournament is such that $\boldsymbol{s} \in \operatorname{rint} P_Q^n$. What is the probability of a strong score vector in an edge model random tournament that does not belong to the β-model?

Theorem 4.2.10. *Let $\boldsymbol{p}$ be an edge model random tournament with expected score vector $\boldsymbol{r} \notin \operatorname{rint} P_Q^n$. Then the probability of a tournament with a strong score vector $\boldsymbol{z}$ is 0.*

Proof. Since $\boldsymbol{r} \notin \operatorname{rint} P_Q^n$, for some nonempty subset $I \subsetneq [n]$ we have

$$\sum_{i \in I} r_i = \binom{|I|}{2}. \tag{4.11}$$

Let X be a tournament valued random variable drawn from $\boldsymbol{p}$, with score vector $\boldsymbol{s} = (s_1(X), \cdots s_n(X))$, and let $f(X)$ be the value of the sum

$$\sum_{i \in I} s_i(X).$$

On the one hand, by 4.11, $\mathbb{E}\, f(X) = \binom{|I|}{2}$. One the other hand, by the definition of expectation,

$$\mathbb{E}\, f(X) = \sum_{K \in \mathbb{T}_n} P(K) f(K). \tag{4.12}$$

Since $f(T) > \binom{|I|}{2}$ for any tournament T with strong score vector $\boldsymbol{z}$, the only way (4.12) can hold is if $P(T) = 0$. □

4.3 Tame Score Sequences

In this section we provide some conditions, which, if satisfied by a score vector $\boldsymbol{s}$, implies the existence of a generalised tournament $\boldsymbol{p} \in GT_n(\boldsymbol{s})$ satisfying 'nice' properties. We will make this more precise, but for now we begin with a condition on the score vector of a generalised tournament so that the likelihood equations of Theorem 4.2.4 have a solution.

Proposition 4.3.1. *Suppose $\boldsymbol{s}$ is a score vector such that for each s_i we have*

$$\alpha < \frac{s_i}{n-1} < \beta,$$

with α and β such that

$$\frac{n}{n-1} \geq 2(\beta - \alpha) \tag{4.13}$$

Then $\boldsymbol{s} \in \operatorname{rint} P_Q^n$, and in particular the likelihood equations of Theorem 4.2.4 have a solution.

Proof. Note that by Proposition 3.3.10, after ordering and re-labelling the s_i if necessary, it suffices to show that for each $t \in [n]$ we have

$$\sum_{i=1}^{t} s_i > \binom{t}{2} = \frac{t(t-1)}{2}.$$

By counting in two ways, using $\sum_{i=1}^{n} s_i = \binom{n}{2}$, this will be true if

$$\max\left(t\alpha(n-1), \frac{n(n-1)}{2} - (n-t)\beta(n-1)\right) \geq \frac{t(t-1)}{2}. \tag{4.14}$$

Define a function $f : \mathbb{R} \to \mathbb{R}$ by $f(t) = \frac{n(n-1)}{2} - (n-t)\beta(n-1) - \frac{t(t-1)}{2}$. Note that, for fixed n and β, f is a quadratic, with maximum value obtained at $t = \beta(n-1) + \frac{1}{2}$ and $f(n) = 0$. It follows that $f(t) \geq 0$ for $2\beta(n-1) + 1 - n \leq t \leq n$.

Suppose then that $t\alpha(n-1) < \frac{t(t-1)}{2}$, so $2\alpha(n-1) + 1 < t$. Then equation (4.14) will be satisfied if

$$2\alpha(n-1) + 1 \geq 2\beta(n-1) + 1 - n \implies n \geq 2(\beta - \alpha)(n-1)$$

which is exactly the condition of (4.13). □

The following corollary is useful, because it provides a condition that that is independent of n.

Corollary 4.3.2. *A vector* $\boldsymbol{s}$ *belongs to* $\operatorname{rint} P_Q^n$ *if for each* s_i *we have*

$$\alpha < \frac{s_i}{n-1} < \beta,$$

with α *and* β *such that* $\beta - \alpha \leq \frac{1}{2}$.

□

Remark 4.3.3. Note that since s_i is a score-vector, we necessarily have $0 < \alpha \leq \frac{1}{2}$ and $\frac{1}{2} \leq \beta < 1$.

Proposition 4.3.1 gives criteria for the existence of a random tournament belonging to the β-model, via Theorem 4.2.4. However, the probability of an edge p_{ij} may be very small, depending on n, and approach 0 in the limit as n goes to infinity.

Definition 4.3.4. For a fixed $\delta > 0$, a score vector $\boldsymbol{s}$ is called δ-tame if there exists a random tournament belonging to the β-model with these expected scores, such that the probability of any edge $p_{ij} > \delta$.

Our final result for this chapter provides a condition for the score vector to be δ-tame.

Proposition 4.3.5. *Suppose δ is a constant with $0 < \delta \leq \frac{1}{2}$ and $\boldsymbol{s} \in P_Q^n$ is a score vector satisfying the condition that for each component s_i*

$$\alpha \leq \frac{s_i}{n-1} \leq \beta, \qquad \beta - \alpha \leq \frac{1}{2} - \delta. \tag{4.15}$$

Then $\boldsymbol{s}$ is δ-tame.

Proof. First note that by Corollary 4.3.2, there exists a β-model random tournament $\boldsymbol{p}$ with score vector $\boldsymbol{s}$. Assume, by relabelling if necessary, that

$$s_1 \leq s_2 \leq \cdots \leq s_n,$$

so that, by Proposition 4.2.8, the strength parameters of $\boldsymbol{p}$ are such that

$$\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n.$$

By invariance under translation by a constant we may assume that $\beta_1 = 0$. Also, set $b := \beta_n$. Then, by (4.15)

$$s_n - s_1 \leq (\beta - \alpha)(n-1) \leq \left(\frac{1}{2} - \delta\right)(n-1).$$

Note that

$$s_n - s_1 = \sum_{j=1}^{n} \left(\frac{e^{b-\beta_j}}{1+e^{b-\beta_j}} - \frac{e^{-\beta_j}}{1+e^{-\beta_j}} \right) = \sum_{j=1}^{n} \left(\frac{1}{1+e^{\beta_j - b}} - \frac{1}{1+e^{\beta_j}} \right).$$

Define a function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(t) = \left(\frac{1}{1+e^{t-b}} - \frac{1}{1+e^{t}} \right).$$

We look to minimise this function on the interval $[0, b]$. Noting the symmetry $f(t) = f(b-t)$ it is evident that this function is minimised on the boundary. Therefore

$$\frac{1}{1+e^{-b}} - \frac{1}{2} \leq \frac{1}{2} - \delta$$

which implies that $b \leq \log \frac{1-\delta}{\delta}$, so that in the random tournament $\boldsymbol{p}$ the probability of any edge is bounded below by δ. $\square$

Chapter 5

Asymptotic Enumeration

5.1 Introduction

5.1.1 Motivation

Suppose n cricket teams enter a round-robin competition and are ranked in a points table based only on their number of wins. Assuming that the result of any game is either a win or a loss, how many possible ways could the competition end with the same points table? In the formalism of tournaments, this problem is that of finding the number of tournaments with a given score vector and is a problem of considerable interest and difficulty. For example, since the conditional distribution of a tournament given its score vector is uniform in the β-model, this would provide a means of computing the probability of a tournament having a given score vector with respect to a β-model random tournament. With regards to the cricket analogy, this would allow one to predict the probability of a particular final points table, based on the rankings of each team.

In this chapter, we consider the more tractable problem of enumerating the number of tournaments with a given score vector asymptotically. Our main result is an asymptotic formula (Theorem 5.4.1) that can be applied to compute the number of tournaments with any δ-tame score vector. In Section 5.6, we prove a number corollaries of this formula, including finding the asymptotic numbers of regular tournaments. We then outline some possible future areas of research in Section 5.7.

We assume familiarity with probability theory and some complex analysis, including complex random variables.

5.1.2 Notation

Notation 5.1.1. Given a complex random variable $\boldsymbol{X}$, we will denote by $\mathbb{V}\boldsymbol{X}$ the *pseudovariance* of $\boldsymbol{X}$, so that

$$\mathbb{V}\boldsymbol{X} = \mathbb{E}\left(\boldsymbol{X} - \mathbb{E}\,\boldsymbol{X}\right)^2.$$

5.1.3 Approach

The main approach to finding an asymptotic formula (Theorem 5.4.1) is as follows:

1. Note that the number of tournaments with a particular score vector is just the coefficient of the monomial $x_1^{s_1} \cdots x_n^{s_n}$ in the generating function $\prod_{1 \le j < k \le n}(x_j + x_k)$.

2. Isolate this coefficient by applying Cauchy's Theorem for n dimensional complex integrals.

3. Observe that the integral inside a small box around the origin is the expectation of a function of the form $e^{f(x)}$ with respect to a Gaussian measure (where $f(x)$ is a function well approximated by a polynomial). Apply methods of Isaev and McKay in [IM16] to approximate this integral.(Subsection 5.5.1)

4. Show that the integral is negligible elsewhere, using a similar integral approximation, and an averaging argument. (Subsection 5.5.2)

Dependencies

Steps 3 and 4 are dependent on the results related to *weighted Laplacian tournament matrices* in Subsection 5.3.2. These steps are also highly dependent on the results from the paper by Isaev and McKay [IM16], and we include the required results from this paper in Appendix A.

The results of Section 5.6 are, for the most part, dependent only on Theorem 5.4.1 and Subsection 5.3.1.

5.2 Review of Literature

The first major results related to the asymptotic enumeration of tournaments with a given score vector appear to have been obtained by Spencer in 1972 [Spe74].
Let $\boldsymbol{r}$ denote the score vector of a regular tournament, that is,

$$\boldsymbol{r} = \begin{cases} \left(\frac{n-1}{2}, \cdots, \frac{n-1}{2}\right), & n \text{ odd}, \\ \left(\frac{n-2}{2}, \cdots, \frac{n-2}{2}, \frac{n}{2}, \cdots, \frac{n}{2}\right), & n \text{ even}. \end{cases}$$

Also, recall that RT_n denotes the set of regular tournaments on n vertices, that is, if $m \in \mathbb{N}$

$$|RT_n| = \begin{cases} |\mathbb{T}(\boldsymbol{r})|, & n = 2m+1, \\ \binom{n}{m} |\mathbb{T}(\boldsymbol{r})|, & n = 2m. \end{cases}$$

In ([Spe74], 1974) Spencer showed that as n goes to infinity

$$\left(\frac{|RT_n|}{2^{\binom{n}{2}}}\right)^{\frac{1}{n}} = \begin{cases} \sqrt{\frac{2}{n\pi}}\left(1 + o(1)\right), & n \text{ odd}, \\ 2\sqrt{\frac{2}{n\pi}}\left(1 + o(1)\right), & n \text{ even}. \end{cases}$$

Recall that $T^d(\boldsymbol{\delta})$ is the set of all tournaments with excess vector $\boldsymbol{\delta} = (\delta_1, .., \delta_n)$. For example, if n is odd, then $T^d(\mathbf{0}) = \mathbb{T}(\boldsymbol{r})$, with $\boldsymbol{r}$ as defined above. In the same paper, Spencer obtained the estimate that

$$T^d(\boldsymbol{\delta}) = |RT_n| \exp\left(-(2+o(1)) \frac{\sum_{i=1}^n \delta_i^2}{n} \right).$$

for tournaments where the maximum score is not too far from $\frac{n}{2}$. In 1990, McKay obtained a more precise estimate of $|RT_n|$. In particular, he showed that for any $\varepsilon > 0$ and odd n sufficiently large:

$$|RT_n| = \left(1 + O\left(n^{-\frac{1}{2}+\varepsilon}\right)\right) \left(\frac{2^{n+1}}{\pi n}\right)^{\frac{n-1}{2}} n^{\frac{1}{2}} e^{-\frac{1}{2}}. \tag{5.1}$$

McKay and Wang, in 1993, extended these results to obtain a formula which approximates $T^d(\boldsymbol{\delta})$ when each $\delta_i = O\left(n^{\frac{3}{4}+\varepsilon}\right)$ for sufficiently small $\varepsilon > 0$. A Corollary of their result showed that when each $\delta_i = o\left(n^{\frac{3}{4}}\right)$,

$$\begin{aligned} \left|T^d(\boldsymbol{\delta})\right| = \left(\frac{2^{n+1}}{\pi n}\right)^{\frac{n-1}{2}} n^{\frac{1}{2}} \exp\Bigg(&-\frac{1}{2} - \frac{\sum_{i=1}^n d_i^2}{2n} + \frac{1}{n^2}\sum_{i=1}^n \delta_j^2 - \frac{1}{12n^3}\sum_{i=1}^n \delta_i^4 - \frac{1}{4n^4}\left(\sum_{i=1}^n \delta_i^2\right) \\ &- \frac{1}{30n^5}\sum_{i=1}^n \delta_i^6 - \frac{1}{6n^6}\left(\sum_{i=1}^n \delta_i^3\right)^2 - \frac{1}{2n^7}\left(\sum_{i=1}^n \delta_i^2\right)^3 + O\left(\frac{\delta_{\max}^4}{n^3} + n^{\frac{1}{4}+\varepsilon}\right)\Bigg), \end{aligned} \tag{5.2}$$

for sufficiently small $\varepsilon > 0$. Note that by Theorem 2.4.7, almost all tournaments satisfy this property. Similar techniques were applied by Gao, McKay and Wang [GMW00] in their enumeration of tournaments containing a specified digraph (Theorem 2.4.5).

In what follows, we will use $\boldsymbol{w}$ to represent the vector $(1, 1, \cdots, 1) \in \mathbb{R}^n$. We will abuse notation slightly, so that $\boldsymbol{w}$ denotes the vector of ones in dimensions smaller than n as well (the dimension will be clear from the context).

5.3 Some Useful Results About Some Matrices

In this section we derive some identities involving special types of matrices that will be useful later on.

5.3.1 Perturbed Identity Matrices

The first type of matrix we will consider, for convenience we will give a name:

Definition 5.3.1. An $(a, b)-$*perturbed identity matrix* is a matrix of the form

$$aI + b\boldsymbol{w}\boldsymbol{w}^T.$$

Lemma 5.3.2. *For all $a \in \mathbb{R} - \{0\}$, $b \neq -\frac{a}{n} \in \mathbb{R}$, every $(a,b)-$perturbed identity matrix is invertible, and moreover has an inverse given by*

$$\frac{1}{a}I - \frac{b}{a^2+abn}\boldsymbol{w}\boldsymbol{w}^T.$$

Proof. We have

$$\begin{aligned}\left(aI + b\boldsymbol{w}\boldsymbol{w}^T\right)\left(\frac{1}{a}I - \frac{b}{a^2+abn}\boldsymbol{w}\boldsymbol{w}^T\right) &= I + \frac{b}{a}\boldsymbol{w}\boldsymbol{w}^T - \frac{ab}{a^2+abn}\boldsymbol{w}\boldsymbol{w}^T - \frac{b^2n}{a^2+abn}\boldsymbol{w}\boldsymbol{w}^T \\ &= I + \left(\frac{ab+b^2n}{a^2+abn} - \frac{ab}{a^2+abn} - \frac{b^2n}{a^2+abn}\right)\boldsymbol{w}\boldsymbol{w}^T \\ &= I.\end{aligned}$$

□

Lemma 5.3.3. *The eigenvalues of an (a,b)-perturbed identity matrix are given by a, which has multiplicity $n-1$ and $nb+a$.*

Proof. It is clear that $\boldsymbol{w}$ is an eigenvector with eigenvalue $a+nb$, and any vector orthogonal to $\boldsymbol{w}$ is an eigenvector with eigenvalue a. □

Corollary 5.3.4. *The determinant of an (a,b)-perturbed identity matrix is*

$$a^{n-1}\left(nb+a\right).$$

Proof. This follows from the fact that the determinant is the product of the eigenvalues. □

5.3.2 Weighted Laplacian Tournament Matrices

The next type of matrix we will find useful later on, we will call *weighted Laplacian tournament matrices* .

Let W be an $n \times n$ symmetric matrix with components w_{ij} such that there exist $c_1, c_2 > 0$ with

$$\begin{cases} c_1 \leq w_{jk} \leq c_2, & j \neq k; \\ w_{jk} = 0, & j = k. \end{cases} \tag{5.3}$$

Definition 5.3.5. We define the *W-Weighted Laplacian Tournament Matrix* as the $n \times n$ matrix P^W such that

$$P^W_{jk} = \begin{cases} -w_{jk}, & j \neq k; \\ \sum_{l=1}^n w_{jl}, & j = k. \end{cases} \tag{5.4}$$

Note that as an immediate consequence of the definition of W we have that

$$c_1(n-1) \leq P^W_{jj} \leq c_2 n, \quad 1 \leq j \leq n. \tag{5.5}$$

For the remainder of this subsection we denote by κ a fixed constant greater that 0.

Lemma 5.3.6. *The matrix $P^W + \kappa \boldsymbol{w}\boldsymbol{w}^T$ has $n-1$ eigenvalues η_i (counting multiplicity) such that*

$$c_1 n \leq \eta_i \leq c_2 n, \tag{5.6}$$

and one eigenvalue equal to $n\kappa$. In particular, if $c_3 = \min(c_1, \kappa)$, then all of the eigenvalues of this matrix are bounded below by $c_3 n$.

Proof. Note that P^W and $\boldsymbol{w}\boldsymbol{w}^T$ are symmetric matrices, thus have real eigenvalues. Now, for $\boldsymbol{v} \in \mathbb{R}^n$ we have

$$\begin{aligned} c_1 \sum_{j<k} (v_j - v_k)^2 &\leq \boldsymbol{v}^T P^W \boldsymbol{v} \\ &= \sum_{j<k} w_{jk} (v_j - v_k)^2 \\ &\leq c_2 \sum_{j<k} (v_j - v_k)^2. \end{aligned}$$

If S is the symmetric matrix corresponding to the quadratic form

$$\boldsymbol{v}^T S \boldsymbol{v} = \sum_{j<k} (v_j - v_k)^2,$$

a routine calculation shows that all eigenvalues of S are n, except for one zero eigenvalue corresponding to the eigenvector $\boldsymbol{w}$. It follows that the eigenvalues of P^W are all bounded below by $c_1 n$ and above by $c_2 n$, except for the 0 eigenvalue corresponding to $\boldsymbol{w}$. If $\boldsymbol{y}$ is an eigenvector of P^W corresponding to an eigenvalue $k > 0$, then since $\boldsymbol{y}$ is orthogonal to $\boldsymbol{w}$, we have

$$\left(P^W + \kappa \boldsymbol{w}\boldsymbol{w}^T\right) \boldsymbol{y} = k\boldsymbol{y} \geq c_1 n \boldsymbol{y},$$

and similarly

$$\left(P^W + \kappa \boldsymbol{w}\boldsymbol{w}^T\right) \boldsymbol{y} = k\boldsymbol{y} \leq c_2 n \boldsymbol{y}.$$

Otherwise, if $\boldsymbol{y}$ is the eigenvector of P^W in the direction of $\boldsymbol{w}$, then

$$\left(P^W + \kappa \boldsymbol{w}\boldsymbol{w}^T\right) \boldsymbol{y} = \kappa n \boldsymbol{y}.$$

The result follows. □

Corollary 5.3.7. *We have $c_1^{n-1} \kappa n^n \leq \det(P^W + \kappa \boldsymbol{w}\boldsymbol{w}^T) \leq c_2^{n-1} \kappa n^n$.*

Lemma 5.3.8. *Let W_1 be the matrix obtained from W by deleting the first row and the first column. Then there is a constant $c > 0$ depending only on c_1 and c_2 such that*

$$\det\left(P^{W_1} + \kappa \boldsymbol{w}\boldsymbol{w}^T\right) \geq \frac{\det\left(P^W + \kappa \boldsymbol{w}\boldsymbol{w}^T\right)}{cn}. \tag{5.7}$$

Proof. If $P^W_{n-1}+\kappa\boldsymbol{w}\boldsymbol{w}^T$ is the matrix obtained by deleting the first row and the first column of $P^W+\kappa\boldsymbol{w}\boldsymbol{w}^T$, note that this matrix coincides with $P^{W_1}+\kappa\boldsymbol{w}\boldsymbol{w}^T$, with the exception of the diagonal entries. Define the $n\times n$ matrix X by

$$X_{jk}=\begin{cases}-w_{1j}+\frac{(\kappa-w_{1j})^2}{P^W_{11}+\kappa}, & \text{For } j=k\neq 1,\\ \frac{(\kappa-w_{1j})(\kappa-w_{1k})}{P^W_{11}+\kappa} & \text{For } j,k\neq 1, j\neq k,\\ 0, & \text{Otherwise.}\end{cases}$$

Then, after performing Gaussian elimination on the matrix $P^W+\kappa\boldsymbol{w}\boldsymbol{w}^T$ to form a pivot in the first column, we find that

$$\det\left(P^W+\kappa\boldsymbol{w}\boldsymbol{w}^T+X\right)=\left(1+P^W_{11}\right)\det\left(P^{W_1}+\kappa\boldsymbol{w}\boldsymbol{w}^T\right). \tag{5.8}$$

By equation (5.5), we have that

$$\|X\|_2\leq\|X\|_1\leq c_4, \tag{5.9}$$

for some $c_4>0$ depending only on c_1 and c_2. Moreover, by Lemma 5.3.6, we find that

$$\left\|X\left(P^W+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)^{-1}\right\|_2\leq\|X\|_2\left\|\left(P^W+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)^{-1}\right\|_2\leq\frac{c_4}{c_3n}. \tag{5.10}$$

Therefore,

$$\det\left(I+X\left(P^W+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)^{-1}\right)\geq\left(1-\frac{c_4}{c_3n}\right)^n. \tag{5.11}$$

By combining equations (5.5), (5.8) and (5.11), we obtain (5.7). □

Corollary 5.3.9. *Let W_r be the matrix obtained from W by deleting some r rows and the corresponding r columns. Then, for $r\leq n-1$ there is a constant $c>0$ depending only on c_1 and c_2 such that*

$$\det\left(P^{W_r}+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)\geq\frac{\det\left(P^W+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)}{(cn)^r}. \tag{5.12}$$

Proof. The result follows by applying Lemma 5.3.8 r times. □

5.4 The Main Theorem

In this chapter we apply the methods of Isaev and McKay in [IM16] to derive an asymptotic formula for the number of tournaments with a given score vector if the score vector is δ-tame.

Let $\boldsymbol{s}$ be a δ-tame score vector, for $\delta>0$. Then, by equations (4.1) and (4.4) there exist β_j satisfying

$$\sum_{k\neq j}\frac{e^{\beta_j}}{e^{\beta_j}+e^{\beta_k}}=s_j,\quad 1\leq j\leq n. \tag{5.13}$$

Suppose we approximate each β_j by $\tilde{\beta}_j$ so that so that, if for $1 \le j < k \le n$ we set

$$\lambda_{jk} = \frac{e^{\beta_j}}{e^{\beta_j} + e^{\beta_k}} \tag{5.14}$$

and

$$\tilde{\lambda}_{jk} = \frac{e^{\tilde{\beta}_j}}{e^{\tilde{\beta}_j} + e^{\tilde{\beta}_k}} \tag{5.15}$$

then

$$\sup_{j,k} \left| \tilde{\lambda}_{jk} - \lambda_{jk} \right| = O\left(n^{-2}\right). \tag{5.16}$$

Moreover, let

$$\tilde{P}(T) = \frac{e^{\tilde{\beta}_1 s_1 + \cdots + \tilde{\beta}_n s_n}}{\prod_{1\le j<k\le n}\left(e^{\tilde{\beta}_j} + e^{\tilde{\beta}_k}\right)} \tag{5.17}$$

approximate the probability of a tournament T with score vector $\boldsymbol{s}$ in the β-model.

Also, let

$$f(\boldsymbol{\theta}) = \sum_{j<k} -\frac{1}{6} i\tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk})(1-2\tilde{\lambda}_{jk})(\theta_j-\theta_k)^3 + \frac{1}{24}\tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk})\left(1-6\tilde{\lambda}_{jk}+6\tilde{\lambda}_{jk}^2\right)(\theta_j-\theta_k)^4,$$

A be the $n \times n$ symmetric matrix such that

$$\boldsymbol{\theta}^T A \boldsymbol{\theta} = \frac{1}{2}\sum_{j<k} \tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk})(\theta_j-\theta_k)^2, \tag{5.18}$$

and $\boldsymbol{x}$ be a random variable with density

$$\pi^{-\frac{n}{2}} \det\left(A + \kappa \boldsymbol{w}\boldsymbol{w}^T\right)^{\frac{1}{2}} e^{-\boldsymbol{x}^T\left(A+\boldsymbol{w}\boldsymbol{w}^T\right)\boldsymbol{x}}, \tag{5.19}$$

with $\kappa \ge \frac{1}{2}\delta(1-\delta)$ a constant. Our main result this section is then the following theorem:

Theorem 5.4.1. *For $n \to \infty$, if $\boldsymbol{s}$ is a δ-tame score vector then*

$$|\mathbb{T}(\boldsymbol{s})| = \left(1 + O\left(n^{-\frac{1}{2}+9\varepsilon}\right)\right) \frac{\sqrt{\kappa}\det\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)^{-\frac{1}{2}}}{\left(2\sqrt{\pi}\right)^{n-1}\tilde{P}(T)} e^{\mathbb{E}\left(f(\boldsymbol{x}) - \frac{1}{2}(\Im f(\boldsymbol{x}))^2\right)}, \tag{5.20}$$

for any $\varepsilon > 0$, where $\tilde{P}(T)$ is as given by (5.17).

The method we use is to estimate the number of tournaments with a given score vector by a contour integral over the complex plane. We note that the integral simplifies to computing the expectation of a quartic polynomial with respect to a Gaussian measure, and can be estimated asymptotically using the methods of Isaev and McKay [IM16].

5.5 The Contour Integral

Let T be a tournament on n vertices with a δ-tame score vector $(s_1, \cdots, s_n)$. Note that the number of tournaments with this score vector is given by the coefficient of $x_1^{s_1} x_2^{s_2} \cdots x_n^{s_n}$ in the generating function

$$\prod_{1 \le j < k \le n} (x_j + x_k).$$

This follows from the fact that if v_j and v_k are two vertices such that v_j dominates v_k, then the score of v_j increases by 1 whilst the score of v_k remains the same.

Now, by applying Cauchy's theorem for n dimensional complex integrals ([Kni96], Theorem 3.1), we see that the value of this coefficient is given by

$$\frac{1}{(2\pi i)^n} \oint \cdots \oint \frac{\prod_{1 \le j < k \le n} (x_j + x_k)}{x_1^{s_1+1} \cdots x_n^{s_n+1}} dx_1 dx_2 \cdots dx_n.$$

Since the pole is at the origin, we choose circular contours centred at the origin. By making the substitution

$$x_j = e^{\tilde{\beta}_j + i\theta_j} \tag{5.21}$$

for each $1 \le j \le n$, we have

$$\frac{dx_j}{d\theta_j} = i e^{\tilde{\beta}_j + i\theta_j},$$

so that the integral reduces to

$$\begin{aligned} \frac{1}{(2\pi)^n} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \frac{\prod_{1 \le j < k \le n} (e^{\tilde{\beta}_j + i\theta_j} + e^{\tilde{\beta}_k + i\theta_k})}{e^{\tilde{\beta}_1 s_1 + \cdots + \tilde{\beta}_n s_n} e^{i(s_1\theta_1 + \cdots + s_n\theta_n)}} d\theta_1 \cdots d\theta_n \\ = \frac{\prod_{1 \le j < k \le n} (e^{\tilde{\beta}_j} + e^{\tilde{\beta}_k})}{(2\pi)^n e^{\tilde{\beta}_1 s_1 + \cdots + \tilde{\beta}_n s_n}} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} F(\boldsymbol{\theta}) d\boldsymbol{\theta} \\ = \frac{1}{(2\pi)^n \, \tilde{P}(T)} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} F(\boldsymbol{\theta}) d\boldsymbol{\theta}, \end{aligned} \tag{5.22}$$

where

$$F(\boldsymbol{\theta}) = \prod_{1 \le j < k \le n} \left(\frac{e^{\tilde{\beta}_j + i\theta_j} + e^{\tilde{\beta}_k + i\theta_k}}{e^{\tilde{\beta}_j} + e^{\tilde{\beta}_k}} \right) / e^{i(s_1\theta_1 + \cdots + s_n\theta_n)}. \tag{5.23}$$

Let $\rho > 0$ and let

$$U_n(\rho) := \{ \boldsymbol{x} \in \mathbb{R}^n : \, |x_j| \le \rho \text{ for } 1 \le j \le n \}$$

denote the n dimensional cube with side length 2ρ. Then, we are left with estimating the integral

$$\int_{U_n(\pi)} F(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{5.24}$$

The remainder of this section will deal with solving this problem.

5.5.1 The Main Contribution of the Integral

The first thing to note when evaluating equation (5.24) is that, since $\boldsymbol{s}$ is a score vector, for any constant k we have

$$\sum_{i=1}^{n} iks_i = ik\binom{n}{2},$$

which implies the symmetry

$$F(\boldsymbol{\theta}) = F(\boldsymbol{\theta} + k\boldsymbol{w}). \tag{5.25}$$

Moreover, $F(\boldsymbol{\theta})$ is invariant upon adding 2π to any variable, so depends only on the value of each variable modulo 2π. This motivates the definition of the pseudometric $|.|_\pi$: $\mathbb{R} \to [0, \infty)$ such that

$$|x|_\pi = \min_{K\in\mathbb{Z}} |x + 2\pi K|. \tag{5.26}$$

For $\varepsilon > 0$, let Ω be the set

$$\left\{\boldsymbol{\theta} \in \mathbb{R}^n : \ |\theta_j - \theta_n|_\pi \le n^{-\frac{1}{2}+\varepsilon} \quad 1 \le j \le n\right\}.$$

As we will show in the next subsection, the dominant part of the integral comes from Ω, so that

$$\int_{U_n(\pi)} F(\boldsymbol{\theta})d\boldsymbol{\theta} = \left(1 + O\left(e^{-b'n^{2\varepsilon}}\right)\right)\int_\Omega F(\boldsymbol{\theta})d\boldsymbol{\theta}, \tag{5.27}$$

for $b' > 0$ a constant.
Let Q and W be matrices such that for $\boldsymbol{y} \in \mathbb{R}^n$

$$Q\boldsymbol{y} = \boldsymbol{y} - y_n\boldsymbol{w}$$

and

$$W\boldsymbol{y} = \sqrt{\frac{\kappa}{n}}(y_1 + \cdots + y_n)\boldsymbol{w},$$

with $\kappa \ge \frac{\delta(1-\delta)}{2}$. Note that by equation (5.25) we have $F(Q\boldsymbol{\theta}) = F(\boldsymbol{\theta})$, so that

$$\int_\Omega F(\boldsymbol{\theta})d\boldsymbol{\theta} = \int_\Omega F(Q\boldsymbol{\theta})d\boldsymbol{\theta}.$$

Now, the integral on the right does not depend on θ_n, which can range over 2π values in Ω. We therefore have

$$\int_\Omega F(\boldsymbol{\theta})d\boldsymbol{\theta} = 2\pi\int_{\Omega\cap Q(\mathbb{R}^n)} F(\boldsymbol{x})d\boldsymbol{x}.$$

Noting that $\det\left(Q^TQ + W^TW\right) = \kappa n^2$, we apply Lemma 4.6 of [IM16] (Lemma A.0.4) with $\rho = \sqrt{\kappa}n^{\varepsilon}$ to find that

$$2\pi \int_{\Omega \cap Q(\mathbb{R}^n)} F(\boldsymbol{x})d\boldsymbol{x} = \left(1 + O(e^{-c'n^{2\varepsilon}})\right) 2\sqrt{\pi\kappa}n \int_{\Omega_\rho} F(\boldsymbol{\theta})e^{-\boldsymbol{\theta}^T W^T W \boldsymbol{\theta}} d\boldsymbol{\theta}, \tag{5.28}$$

for some $c' > 0$, where Ω_p is such that

$$\Omega_p = \{\boldsymbol{x} \in \mathbb{R}^n : Q\boldsymbol{x} \in \Omega \text{ and } W\boldsymbol{x} \in U_n(\rho)\}. \tag{5.29}$$

Lemma 5.5.1. *The set Ω_ρ is between two boxes of side length $O\left(n^{-\frac{1}{2}+\varepsilon}\right)$, in particular*

$$U_n\left(\frac{1}{2}n^{-\frac{1}{2}+\varepsilon}\right) \subseteq \Omega_\rho \subseteq U_n\left(3n^{-\frac{1}{2}+\varepsilon}\right).$$

Proof. For the first inclusion, note that if $\boldsymbol{x} \in U_n\left(\frac{1}{2}n^{-\frac{1}{2}+\varepsilon}\right)$, then

$$\sup_{j<k\in[n]} |x_j - x_k| \le n^{-\frac{1}{2}+\varepsilon},$$

so that $Q\boldsymbol{x} \in \Omega$ and $W\boldsymbol{x} \in U_n(\rho)$.

For the second inclusion, if $\boldsymbol{x} \in \Omega_\rho$, then $\boldsymbol{x}$ satisfies

$$\sup_{j<k\in[n]} |x_j - x_k| \le 2n^{-\frac{1}{2}+\varepsilon}. \tag{5.30}$$

Suppose the $\boldsymbol{x} \notin U_n\left(3n^{-\frac{1}{2}+\varepsilon}\right)$. Then for some $j \in [n]$ either $x_j > 3n^{-\frac{1}{2}+\varepsilon}$ or $x_j < -3n^{-\frac{1}{2}+\varepsilon}$. Assume that for some $x_j > 3n^{-\frac{1}{2}+\varepsilon}$. Then (5.30) implies that for all $k \neq j$ $x_k > n^{-\frac{1}{2}+\varepsilon}$, which implies that $W\boldsymbol{x} \notin U_n(\rho)$. The case where $x_j < -3n^{-\frac{1}{2}+\varepsilon}$ is similar. □

Now, by (5.15), we have

$$\begin{aligned} F(\boldsymbol{\theta})e^{i(s_1\theta_1+\cdots+s_n\theta_n)} &= \prod_{j<k}\left(\tilde{\lambda}_{jk}e^{i\theta_j} + (1-\tilde{\lambda}_{jk})e^{i\theta_k}\right) \\ &= \prod_{j<k} e^{i\theta_k} \prod_{j<k}\left(1 + \tilde{\lambda}_{jk}\left(e^{i(\theta_j-\theta_k)} - 1\right)\right). \end{aligned} \tag{5.31}$$

By Taylor's Theorem, for $\boldsymbol{\theta} \in \Omega_\rho$

$$\begin{aligned} \log\left(1 + \tilde{\lambda}_{jk}\left(e^{i(\theta_j-\theta_k)} - 1\right)\right) &= i\tilde{\lambda}_{jk}(\theta_j-\theta_k) - \frac{1}{2}\tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk})(\theta_j-\theta_k)^2 \\ &\quad - \frac{1}{6}i\tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk})(1-2\tilde{\lambda}_{jk})(\theta_j-\theta_k)^3 \\ &\quad + \frac{1}{24}\tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk})\left(1 - 6\tilde{\lambda}_{jk} + 6\tilde{\lambda}_{jk}^2\right)(\theta_j-\theta_k)^4 \\ &\quad + O\left(n^{-\frac{5}{2}+5\varepsilon}\right) \qquad 1 \le j < k \le n. \end{aligned}$$

Thus

$$F(\boldsymbol{\theta})e^{i(s_1\theta_1+\cdots+s_n\theta_n)} = \exp\left(\sum_{j<k}\left(\tilde{\lambda}_{jk}\theta_j + (1-\tilde{\lambda}_{jk})\theta_k\right) - \boldsymbol{\theta}^T A\boldsymbol{\theta} + f(\boldsymbol{\theta}) + O\left(n^{-\frac{1}{2}+5\varepsilon}\right)\right) \tag{5.32}$$

where

$$f(\boldsymbol{\theta}) = \sum_{j<k} -\frac{1}{6}i\tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk})(1-2\tilde{\lambda}_{jk})(\theta_j-\theta_k)^3 + \frac{1}{24}\tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk})\left(1-6\tilde{\lambda}_{jk}+6\tilde{\lambda}_{jk}^2\right)(\theta_j-\theta_k)^4,$$

and A is the $n\times n$ symmetric matrix such that

$$\boldsymbol{\theta}^T A\boldsymbol{\theta} = \frac{1}{2}\sum_{j<k}\tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk})(\theta_j-\theta_k)^2. \tag{5.33}$$

Now, by (5.13) and (5.16)

$$\begin{aligned}\exp\left(i\sum_{j<k}\left(\tilde{\lambda}_{jk}\theta_j + (1-\tilde{\lambda}_{jk})\theta_k\right)\right) &= \exp\left(i\sum_{j<k}\left(\tilde{\lambda}_{jk}\theta_j + \tilde{\lambda}_{kj}\theta_k\right)\right)\\ &= \exp\left(i\sum_{j=1}^{n}\left(\sum_{k\neq j}\tilde{\lambda}_{jk}\right)\theta_j\right)\\ &= \left(1+O\left(n^{-\frac{1}{2}+\varepsilon}\right)\right)e^{i(s_1\theta_1+\cdots+s_n\theta_n)};\end{aligned} \tag{5.34}$$

so that

$$\begin{aligned}&\int_{\Omega_\rho} F(\boldsymbol{\theta})e^{-\boldsymbol{\theta}^T W^T W\boldsymbol{\theta}}d\boldsymbol{\theta}\\ &= \left(1+O\left(n^{-\frac{1}{2}+\varepsilon}\right)\right)\int_{\Omega_\rho}\exp\left(-\boldsymbol{\theta}^T(A+\kappa\boldsymbol{w}\boldsymbol{w}^T)\boldsymbol{\theta} + f(\boldsymbol{\theta}) + O\left(n^{-\frac{1}{2}+5\varepsilon}\right)\right)d\boldsymbol{\theta}.\end{aligned}$$

The following sequence of Lemmas allow us to apply the main results of [IM16] to approximate this integral.

Note that A is a Λ-Weighted Laplacian Tournament matrix, with Λ such that

$$\Lambda_{jk} = \frac{1}{2}\tilde{\lambda}_{jk}(1-\tilde{\lambda}_{jk}) \tag{5.35}$$

since by the condition of δ-tameness, for n sufficiently large,

$$\frac{1}{2}\delta(1-\delta) \leq \tilde{\lambda}_{jk} \leq \frac{1}{8}. \tag{5.36}$$

Lemma 5.5.2. *Suppose D is the diagonal matrix with the same diagonal as A. Then for some constant a_1 we have*

$$\left\|\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)^{-1} - D^{-1}\right\|_{\max} \leq a_1 n^{-2}.$$

Moreover, there exists a matrix T with $T^T\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)T = I$ and some constants a_2 and a_3 such that $\|T\|_1$, $\|T\|_\infty \leq a_2 n^{-\frac{1}{2}}$ and $\|T^{-1}\|_\infty \leq a_3 n^{\frac{1}{2}}$.

Proof. By equation (5.36), we have $\|A - D\|_{\max} \leq \frac{1}{8}$. Also, by Lemma 5.3.6, since $\kappa \geq \frac{1}{2}\delta(1-\delta)$, for $\boldsymbol{x} \in \mathbb{R}^n$ we have

$$\boldsymbol{x}^T \left(A + \kappa \boldsymbol{w}\boldsymbol{w}^T\right) \boldsymbol{x} \geq \frac{1}{2}\delta(1-\delta)n\boldsymbol{x}^T\boldsymbol{x}.$$

Noting that

$$\max D_{jj} \leq \frac{1}{8}(n-1) \text{ and } \min D_{jj} \geq \frac{1}{2}\delta(1-\delta)(n-1),$$

we apply Lemma 4.9 of [IM16] (Lemma A.0.5) with $r = \frac{n}{4\delta(1-\delta)(n-1)}$ and $\gamma = 4\delta(1-\delta)\frac{n}{n-1}$ to finish the proof (Note that the condition $\gamma \leq 1$ in Lemma 4.9 is not needed). □

Let $g(\boldsymbol{\theta}) = \Re f(\boldsymbol{\theta})$, and set $\rho_1, \rho_2 = O(n^{\varepsilon})$.

Now, for $\boldsymbol{\theta} \in \Omega_p$, we have

1. $|f_j(\boldsymbol{\theta})| = O(n^{2\varepsilon})$
2. $\|H(f, T(U_n(\rho_1)))\|_\infty = O(n^{\frac{1}{2}+\varepsilon})$
3. $|g_j(\boldsymbol{\theta})| = O(n^{-\frac{1}{2}+3\varepsilon})$
4. $\|H(g, T(U_n(\rho_2)))\|_\infty = O(n^{2\varepsilon})$.

Let $\boldsymbol{x}$ be a random variable with density

$$\pi^{-\frac{n}{2}} \det \left(A + \kappa \boldsymbol{w}\boldsymbol{w}^T\right)^{\frac{1}{2}} e^{-\boldsymbol{x}^T\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)\boldsymbol{x}}. \tag{5.37}$$

Then by applying Theorem 4.4 (Theorem A.0.2) and Remark 4.5 (Remark A.0.3) of [IM16] with $\phi_1 = O(n^{-\frac{1}{6}+3\varepsilon})$, $\phi_2 = O(n^{-\frac{2}{3}+4\varepsilon})$ and $h(\boldsymbol{x}) = O\left(n^{-\frac{5}{2}+5\varepsilon}\right)$, we get

$$\begin{aligned}\int_{\Omega_\rho} &\exp\left(-\boldsymbol{x}^T(A+\kappa\boldsymbol{w}\boldsymbol{w}^T)\boldsymbol{x} + f(\boldsymbol{x}) + O\left(n^{-\frac{5}{2}+5\varepsilon}\right)\right)d\boldsymbol{x} \\ &= (1+K)\pi^{\frac{n}{2}} \det\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)^{-\frac{1}{2}} e^{\mathbb{E}\, f(\boldsymbol{x}) + \frac{1}{2}\mathbb{V} f(\boldsymbol{x})},\end{aligned} \tag{5.38}$$

where $|K| \leq Ce^{\frac{1}{2}\operatorname{Var}\Im f(\boldsymbol{x})}\left(e^{O\left(n^{-\frac{1}{2}+9\varepsilon}\right)} - 1\right)$.

Lemma 5.5.3. *We have*

$$\mathbb{E}\, f(\boldsymbol{x}) = \mathbb{E}\,\Re f(\boldsymbol{x}) = O(1) \tag{5.39}$$

and

$$\begin{aligned}\mathbb{V} f(\boldsymbol{x}) &= \operatorname{Var}\Re f(\boldsymbol{x}) - \operatorname{Var}\Im f(\boldsymbol{x}) \\ &= \operatorname{Var}\Re f(\boldsymbol{x}) - \mathbb{E}\left(\Im f(\boldsymbol{x})\right)^2 \\ &= O\left(1\right).\end{aligned} \tag{5.40}$$

Proof. Let $\boldsymbol{x} = (X_1, \cdots, X_n)$ be the random variable defined by equation (5.37). If we denote the entries of the covariance matrix $\left(2\left(A + \kappa \boldsymbol{w}\boldsymbol{w}^T\right)\right)^{-1}$ by σ_{jk}, Lemma 5.5.2 implies that, $\sigma_{jj} = O\left(n^{-1}\right)$ and $\sigma_{jk} = O\left(n^{-2}\right)$ for all $j \neq k$. By applying Theorem 4.2 of [IM16] (Theorem A.0.1), the expected value of any odd monomial is 0, and for $p, q \in \mathbb{N}$ and $j \neq k$, $j' \neq k'$, we find that

$$\mathbb{E}\left(X_j - X_k\right)^{2p} = O\left(n^{-p}\right);$$

$$\mathrm{Cov}\left(\left(X_j - X_k\right)^{2p+1}, \left(X_{j'} - X_{k'}\right)^{2q+1}\right) = \begin{cases} O\left(n^{-p-q-1}\right), & \text{if } \{j,k\} \cap \{j',k'\} \neq \emptyset, \\ O\left(n^{-p-q-2}\right), & \text{otherwise;} \end{cases}$$

$$\mathrm{Cov}\left(\left(X_j - X_k\right)^{2p}, \left(X_{j'} - X_{k'}\right)^{2q}\right) = \begin{cases} O\left(n^{-p-q}\right), & \text{if } \{j,k\} \cap \{j',k'\} \neq \emptyset, \\ O\left(n^{-p-q-2}\right), & \text{otherwise.} \end{cases}$$

Equations (5.39) and (5.40) then follow from the above, using the useful identity

$$\mathrm{Var}\left(\sum_j R_j\right) = \sum_{j,k} \mathrm{Cov}\left(R_j, R_k\right)$$

to show

$$\mathrm{Var}\, \Re f(\boldsymbol{x}) = O\left(n^{-1}\right)$$

and

$$\mathrm{Var}\, \Im f(\boldsymbol{x}) = \mathbb{E}\left(\Im f(\boldsymbol{x})\right)^2 = O(1)$$

(since $\mathbb{E}\, \Im f(\boldsymbol{x}) = 0$). □

Putting all this together we get

$$\begin{aligned} &\int_{\Omega_\rho} \exp\left(-\boldsymbol{x}^T(A + \kappa \boldsymbol{w}\boldsymbol{w}^T)\boldsymbol{x} + f(\boldsymbol{x}) + O\left(n^{-\frac{5}{2}+5\varepsilon}\right)\right) d\boldsymbol{x} \\ &\quad = \left(1 + O\left(n^{-\frac{1}{2}+9\varepsilon}\right)\right) \pi^{\frac{n}{2}} \det\left(A + \kappa \boldsymbol{w}\boldsymbol{w}^T\right)^{-\frac{1}{2}} e^{\mathbb{E}\left(f(\boldsymbol{x}) - \frac{1}{2}(\Im f(\boldsymbol{x}))^2\right)}. \end{aligned} \tag{5.41}$$

By combining (5.41), (5.28) and (5.22) we get the main result of the chapter ((5.20)). The only gap left to fill is to prove Proposition 5.5.4, which shows that the rest of the integral is negligible.

5.5.2 The Negligible Part of the Integral

In this sub-section we will prove the following Proposition:

Proposition 5.5.4. *For $\varepsilon > 0$, and $b' > 0$ dependent only on δ,*

$$\int_{U_n(\pi)} F(\boldsymbol{\theta}) d\boldsymbol{\theta} = \left(1 + O\left(e^{-b' n^{2\varepsilon}}\right)\right) \int_{\Omega} F(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{5.42}$$

Proof. Note that

$$\left|\int_{U_n(\pi)} F(\boldsymbol{\theta})d\boldsymbol{\theta} - \int_{\Omega} F(\boldsymbol{\theta})d\boldsymbol{\theta}\right| = \left|\int_{U_n(\pi)-\Omega} F(\boldsymbol{\theta})d\boldsymbol{\theta}\right| \leq \int_{U_n(\pi)-\Omega} |F(\boldsymbol{\theta})|\, d\boldsymbol{\theta}.$$

We will show that the right side of this equation is small compared to the dominant term of (5.38). By applying Corollary 5.3.7, we see that this dominant term satisfies

$$\pi^{\frac{n}{2}} \det\left(A + \kappa \boldsymbol{w}\boldsymbol{w}^T\right)^{-\frac{1}{2}} \geq \kappa^{-1} 2\pi^{\frac{n}{2}} 8^{\frac{n-1}{2}} n^{-\frac{n}{2}}. \tag{5.43}$$

We start by observing that

$$|F(\boldsymbol{\theta})| = \prod_{j<k} \left|1 + \tilde{\lambda}_{jk}\left(e^{i(\theta_j - \theta_k)} - 1\right)\right|.$$

Lemma 5.5.5. *For all* $0 \leq \lambda \leq 1$ *and* $x \in \mathbb{R}$

$$\begin{aligned}\left|1 + \lambda\left(e^{ix} - 1\right)\right| &= \sqrt{1 - 2\lambda(1-\lambda)(1-\cos x)} \\ &\leq \exp\left(-\frac{1}{2}\lambda(1-\lambda)x^2 + \frac{1}{24}\lambda(1-\lambda)x^4\right).\end{aligned}$$

Proof. This follows from comparing the functions using Taylor's Theorem. □

For $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n) \in U_n(\pi)$ suppose that there are m pairs $\theta_j, \theta_k,\ i < j$, such that

$$|\theta_j - \theta_k|_\pi > \frac{1}{2} n^{-\frac{1}{2}+\varepsilon}. \tag{5.44}$$

Then, if B is the set of all such (unordered) pairs

$$\begin{aligned}|F(\boldsymbol{\theta})| &= \prod_{j<k} \sqrt{1 - 2\tilde{\lambda}_{jk}\left(1 - \tilde{\lambda}_{jk}\right)(1 - \cos(\theta_j - \theta_k))} \\ &\leq \prod_{(j,k)\in B} \sqrt{1 - 2\tilde{\lambda}_{jk}\left(1 - \tilde{\lambda}_{jk}\right)\left(1 - \cos\left(\frac{1}{2}n^{-\frac{1}{2}+\varepsilon}\right)\right)} \\ &\leq \exp\left(-\delta(1-\delta)\frac{11}{96} m n^{-1+2\varepsilon}\right).\end{aligned} \tag{5.45}$$

Suppose $m \leq n^{2-\varepsilon}$. Then, by an averaging argument, there exists some θ_s such that the number of $\theta_j,\ j \neq s$ with

$$|\theta_s - \theta_j|_\pi > \frac{1}{2} n^{-\frac{1}{2}+\varepsilon}$$

is at most $n^{1-\varepsilon}$.

It then follows that the set $U_n(\pi) - \Omega$ is covered by the set

$$M := \left\{\boldsymbol{\theta} \in U_n(\pi) : \text{the number of pairs satisfying (5.44) is more than } n^{2-\varepsilon}\right\},$$

together with the collection of sets

$$\Theta_r := \left\{ \boldsymbol{\theta} \in U_n(\pi) : \exists\, \theta_s \text{ such that } \begin{array}{ll} |\theta_s - \theta_j| > n^{-\frac{1}{2}+\varepsilon}, & \text{for exactly } r \text{ values of } j, \\ |\theta_s - \theta_j| \le \frac{1}{2} n^{-\frac{1}{2}+\varepsilon}, & \text{otherwise.} \end{array} \right\},$$

with $1 \le r \le n^{1-\varepsilon}$.

By (5.45), we have

$$\int_M |F(\boldsymbol{\theta})|\, d\boldsymbol{\theta} \le (2\pi)^n \exp\left(-\delta\,(1-\delta)\,\frac{11}{96} n^{1+\varepsilon}\right) = \exp\left(-bn^{1+\varepsilon}\right). \tag{5.46}$$

for some $b > 0$. For n large, this is negligible compared to the dominant term in (5.43).

Now if $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n) \in \Omega_r$, and I is a subset of $[n]$ such that $|I| = r$, $r \le n^{1-\varepsilon}$ define $G_I(\boldsymbol{\theta})$ such that

$$G_I(\boldsymbol{\theta}) = \prod_{\substack{j<k, \\ j,k \in [n]-I}} \exp\left(-\frac{1}{2}\tilde{\lambda}_{jk}\,(\theta_j - \theta_k)^2 + \frac{1}{24}\tilde{\lambda}_{jk}\,(\theta_j - \theta_k)^4\right).$$

We then have that, by (5.45)

$$\int_{\Theta_r} |F(\boldsymbol{\theta})|\, d\boldsymbol{\theta} \le (2\pi)^r e^{-grn^{2\varepsilon}} \sum_{I \subset [n]} \int_{U_{n-r}\left(\frac{1}{2} n^{-\frac{1}{2}+\varepsilon}\right)} G_I(\boldsymbol{\theta}), \tag{5.47}$$

where $g > 0$ depends only on δ. For each I, define the matrix A_I such that

$$\boldsymbol{\theta}^T A_I \boldsymbol{\theta} = \sum_{\substack{j<k, \\ j,k \in [n]-I}} -\frac{1}{2}\tilde{\lambda}_{jk}\left(1 - \tilde{\lambda}_{jk}\right)(\theta_j - \theta_k)^2.$$

Now, by a very similar approach to the evaluation of the integral in the previous subsection (see (5.38)), we find that

$$\int_{U_{n-r}\left(\frac{1}{2} n^{-\frac{1}{2}+\varepsilon}\right)} G_I(\boldsymbol{\theta}) = O(1)\, \pi^{\frac{n-r}{2}} \det\left(A_I + \kappa \boldsymbol{w}\boldsymbol{w}^T\right)^{-\frac{1}{2}}, \qquad r \le n^{1-\varepsilon}.$$

Observe that each A_I is a matrix obtained from the Λ-Weighted Laplacian Tournament matrix A by deleting some r rows and the corresponding r columns. Applying Corollary 5.3.9, we find that

$$\int_{U_{n-r}\left(\frac{1}{2} n^{-\frac{1}{2}+\varepsilon}\right)} G_I(\boldsymbol{\theta}) = O(1)\, \pi^{-\frac{r}{2}} (cn)^{-\frac{r}{2}} \pi^{\frac{n}{2}} \det\left(A + \kappa \boldsymbol{w}\boldsymbol{w}^T\right)^{-\frac{1}{2}},$$

where c depends only on δ. Substituting this in equation (5.47), we find

$$\begin{aligned}\int_{\Theta_r}|F(\boldsymbol{\theta})|\,d\boldsymbol{\theta} &= O\left(1\right)\left(2\pi\right)^{r} e^{-gn^{2\varepsilon}}\binom{n}{r}\pi^{-\frac{r}{2}}\left(cn\right)^{-\frac{r}{2}}\pi^{\frac{n}{2}}\det\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^{T}\right)^{-\frac{1}{2}}\\ &= O\left(1\right)\left(2\pi e^{-gn^{2\varepsilon}}n\pi^{-\frac{1}{2}}\left(cn\right)^{-\frac{1}{2}}\right)^{r}\pi^{\frac{n}{2}}\det\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^{T}\right)^{-\frac{1}{2}};\end{aligned}$$

whence we have, for n large enough,

$$\sum_{r=1}^{n^{1-\varepsilon}}\int_{\Theta_r}|F(\boldsymbol{\theta})|\,d\boldsymbol{\theta} = O\left(e^{-g'n^{2\varepsilon}}\right)\pi^{\frac{n}{2}}\det\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^{T}\right)^{-\frac{1}{2}}, \tag{5.48}$$

where $g' > 0$ is a constant, dependent only on δ. By adding (5.46) and (5.48), we get (5.42), and this completes the proof. □

5.6 Some Basic Corollaries

In this section we prove some basic Corollaries of Theorem 5.4.1.

5.6.1 Growth rate of $|\mathbb{T}(\boldsymbol{s})|$

Our first Corollary gives a rough indication of the growth rate of $|T(\boldsymbol{s})|$.

Corollary 5.6.1. *For any δ-tame score vector $\boldsymbol{s}$ there exists a constant C, such that*

$$C\left(\frac{2}{\pi}\right)^{\frac{n-1}{2}}\left(\frac{1}{1-\delta}\right)^{\binom{n}{2}}n^{-\frac{n}{2}} \leq |\mathbb{T}(\boldsymbol{s})| \leq C\left(\frac{2^{n-1}}{\pi\delta(1-\delta)}\right)^{\frac{n-1}{2}}n^{-\frac{n}{2}}. \tag{5.49}$$

Proof. By (5.36) and Corollary (5.36), if we set $\kappa = 1$, we have

$$2^{\frac{3n-3}{2}}n^{-\frac{n}{2}} \leq \det\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^{T}\right)^{-\frac{1}{2}} \leq \left(\frac{1}{2}\delta(1-\delta)\right)^{\frac{1-n}{2}}n^{-\frac{n}{2}}. \tag{5.50}$$

Moreover, since by Theorem 4.2.4 the β-model is the probability distribution that maximises $P(T)$, we have, in particular, that $P(T)$ is larger than the probability of a tournament with score vector $\boldsymbol{s}$ coming from a uniform random tournament. Thus

$$P(T) \geq \frac{1}{2^{\binom{n}{2}}}.$$

Moreover, by δ-tameness,

$$P(T) \leq (1-\delta)^{\binom{n}{2}}.$$

Therefore

$$\left(\frac{1}{1-\delta}\right)^{\binom{n}{2}} \leq \frac{1}{P(T)} \leq 2^{\binom{n}{2}}. \tag{5.51}$$

Applying Theorem 5.4.1, if we set C to be the $O(1)$ terms in (5.20), the result then follows from (5.50) and (5.51). □

5.6.2 Odd Regular Tournaments

Another easy application is to use the formula (5.20) to compute the number of regular tournaments on n vertices, when n is odd. By Corollary 5.6.1, we need only compute the $O(1)$ term in (5.20).

Note that if n is odd, the score vector for a regular tournament is the expected score vector of a uniform random tournament, which in turn corresponds to a β-model random tournament in which all the values β_i are equal. Since each $\lambda_{jk} = \frac{1}{2}$, we find

$$f(\boldsymbol{\theta}) = -\frac{1}{192}\sum_{j<k} (\theta_j - \theta_k)^4 .$$

In applying formula (5.20), we find that a good choice of κ is $\frac{1}{8}$ so that $A + \kappa \boldsymbol{w}\boldsymbol{w}^T$ is a diagonal matrix with each diagonal entry equal to $\frac{1}{8}n$. Thus, the random variable defined by 5.37 has covariance matrix given by

$$\left(2\left(A + \kappa \boldsymbol{w}\boldsymbol{w}^T\right)\right)^{-1} = 4n^{-1} I_n,$$

where I is the identity matrix. In this case, the fact that the random variables are uncorrelated (and hence independent) makes the expectation in formula (5.20) easy to compute. In particular, if the terms of the covariance matrix are σ_{ij}, we have

$$\mathbb{E}\left(X_j - X_k\right)^4 = 3\left(\sigma_{jj}^2 + \sigma_{kk}^2\right) + 6\sigma_{jj}\sigma_{kk} = 192 n^{-2},$$

so that

$$\mathbb{E}\, f(\boldsymbol{x}) = -\frac{n-1}{2n}.$$

Therefore, we have the following Corollary:

Corollary 5.6.2. *For odd $n \to \infty$, the number of regular tournaments is*

$$\left(1 + O\left(n^{-\frac{1}{2}+9\varepsilon}\right)\right)\left(\frac{2^{n+1}}{\pi n}\right)^{\frac{n-1}{2}} n^{\frac{1}{2}} e^{-\frac{1}{2}},$$

for any $\varepsilon > 0$.

This is equivalent to McKay's result (5.1).

5.6.3 Even Regular Tournaments

Another interesting application of formula (5.20) is to compute the number of regular tournaments on n vertices when n is even. This is a special case of (5.2). Recall that an almost regular tournament, for n even, is a tournament with the score vector

$$\left(\frac{n-2}{2}, \frac{n-2}{2}, \ldots, \frac{n-2}{2}, \frac{n}{2}, \ldots, \frac{n}{2}, \frac{n}{2}\right).$$

Let $U = \left[\frac{n}{2}\right]$ and $V = [n] - \left[\frac{n}{2}\right]$. Then, this score vector is the expected score vector of an edge model random tournament where the probabilities (denoted by λ_{jk}) that a vertex j dominates vertex k are given by

$$\lambda_{jk} = \begin{cases} \frac{1}{2} - \frac{1}{n}, & \text{if } j \in U, k \in V, \\ \frac{1}{2} + \frac{1}{n}, & \text{if } j \in V, k \in U, \\ \frac{1}{2}, & \text{otherwise,} \end{cases}$$

so that

$$\begin{aligned} f(\boldsymbol{\theta}) = \sum_{j \in U,\, k \in V} & -\frac{1}{3n}\left(\frac{1}{4} - \frac{1}{n^2}\right) i\left(\theta_j - \theta_k\right)^3 - \frac{1}{24}\left(\frac{1}{2} - \frac{6}{n^2}\right)\left(\frac{1}{4} - \frac{1}{n^2}\right)\left(\theta_j - \theta_k\right)^4 \\ & - \frac{1}{192}\left(\sum_{j<k \in U}\left(\theta_j - \theta_k\right)^4 + \sum_{j<k \in V}\left(\theta_j - \theta_k\right)^4\right). \end{aligned}$$

By choosing $\kappa = \frac{1}{2}\left(\frac{1}{4} - \frac{1}{n^2}\right)$, the result is that $A + \kappa \boldsymbol{w}\boldsymbol{w}^T$ is a block diagonal matrix:

$$A + \kappa \boldsymbol{w}\boldsymbol{w}^T = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}$$

with

$$C_{ij} = \begin{cases} \frac{n}{8} - \frac{1}{4n} - \frac{1}{2n^2}, & \text{if } i = j, \\ -\frac{1}{2n^2}, & \text{otherwise.} \end{cases}$$

Noting that $2C$ is a $\left(\frac{n}{4} - \frac{1}{2n}, -\frac{1}{n^2}\right)$-perturbed identity matrix, by Lemma 5.3.2, the matrix $2C$ has an inverse given by

$$(2C)^{-1}_{ij} = \begin{cases} \left(\frac{n}{4} - \frac{1}{2n}\right)^{-1} + \frac{16}{(n^2-6)(n^2-2)}, & \text{if } i = j, \\ \frac{16}{(n^2-6)(n^2-2)}, & \text{otherwise,} \end{cases}$$

so that the random variable defined by (5.37) has covariance matrix

$$\begin{bmatrix} (2C)^{-1} & 0 \\ 0 & (2C)^{-1} \end{bmatrix}.$$

We see from (5.6.3) and Lemma 5.5.3, that

$$\mathbb{E} f(\boldsymbol{x}) = \mathbb{E}\left(-\frac{1}{192}\sum_{j<k}\left(X_j - X_k\right)^4\right) + O\left(n^{-2}\right)$$

and

$$\mathbb{E}\left(\Im f(\boldsymbol{x})\right)^2 = O\left(n^{-1}\right).$$

Moreover, since the main contributing factors from the covariance matrix in computing this expectation are the $O\left(n^{-1}\right)$ terms, we find that

$$\mathbb{E}\left(f\left(\boldsymbol{x}\right)-\frac{1}{2}\left(\Im f(\boldsymbol{x})\right)^2\right)=-\frac{n-1}{2n}.$$

Moreover, since $A+\kappa\boldsymbol{w}\boldsymbol{w}^T$ is a block matrix, its eigenvalues are two copies of those of C (including multiplicity). By Corollary 5.3.4,

$$\det\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)=\left(\frac{n}{8}-\frac{1}{4n}\right)^{n-2}\left(\frac{n}{8}+\frac{1}{4n}\right)^2,$$

so that

$$\det\left(A+\kappa\boldsymbol{w}\boldsymbol{w}^T\right)^{-\frac{1}{2}}=\left(1+O\left(n^{-2}\right)\right)2^{\frac{3n}{2}}n^{-\frac{n}{2}}.$$

Finally, for the term $P(T)$, we need only consider the probability of a particular tournament with score vector $\boldsymbol{s}$. By counting edges in such a tournament, we get

$$\begin{aligned}
P(T)&=2^{-\frac{n}{4}\left(\frac{n}{2}-1\right)}\left(\frac{1}{2}-\frac{1}{n}\right)^{\frac{n}{4}\left(\frac{n}{2}-1\right)}2^{-\frac{n}{4}\left(\frac{n}{2}-1\right)}\left(\frac{1}{2}+\frac{1}{n}\right)^{\frac{n}{4}\left(\frac{n}{2}+1\right)}\\
&=2^{-\binom{n}{2}}\left(1-\frac{4}{n^2}\right)^{\frac{n^2}{8}}\left(1+\frac{2}{n}\right)^{\frac{n}{2}}\left(1+O\left(n^{-1}\right)\right)\\
&=2^{-\binom{n}{2}}e^{\frac{1}{2}}\left(1+O\left(n^{-1}\right)\right).
\end{aligned}$$

We then get

$$\frac{1}{P(T)}=2^{\binom{n}{2}}e^{-\frac{1}{2}}\left(1-O\left(n^{-1}\right)\right).$$

This gives us our next corollary:

Corollary 5.6.3. *For even $n\to\infty$, the number of regular tournaments is*

$$\left(1+O\left(n^{-\frac{1}{2}+9\varepsilon}\right)\right)\left(\frac{2^{\frac{n+1}{2}}}{\pi n}\right)^{\frac{n-1}{2}}n^{\frac{1}{2}}e^{-1},$$

for any $\varepsilon>0$.

Remark 5.6.4. Note that a similar approach to that outlined in this section may be applied to find the number of tournaments with exactly 2 scores s_1 and s_2, at least up to a constant term. The difficulty lies in approximating the covariance and expectation in (5.20).

5.7 Future Direction

5.7.1 Finding the Number of Tournaments with a Given Score Vector

Suppose one were to observe a tournament with n players with δ-tame score vector $\boldsymbol{s}$. How many other ways could the tournament result in score vector $\boldsymbol{s}$?

If we can find explicit bounds for the constant terms in formula (5.20), this formula provides an approach for approximating this question. Such a method would proceed as follows:

1. Use an optimisation algorithm to find the $\tilde{\lambda}_{jk}$ (by maximising (4.5)).
2. Compute $A + \kappa \boldsymbol{w}\boldsymbol{w}^T$, for a suitable choice of κ, and use this to find the determinant, the expectation and covariance terms in (5.20)
3. Find a tournament with score vector $\boldsymbol{s}$, and use this to compute the term $\tilde{P}(T)$ using the $\tilde{\lambda}_{jk}$.

Alternatively, the bounds in Corollary 5.6.1 provide a rough approximation that can be computed much more quickly. This simply requires computing the value of δ, and applying equation (5.49).

This result would allow one to compute the probability that a tournament has score vector $\boldsymbol{s}$, with respect to a β-model random tournament.

5.7.2 Finding the Number of Tournaments with a Specified Oriented Graph

A similar approach outlined in this section would allow the enumeration of tournaments with a δ-tame score vector $\boldsymbol{s}$ containing a fixed oriented graph H. The main difference would be in changing $\tilde{\lambda}_{jk}$ so that

$$\tilde{\lambda}_{jk} = \begin{cases} 1, & (j,k) \in H, \\ 0, & (k,j) \in H, \\ \frac{e^{\tilde{\beta}_j}}{e^{\tilde{\beta}_j} + e^{\tilde{\beta}_k}}, & \text{otherwise.} \end{cases}$$

This change increases the complexity of many parts of this section, in particular, there may be difficulty in bounding the negligible contribution of the integral. If such an enumeration result can be found, then this would allow one to find the probability that a tournament with score vector $\boldsymbol{s}$ contains an oriented graph H, with respect to a β-model random tournament.

Appendix A

Results Needed for Chapter 5

For completeness, we reproduce the results from [IM16] required for Chapter 5 below.

Theorem A.0.1 (Theorem 4.2 [IM16], due to Isserlis [Iss18])**.** *Let A be a positive-definite real symmetric matrix of order n and let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random variable with the normal density $\pi^{-n/2}|A|^{-1/2}e^{-\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}}$. Let $\Sigma = (\sigma_{jk}) = (2A)^{-1}$ be the corresponding covariance matrix. Consider a product $Z = X_{j_1}X_{j_2}\cdots X_{j_k}$, where the subscripts do not need to be distinct. If k is odd, then $\mathbb{E}\, Z = 0$. If k is even, then*

$$\mathbb{E}\, Z = \sum_{(i_1,i_2),(i_2,i_3),\ldots,(i_{k-1},i_k)} \sigma_{j_{i_1}j_{i_2}} \cdots \sigma_{j_{i_{k-1}}j_{i_k}},$$

where the sum is over all unordered partitions of $\{1, 2, \ldots, k\}$ into $k/2$ disjoint unordered pairs. The number of terms in the sum is $(k-1)(k-3)\cdots 3\cdot 1$.

□

Theorem A.0.2 (Theorem 4.4 [IM16])**.** *Let $c_1, c_2, c_3, \varepsilon, \rho_1, \rho_2, \phi_1, \phi_2$ be nonnegative real constants with $c_1, \varepsilon > 0$. Let A be an $n \times n$ positive-definite symmetric real matrix and let T be a real matrix such that $T^{\mathrm{T}}AT = I$. Let Ω be a measurable set such that $U_n(\rho_1) \subseteq T^{-1}(\Omega) \subseteq U_n(\rho_2)$, and let $f : \mathbb{R}^n \to \mathbb{C}$, $g : \mathbb{R}^n \to \mathbb{R}$ and $h : \Omega \to \mathbb{C}$ be measurable functions. We make the following assumptions.*

(a) $c_1(\log n)^{1/2+\varepsilon} \le \rho_1 \le \rho_2$.

(b) For $\boldsymbol{x} \in T(U_n(\rho_1))$, $2\rho_1 \|T\|_1 |f_j(\boldsymbol{x})| \le \phi_1 n^{-1/3} \le \frac{2}{3}$ *for* $1 \le j \le n$ *and* $4\rho_1^2 \|T\|_1 \|T\|_\infty \|H(f, T(U_n(\rho_1)))\|_\infty \le \phi_1 n^{-1/3}$.

(c) For $\boldsymbol{x} \in \Omega$, $\Re f(\boldsymbol{x}) \le g(\boldsymbol{x})$. *For* $\boldsymbol{x} \in T(U_n(\rho_2))$, *either*
(i) $2\rho_2 \|T\|_1 |g_j(\boldsymbol{x})| \le (2\phi_2)^{3/2} n^{-1/2}$ *for* $1 \le j \le n$, *or*
(ii) $2\rho_2 \|T\|_1 |g_j(\boldsymbol{x})| \le \phi_2 n^{-1/3}$ *for* $1 \le j \le n$ *and* $4\rho_2^2 \|T\|_1 \|T\|_\infty \|H(g, T(U_n(\rho_2)))\|_\infty \le \phi_2 n^{-1/3}$.

(d) $|f(\boldsymbol{x})|, |g(\boldsymbol{x})| \le n^{c_3} e^{c_2 \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}/n}$ *for* $\boldsymbol{x} \in \mathbb{R}^n$.

Let $\boldsymbol{X}$ be a random variable with the normal density $\pi^{-n/2}\det(A)^{1/2}e^{-\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}}$. Then, provided $\mathbb{V}f(\boldsymbol{X})$ and $\mathbb{V}g(\boldsymbol{X})$ are finite and h is bounded in Ω,

$$\int_{\Omega} e^{-\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}+f(\boldsymbol{x})+h(\boldsymbol{x})}\,d\boldsymbol{x} = (1+K)\pi^{n/2}\det(A)^{-1/2}e^{\mathbb{E}\,f(\boldsymbol{X})+\frac{1}{2}\mathbb{V}f(\boldsymbol{X})},$$

where, for some constant C depending only on $c_1, c_2, c_3, \varepsilon$,

$$|K| \le Ce^{\frac{1}{2}\operatorname{Var}\Im f(\boldsymbol{X})}\Big(e^{\phi_1^3+e^{-\rho_1^2/2}}-1 + \big(2e^{\phi_2^3+e^{-\rho_1^2/2}}-2+\sup_{\boldsymbol{x}\in\Omega}|e^{h(\boldsymbol{x})}-1|\big)\,e^{\mathbb{E}(g(\boldsymbol{X})-\Re f(\boldsymbol{X}))+\frac{1}{2}(\operatorname{Var}g(\boldsymbol{X})-\operatorname{Var}\Re f(\boldsymbol{X}))}\Big).$$

In particular, if $n \ge (1+2c_2)^2$ and $\rho_1^2 \ge 15+4c_2+(3+8c_3)\log n$, we can take $C=1$.

□

Remark A.0.3 (Remark 4.5 [IM16])**.** Note that $U_n(\rho_1) \subseteq T^{-1}(\Omega) \subseteq U_n(\rho_2)$ is implied by $U_n(\rho_1\|T\|_\infty) \subseteq \Omega \subseteq U_n(\rho_2\|T^{-1}\|_\infty^{-1})$, so the latter condition could be used instead of the former.

Lemma A.0.4 (Lemma 4.6 [IM16])**.** *Let $Q, W : \mathbb{R}^n \to \mathbb{R}^n$ be linear operators such that* $\ker Q \cap \ker W = \{\mathbf{0}\}$ *and* $\operatorname{span}(\ker Q, \ker W) = \mathbb{R}^n$. *Let $n^{\perp}$ denote the dimension of* $\ker Q$. *Suppose $\Omega \subseteq \mathbb{R}^n$ and $F : \Omega \cap Q(\mathbb{R}^n) \to \mathbb{C}$. For any $\rho > 0$, define*

$$\Omega_\rho = \{\boldsymbol{x} \in \mathbb{R}^n \mid Q\boldsymbol{x} \in \Omega \text{ and } W\boldsymbol{x} \in U_n(\rho)\}.$$

Then, if the integrals exist,

$$\int_{\Omega\cap Q(\mathbb{R}^n)} F(\boldsymbol{y})\,d\boldsymbol{y} = (1-K)^{-1}\,\pi^{-n_\perp/2}\,\det\big(Q^{\mathrm{T}}Q+W^{\mathrm{T}}W\big)^{1/2}\int_{\Omega_\rho} F(Q\boldsymbol{x})\,e^{-\boldsymbol{x}^{\mathrm{T}}W^{\mathrm{T}}W\boldsymbol{x}}\,d\boldsymbol{x},$$

where

$$0 \le K < \min(1, ne^{-\rho^2/\kappa^2}), \quad \kappa = \sup_{W\boldsymbol{x}\ne 0}\frac{\|W\boldsymbol{x}\|_\infty}{\|W\boldsymbol{x}\|_2} \le 1.$$

Moreover, if $U_n(\rho_1) \subseteq \Omega \subseteq U_n(\rho_2)$ for some $\rho_2 \ge \rho_1 > 0$ then

$$U_n\bigg(\min\Big(\frac{\rho_1}{\|Q\|_\infty}, \frac{\rho}{\|W\|_\infty}\Big)\bigg) \subseteq \Omega_\rho \subseteq U_n\Big(\|P\|_\infty\,\rho_2 + \|R\|_\infty\,\rho\Big)$$

for any linear operators $P, R : \mathbb{R}^n \to \mathbb{R}^n$ such that $PQ+RW$ is equal to the identity operator on $\mathbb{R}^n$.

□

Lemma A.0.5 (Lemma 4.9 [IM16])**.** *Let D be an $n\times n$ real diagonal matrix with* $d_{\min} = \min_j d_{jj} > 0$ *and* $d_{\max} = \max_j d_{jj}$. *Let A be a real symmetric positive-semidefinite $n\times n$ matrix with*

$$\|A-D\|_{\max} \le \frac{rd_{\min}}{n} \qquad \text{and} \qquad \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} \ge \gamma\boldsymbol{x}_\parallel^{\mathrm{T}}D\boldsymbol{x}_\parallel = \gamma\boldsymbol{x}^{\mathrm{T}}D^{1/2}(I-P_D)D^{1/2}\boldsymbol{x}$$

for some $1 \ge \gamma > 0$, $r > 0$ and all $\boldsymbol{x} \in \mathbb{R}^n$. Let $n_\perp$ denote the dimension of $\ker A$. *Then the following are true.*

(a) $\|A_D - A\|_\infty \leq rn_\perp d_{\max}^{1/2} d_{\min}^{1/2}$, $\|A_D - A\|_{\max} \leq \dfrac{r^2 n_\perp d_{\min}}{n}$ *and* $n_\perp \leq r^2$.

(b) A_D *is symmetric and positive-definite. Moreover,*

$$\|A_D^{-1}\|_\infty \leq \frac{r+\gamma}{\gamma d_{\min}} \quad \text{and} \quad \|A_D^{-1} - D^{-1}\|_{\max} \leq \frac{(r+\gamma)r}{\gamma n d_{\min}}(1 + rn_\perp).$$

(d) *There exists a matrix* T *such that* $T^{\mathrm{T}} A_D T = I$ *and*

$$\|T\|_1, \|T\|_\infty \leq \frac{r + \gamma^{1/2}}{\gamma^{1/2} d_{\min}^{1/2}} \quad \text{and} \quad \|T^{-1}\|_1, \|T^{-1}\|_\infty \leq \left(\frac{(r+1)(r+\gamma^{1/2})}{\gamma^{1/2}} + rn_\perp\right) d_{\max}^{1/2}.$$

□

Bibliography

[ASE92] Noga Alon, Joel H Spencer, and Paul Erdős, *The probabilistic method*, Wiley, 1992.

[Bar12] J.A. Barvinok, Alexander Hartigan, *The number of graphs and a random graph with a given degree sequence*, Random Structures & Algorithms **42** (2012), no. 3, 301–348.

[BGQ97] A. D. Barbour, Anant P. Godbole, and Jinghua Qian, *Imperfections in random tournaments*, Comb. Probab. Comput. **6** (1997), no. 1, 1–15.

[BjS77] Chang M. Bang and Henry jun. Sharp, *An elementary proof of Moon's theorem on generalized tournaments.*, J. Comb. Theory, Ser. B **22** (1977), 299–301 (English).

[Bol01] Béla Bollobás, *Random graphs*, Cambridge University Press, 2001.

[Bow72] V. J. Bowman, *Permutation polyhedra*, SIAM Journal on Applied Mathematics **22** (1972), no. 4, 580–589.

[BS15] Béla Bollobás and Alex Scott, *Intersections of random hypergraphs and tournaments*, European Journal of Combinatorics **44 Part A** (2015), 125 – 139.

[Cha81] Gary Chartrand, *The theory and applications of graphs*, Wiley, 1981.

[CHK$^+$12] Villő Csiszár, Péter Hussami, János Komlós, Tamás F. Móri, Lídia Rejtő, and Gábor Tusnády, *When the degree sequence is a sufficient statistic*, Acta Mathematica Hungarica **134** (2012), no. 1, 45–53.

[Chu91] R. L. Chung, F. R. K.Graham, *Quasi-random tournaments*, J. Graph Theory **15** (1991), no. 2, 173–198.

[EL09] S. Erick Alm and S. Linusson, *A counter-intuitive correlation in a random tournament*, ArXiv e-prints (2009).

[Fel68] William Feller, *An introduction to probability theory and its applications*, J. Wiley, 1968.

[FH66] Leo Moser Frank Harary, *The theory of round robin tournaments*, The American Mathematical Monthly **73** (1966), no. 3, 231–246.

[For57] L. R. Ford, *Solution of a ranking problem from binary comparisons*, The American Mathematical Monthly **64** (1957), no. 8, 28–33.

[Fra68] O. Frank, *Stochastic competition graphs*, Revue de l'Institut International de Statistique / Review of the International Statistical Institute **36** (1968), no. 3, 319–326.

[GMW00] Zhicheng Gao, Brendan D. McKay, and Xiaoji Wang, *Asymptotic enumeration of tournaments with a given score sequence containing a specified digraph.*, Random Struct. Algorithms **16** (2000), no. 1, 47–57 (English).

[HU01] Claude Hiriart-Urruty, Jean-BaptisteLemarechal, *Fundamentals of convex analysis*, Springer, 2001.

[Hub63] Peter J. Huber, *A remark on a paper of trawinski and david entitled: "selection of the best treatment in a paired-comparison experiment"*, The Annals of Mathematical Statistics **34** (1963), no. 1, 92–94.

[Hun04] David R. Hunter, *Mm algorithms for generalized bradley-terry models*, Ann. Statist. **32** (2004), no. 1, 384–406.

[IM16] M. Isaev and B. D. McKay, *Complex martingales and asymptotic enumeration*, ArXiv e-prints (2016).

[Iss18] L. Isserlis, *On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables*, Biometrika **12** (1918), no. 1/2, 134–139.

[Jan95] Svante Janson, *Hamilton cycles in a random tournament*, Random Structures & Algorithms **6** (1995), no. 2-3, 213–220.

[Jec83] Thomas Jech, *The ranking of incomplete tournaments: A mathematician's guide to popular sports*, The American Mathematical Monthly **90** (1983), no. 4, 246–266.

[Ken55] M. G. Kendall, *Further contributions to the theory of paired comparisons*, Biometrics **11** (1955), no. 1, 43–62.

[Kni96] Oliver Knill, *A short introduction to several complex variables*, `http://www.math.harvard.edu/~knill/teaching/severalcomplex_1996/severalcomplex.pdf`, 1996.

[KO13] Daniela Kühn and Deryk Osthus, *Hamilton decompositions of regular expanders: A proof of kellys conjecture for large tournaments*, Advances in Mathematics **237** (2013), 62 – 146.

[Lan53] H. G. Landau, *On dominance relations and the structure of animal societies: III the condition for a score structure*, The bulletin of mathematical biophysics **15** (1953), no. 2, 143–148.

[Mau80] Stephen B. Maurer, *The king chicken theorems*, Mathematics Magazine **53** (1980), no. 2, 67.

[McK90] Brendan D. McKay, *The asymptotic numbers of regular tournaments, Eulerian digraphs and Eulerian oriented graphs.*, Combinatorica **10** (1990), no. 4, 367–377 (English).

[Moo62] L. Moon, J. W.Moser, *Almost all tournaments are irreducible*, Canadian Mathematical Bulletin **5** (1962), no. 0, 61–65.

[Moo63] J. W. Moon, *An extension of Landau's theorem on tournaments*, Pacific J. Math. **13** (1963), 1343–1345. MR 0155763

[Moo68] John W Moon, *Topics on tournaments in graph theory*, Holt, Rinehart and Winston, 1968.

[Mor47] P. A. P. Moran, *On the method of paired comparisons*, Biometrika **34** (1947), no. 3/4, 363–365.

[R3́4] L. Rédei, *Ein kombinatorischer satz*, Acta Litt. Szeged. **7** (1934), 39–43.

[RAB52] Milton E. Terry Ralph Allan Bradley, *Rank analysis of incomplete block designs: I. the method of paired comparisons*, Biometrika **39** (1952), no. 3/4, 324–345.

[Sil92] Charles L. Silver, *The voting paradox*, The Mathematical Gazette **76** (1992), no. 477, 387–388.

[SJ80] John Shore and Rodney Johnson, *Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy*, IEEE Transactions on information theory **26** (1980), no. 1, 26–37.

[Spe71] J. Spencer, *Optimal ranking of tournaments*, Networks **1** (1971), no. 2, 135–138.

[Spe74] J. H. Spencer, *Random regular tournaments*, Periodica Mathematica Hungarica **5** (1974), no. 2, 105–120.

[Spe80] J. Spencer, *Optimally ranking unrankable tournaments*, Periodica Mathematica Hungarica **11** (1980), no. 2, 131–144.

[SY99] Gordon Simons and Yi-Ching Yao, *Asymptotics when the number of parameters tends to infinity in the bradley-terry model for paired comparisons*, Ann. Statist. **27** (1999), no. 3, 1041–1060.

[Sze43] T. Szele, *Kombinatorikai vizsgálatok az irányitott teljes gráffal kapcsolatban*, Mat. Fiz. Lapok (1943), no. 50, 223–256.

[Wei52] T. H. Wei, *The algebraic foundations of ranking theory*, Ph.D. thesis, Cambridge University, 1952.

[Wornt] N Wormald, *Tournaments with many hamilton cycles*, http://users.monash.edu.au/~nwormald/papers/hamtourn.pdf, preprint.

[Zer29] E. Zermelo, *Die berechnung der turnier-ergebnisse als ein maximumproblem der wahrscheinlichkeitsrechnung.*, Mathematische Zeitschrift **29** (1929), 436–460.

[Zie95] Gunter M Ziegler, *Lectures on polytopes*, Springer-Verlag, 1995.

[RG96] Tomasz uczak, Andrzej Ruciski, and Jacek Gruszka, *On the evolution of a random tournament*, Discrete Mathematics **148** (1996), no. 1, 311 – 316.