Voting paradoxes and digraphs realizations

Noga Alon *

Abstract

A family of permutations $\mathcal{F}$ forms a realization of a directed graph $T = (V, E)$ if for every directed edge uv of T, u precedes v in more than half of the permutations. The quality $q(\mathcal{F}, T)$ of the realization is the minimum, over all directed edges uv of T, of the ratio $(|\mathcal{F}(u, v)| - |\mathcal{F}(v, u)|)/|\mathcal{F}|$, where $|\mathcal{F}(x, y)|$ is the number of permutations in $\mathcal{F}$ in which x precedes y. The study of this quantity is motivated by questions about voting schemes in which each individual has a linear ordering of all candidates, and the individual preferences are combined to decide between any pair of possible candidates by applying the majority vote. It is shown that every simple digraph T on n vertices, with no anti-parallel edges, admits a realization $\mathcal{F}$ with quality at least $c/\sqrt{n}$ for some absolute positive constant c, and this is tight up to the constant factor c.

1 Introduction

All directed graphs considered here are finite, simple (that is, have no loops and no parallel edges), and have no anti-parallel edges. The densest digraphs of this type are tournaments. A *tournament* on a set V of n vertices is a directed graph on V in which for every pair of distinct vertices $u, v \in V$, either uv or vu is a directed edge, but not both. Let $T = (V, E)$ be a digraph, and let $\mathcal{F}$ be a collection of (not necessarily distinct) permutations of V. We say that $\mathcal{F}$ is a *realization* of T if for every directed edge $uv \in E$, u precedes v in a majority of the permutations in $\mathcal{F}$. The *quality* $q(\mathcal{F}, T)$ of the realization is given by

$$q(\mathcal{F}, T) = min_{uv \in E} \frac{|\mathcal{F}(u, v)| - |\mathcal{F}(v, u)|}{|\mathcal{F}|},$$

where $\mathcal{F}(x, y)$ is the set of all permutations in $\mathcal{F}$ in which x precedes y.

McGarvey [9] proved that every tournament (and hence every digraph) has a realization by permutations, and subsequent results by Stearns [11] and Erdős and Moser [4] imply that every

*Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. Email: noga@math.tau.ac.il. Research supported in part by a USA-Israeli BSF grant, by the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.

tournament on n vertices can be realized by $O(n/\log n)$ permutations, and some tournaments on n vertices cannot be realized by less than $\Omega(n/\log n)$ permutations. In this paper we study the maximum possible real $q = q(n)$, such that every tournament (and hence every digraph) on n vertices has a realization with quality at least q. It turns out that $q(n) = \Theta(1/\sqrt{n})$, as stated in the following theorem.

Theorem 1.1 *There are three absolute positive constants c_1, c_2, c_3 such that the following holds for every integer n.*
(i) For every digraph $T = (V, E)$ on n vertices there is a set $\mathcal{F}$ of permutations of V such that $q(\mathcal{F}, T) \geq \frac{c_1}{\sqrt{n}}$. Moreover, there is such an $\mathcal{F}$ of cardinality $|\mathcal{F}| \leq c_3 n \log n$.
(ii) There exists a tournament $T = (V, E)$ on n vertices, such that for every realization $\mathcal{F}$ of T, $q(\mathcal{F}, T) \leq \frac{c_2}{\sqrt{n}}$. In fact, this holds for almost all tournaments on n vertices.

The question of realizing digraphs by permutations arises in Social Choice Theory (see, e.g., [5]). Thus, for example, the well known Condorcet Paradox which asserts that the majority might prefer option A over option B, prefer option B over option C, and yet prefer option C over option A, even if each individual has a linear order over the options, is simply the fact that the cyclic triangle can be realized by permutations. Realizations of this type occur in the study of voting schemes in which each individual has a linear ordering of all candidates, and the individual preferences are combined to decide between any pair of possible candidates by applying the majority vote. The quality of a realization is thus a measure for the smallest gap between a pair of candidates, with a given set of voters. Recent results of Kalai [7] about schemes in which other rules are applied instead of majority, provide further motivation to study the quantity $q(n)$ defined above, and indeed the problem of estimating this quantity was raised by Kalai.

The proof of Theorem 1.1, part (i) combines some probabilistic arguments with the minmax Theorem, and is based on an extremal result about transitive subgraphs in weighted directed graphs, which may be of independent interest. We present this proof in Section 2, together with an extension of it dealing with digraphs of small maximum degree. The assertion of part (ii) can be derived from a known result of de la Vega [3], as described in Section 3. The final Section 4 contains some concluding remarks.

2 Transitive subgraphs in weighted digraphs

Let $D = (V, E)$ be a loopless directed graph on the set of vertices $V = \{1, 2, \ldots, n\}$ in which every pair of vertices are joined by at most one oriented edge, and let $w : E \mapsto R^+$ be a weight function assigning to each directed edge a positive real weight. Let $w(E) = \sum_{ab \in E} w(ab)$ denote the total weight of the edges of D. For a permutation σ of V let $FIT(D, \sigma)$ denote the total weight of all edges ij of D for which i precedes j in σ, that is

$$FIT(D,\sigma) = \sum_{ij\in E,\sigma(i)<\sigma(j)} w(ij).$$

Our first result in this section is the following.

Theorem 2.1 *There is an absolute, positive constant c such that the following holds. For every weighted, simple, directed graph on n vertices $D = (V, E)$ with no anti-parallel edges, there is a permutation σ of V such that*

$$FIT(D,\sigma) \geq (\frac{1}{2} + \frac{c}{\sqrt{n}})w(E).$$

This extends a result of Spencer [10], who proved the above result for non-weighted tournaments. The result is tight, up to the constant c, as follows from the discussion in the next section.

In the proof of the theorem, we apply the following result of Szarek. See also [8] for a shorter proof of a more general result.

Lemma 2.2 ([12]) *For every set of m reals $c_1, c_2, \ldots, c_m$, the expected value of the random variable $|\sum_{i=1}^{m} \epsilon_i c_i|$, where the variables ϵ_i are independent, identically distributed random variables, each distributed uniformly on $\{-1, 1\}$, is at least $2^{-1/2}(c_1^2 + \ldots + c_m^2)^{1/2}$.*

Proof of Theorem 2.1. We make no attempt to optimize the multiplicative constant c and prove the theorem with $c = 1/16$. It is convenient to extend the definition of the function w to $V \times V$ by putting $w(uv) = 0$ for every ordered pair of vertices u, v for which $uv \notin E$. For any two disjoint sets of vertices $A, B \subset V$, define $w(A, B) = \sum_{a\in A,b\in B} w(ab)$. Let $V = A \cup B$ be a random partition of V into two disjoint sets, obtained by choosing each member of V, randomly and independently, to lie in A or in B with equal probability. By linearity of expectation, the expected value of $w(A, B) + w(B, A)$ is precisely $w(E)/2$, and hence there are A and B for which

$$w(A, B) + w(B, A) \geq w(E)/2. \tag{1}$$

Fix such a partition $V = A \cup B$ and assume, without loss of generality, that $|B| \leq n/2$. With these fixed A and B, let $B = X \cup Y$ be a random partition of B obtained by choosing each member of B, randomly and independently, to lie in X or in Y with equal probability. For each vertex $a \in A$ let S_a be the random variable

$$S_a = |\sum_{x\in X} w(xa) - \sum_{x\in X} w(ax) + \sum_{y\in Y} w(ay) - \sum_{y\in Y} w(ya)|$$

By Lemma 2.2 (with $m = |B|$, the reals c_b, $b \in B$ being given by $c_b = w(ba)$ if $ba \in E$, $c_b = -w(ab)$ if $ab \in E$, and $c_b = 0$ otherwise, $\epsilon_b = +1$ if $b \in X$ and $\epsilon_b = -1$ if $b \in Y$) we conclude that the expectation of S_a satisfies

$$E(S_a) \geq \frac{1}{\sqrt{2}}[\sum_{b\in B} w^2(ab) + w^2(ba)]^{1/2}.$$

By the Cauchy Schwarz Inequality, and using the fact that $|B| \le n/2$ it follows that

$$E(S_a) \geq \frac{1}{\sqrt{2}} \frac{\sum_{b \in B} w(ab) + w(ba)}{|B|^{1/2}} \geq \frac{\sum_{b \in B} w(ab) + w(ba)}{\sqrt{n}}. \qquad (2)$$

Summing over all $a \in A$, and using (1) we obtain, by linearity of expectation

$$E(\sum_{a \in A} S_a) \geq \frac{w(E)}{2\sqrt{n}}.$$

Therefore, there is a fixed choice of X and Y such that

$$\sum_{a \in A} |\sum_{x \in X} w(xa) - \sum_{x \in X} w(ax)| + \sum_{a \in A} |\sum_{y \in Y} w(ay) - \sum_{y \in Y} w(ya)| \geq \frac{w(E)}{2\sqrt{n}},$$

where here we used the triangle inequality. It follows that either the first summand or the second one is at least $\frac{w(E)}{4\sqrt{n}}$, and in that summand, either the contribution of the positive terms or that of the absolute values of the negative terms is at least $\frac{w(E)}{8\sqrt{n}}$. In any case, we get two disjoint sets of vertices, say C and Z, one of which is either X or Y and the other is a subset of A, such that

$$w(Z, C) - w(C, Z) = \sum_{c \in C} \sum_{z \in Z} (w(zc) - w(cz)) \geq \frac{w(E)}{8\sqrt{n}}.$$

Put $F = V \setminus (C \cup Z)$. Let σ_C be a random permutation of the elements of C (chosen uniformly among all possible permutations). Similarly, let σ_Z be a random permutation of the elements of Z and let σ_F be a random permutation of the elements of F. Finally, let σ be the (random) permutation of V obtained from the permutation σ_Z followed by the permutation σ_C by putting the permutation σ_F either before or after all elements of $C \cup Z$, where each of the two choices is equally likely. It is easy to see that the contribution of each directed edge which is not connecting a vertex in C with one in Z to the expected value of $FIT(D, \sigma)$ is precisely half its weight, whereas the total contribution of the edges between C and Z to this expected value exceeds half their total weight by $(W(Z, C) - W(C, Z))/2 \geq \frac{w(E)}{16\sqrt{n}}$. Therefore, the expected value of $FIT(D, \sigma)$ is at least $(\frac{1}{2} + \frac{1}{16\sqrt{n}})w(E)$, implying that there is a permutation σ with

$$FIT(D, \sigma) \geq (\frac{1}{2} + \frac{1}{16\sqrt{n}})w(E).$$

This completes the proof. ∎

Remark: The part of the proof following the construction of C and Z can be described without the random choices, by a simple greedy procedure. It is also possible to choose σ_C, σ_Z and σ_F more carefully (by applying the construction in the proof recursively) in order to get a somewhat better value of c in the statement of the theorem.

Proof of Theorem 1.1, part (i). Let $T = (V, E)$ be a digraph, where $V = \{1, 2, \ldots, n\}$. Consider the following two-person zero-sum game. The first player, called the *edge player*, chooses a directed edge $ij \in E$, and the second player, called the *order player* chooses a permutation σ of V. The edge player then pays the order player 1 if and only if $\sigma(i) < \sigma(j)$, that is, iff i precedes j in the order σ. A mixed strategy for the edge player is a probability distribution on the edges of T. By Theorem 2.1, for every such mixed strategy, there is a pure strategy of the order player that ensures him an expected payoff of at least Val, where $Val = \frac{1}{2} + \frac{c}{\sqrt{n}}$ and $c \geq 1/16$ is the constant in the assertion of the theorem. It follows that the value of the game is at least Val and hence, by the minmax theorem, there is a mixed strategy of the order player whose expected payoff for every pure strategy of the edge player is at least Val. A mixed strategy of the order player is a probability distribution P on the permutations of V, and the fact that its expected payoff is at least Val means that for every directed edge ij of the tournament, the probability that $\sigma(i) < \sigma(j)$ when σ is chosen according to the distribution P, is at least Val. Put $t = 4n \log n / c^2$, and let $\mathcal{F}$ be a random collection of t permutations of V, where each member of $\mathcal{F}$ is chosen, randomly and independently, according to the distribution P. Fix a directed edge ij of T, and let A_{ij} be the event that i precedes j in at least $(\frac{1}{2} + \frac{c}{2\sqrt{n}})t$ permutations in $\mathcal{F}$. The expected number of permutations in $\mathcal{F}$ in which i precedes j is at least $(\frac{1}{2} + \frac{c}{\sqrt{n}})t$. Therefore, by the standard estimates of Chernoff (cf., e.g., [2], Theorem A.1.4), the probability that the event A_{ij} does not hold is at most

$$e^{-2c^2 t/(4n)} \leq 1/n^2.$$

It follows that with positive probability, all the events A_{ij} hold, and hence there is a collection $\mathcal{F}$ of $4n \log n / c^2$ ($\leq 1024 n \log n$) permutations of V, such that for every directed edge ij of T,

$$\frac{|\mathcal{F}(i,j)| - |\mathcal{F}(j,i)|}{|\mathcal{F}|} \geq \frac{c}{\sqrt{n}} \ (\ \geq \frac{1}{16\sqrt{n}}).$$

Thus, $q(\mathcal{F}, T) \geq \frac{c}{\sqrt{n}}$, as needed. ■

A close look at the proof of Theorem 2.1 shows that for digraphs with small maximum degree it gives a stronger statement. The *indegree* $d^-(v)$ of a vertex v in a digraph $D = (V, E)$ is the number of vertices u such that $uv \in E$. Similarly, the *outdegree* $d^+(v)$ of v is the number of vertices w such that $vw \in E$, and the *degree* $d(v)$ is the sum $d(v) = d^-(v) + d^+(v)$.

Theorem 2.3 *There is an absolute, positive constant c such that the following holds. For every weighted, simple, directed graph $D = (V, E)$ with no anti-parallel edges and with maximum degree at most d, there is a permutation σ of V such that*

$$FIT(D, \sigma) \geq (\frac{1}{2} + \frac{c}{\sqrt{d}}) w(E). \quad \Box$$

Indeed, this follows by repeating the proof of Theorem 2.1, and by observing that the $|B|^{1/2}$ term in (2) can be replaced by $d^{1/2}$. This implies the following strengthening of Theorem 1.1, part (i).

Theorem 2.4 *There is an absolute positive constant c_1 such that the following holds for every integer d. For every digraph $T = (V, E)$ with maximum degree d there is a set $\mathcal{F}$ of permutations of V such that $q(\mathcal{F}, T) \geq \frac{c_1}{\sqrt{d}}$.* $\square$

3 Random tournaments do not admit a high quality realization

In this section we present the (simple) proof of Theorem 1.1, part (ii). We need the following result of de-la Vega.

Theorem 3.1 ([3]) *There exists an absolute constant b such that the following holds. Let $T = (V, E)$ be a random tournament on the set $V = \{1, 2, \ldots, n\}$, obtained by choosing, for each $1 \leq i < j \leq n$, randomly and independently, either ij or ji to be a directed edge of T with equal probability. Assign to each edge of T weight 1. Then almost surely (that is, with probability that tends to 1 as n tends to infinity), for every permutation σ of V,*

$$FIT(T, \sigma) \leq \frac{1}{2}\binom{n}{2} + bn^{3/2}. \tag{3}$$

Proof of Theorem 1.1, part (ii). Let $T = (V, E)$ be a tournament on n vertices with weight 1 assigned to each of its edges, and suppose that (3) holds for each permutation σ of V. Let $\mathcal{F}$ be a realization of T of quality $q = q(\mathcal{F}, T)$. Let σ be a random member of $\mathcal{F}$, chosen uniformly. For each directed edge ij of T, the probability that $\sigma(i) < \sigma(j)$ is at least $\frac{1}{2} + \frac{1}{2}q$. Therefore, by linearity of expectation, the expected value of $FIT(T, \sigma)$ is at least $(\frac{1}{2} + \frac{1}{2}q)\binom{n}{2}$. By (3) it follows that

$$\frac{1}{2}q\binom{n}{2} \leq bn^{3/2},$$

implying that $q \leq O(n^{-1/2})$, as needed. $\square$

4 Concluding remarks

- By Theorem 1.1, part (ii), most tournaments on n vertices do not admit a realization of quality ϵ when ϵ is much bigger than $1/\sqrt{n}$. It is still of interest to estimate the number $f(n, \epsilon)$ of labelled tournaments on n vertices that admit a realization of quality at least ϵ. Repeating the argument appearing at the end of Section 2 it follows that each such tournament can be realized by a set $\mathcal{F}$ of at most $\frac{4}{\epsilon^2} \log n$ permutations, implying that

$$f(n, \epsilon) \leq (n!)^{4 \log n / \epsilon^2} \ (\ \leq 2^{O(\frac{n \log^2 n}{\epsilon^2})} \ ).$$

- Hurlbert and Kierstead [6] have recently considered a different way to realize tournaments. In this realization, each vertex v of the tournament is assigned a set S_v of k integers, where no integer is assigned to more than a single vertex. For two distinct vertices u and v, the pair uv is a directed edge if and only if in the majority of the ordered pairs (x,y) with $x \in S_u,\ y \in S_v$, x exceeds y. This can be viewed as follows: each vertex v is assigned a k-sided die with the numbers in S_v on its sides. Each pair of vertices can now play by rolling their dice, where the bigger number wins. The direction of the edge connecting u and v is from the player who is more likely to win, to the other player. The authors of [6] proved that every tournament can be realized by an appropriate collection of dice. For a tournament T, they defined the *dice dimension* of T, denoted $dd(T)$, to be the minimum k such that T can be realized by k-sided dice. A simple counting argument shows that for some tournament on n vertices, the dice dimension is at least $\Omega(n/\log n)$, and the authors of [6] proved that the dice dimension of any tournament on n vertices is at most $O(n)$. This can be improved to a (tight) bound of $O(n/\log n)$ by applying the result of Erdős and Moser [4], as follows.

  There is a simple way to obtain from any realization of a tournament $T = (V,E)$ by a set $\mathcal{F}$ of k permutations, a realization of the tournament by k-sided dice. Indeed, if $\mathcal{F} = \{\pi_1, \pi_2, \ldots, \pi_k\}$ define, for each $v \in V$, $S_v = \{jn - \pi_j^{-1}(v) + 1 :\ 1 \le j \le k\}$, where $\pi_j^{-1}(v)$ is the place of v in the permutation π_j. It is not difficult to check that if u precedes v in t of the k permutations, then in precisely $k(k-1)/2 + t$ of the pairs $(x,y) \in S_u \times S_v$, $x > y$. Therefore, the result of [4] mentioned in the introduction implies that the dice dimension of any tournament on n vertices is at most $O(n/\log n)$. Similarly, by Theorem 1.1 here, every tournament can be realized by dice so that for every directed edge uv, the probability that the die of u will beat that of v is at least $1/2 + \Omega(\frac{1}{n^{3/2}\log n})$. On the other hand, linearity of expectation together with the result of de-la Vega mentioned in Section 3 easily imply that for most tournaments T on n vertices, in any dice realization there will be directed edges uv such that the probability that the die of u will beat that of v is at most $1/2 + O(\frac{1}{n^{1/2}})$. Indeed, if for every directed edge of T this probability is at least $1/2(1+q)$, choose, for each vertex v of T, a random element $x_v \in S_v$, and let σ be the ordering of the vertices in a decreasing order of the numbers x_v. For each fixed directed edge uv, the probability that $x_u > x_v$ is at least $1/2(1+q)$. Thus, by linearity of expectation, the expected value of $FIT(T,\sigma)$ is at least $1/2(1+q)\binom{n}{2}$. As, by the result of de-la Vega (Theorem 3.1), for most tournaments T on n vertices, $FIT(T,\pi) \le 1/2\binom{n}{2} + O(n^{3/2})$ for every permutation π, the desired upper estimate for q follows.

- The proof of Theorem 2.1 (and the related proof of Theorem 2.3) can be converted into algorithmic proofs. That is, we can prove the following.

  Proposition 4.1 *There is an absolute, positive constant c such that the following holds. There is a deterministic algorithm, that given a weighted, simple, directed graph $D = (V,E)$ with no* *anti-parallel edges and with maximum degree at most d, finds, in polynomial time, a permutation σ of V such that*

$$FIT(D,\sigma) \geq (\frac{1}{2} + \frac{c}{\sqrt{d}})w(E).$$

One way to prove the above proposition is to first show, using Hölder's Inequality, that for any real random variable X

$$E(|X|) \geq \frac{E(X^2)^{3/2}}{E(X^4)^{1/2}}.$$

Next, observe that if $c_1, \ldots, c_m$ are reals, $\epsilon_1, \ldots, \epsilon_m$ are 4-wise independent, identically distributed random variables, each distributed uniformly on $\{-1, 1\}$ and $X = \sum_{i=1}^m \epsilon_i c_i$, then $E(X^2) = \sum_{i=1}^m c_i^2$ and

$$E(X^4) = \sum_{i=1}^m c_i^4 + 6 \sum_{1 \leq i < j \leq m} c_i^2 c_j^2 \leq 3(\sum_{i=1}^m c_i^2)^2.$$

Therefore, by the previous inequality, $E(|X|) \geq 3^{-1/2}(c_1^2 + \ldots + c_m^2)^{1/2}$. This shows that the assertion of Szarek's Inequality (Lemma 2.2) holds (with a slight loss in the constant factor) even under the assumption that the variables ϵ_i are 4-wise independent, rather than fully independent. It follows that one can obtain an algorithmic version of the proofs of Theorem 2.1 and Theorem 2.3 by checking all points of a small sample space that supports n 4-wise independent random variables ϵ_i as needed. Constructions of such spaces with $O(n^2)$ points appear in [1] (see also [2]), providing the required efficient, deterministic algorithm.

- The proof in Section 3 provides no explicit example of a tournament T on n vertices that admits no realization of quality $\Omega(1)$. Such examples are given by the quadratic residue tournaments. For a prime $p \equiv 3 \pmod 4$, the tournament T_p is the tournament whose set of vertices are the elements of the finite field Z_p, and ij forms a directed edge iff $i - j$ is a quadratic residue. It is proved in [2], Chapter 9, that if we assign weight 1 for each edge of T_p, then, for every permutation σ of Z_p,

$$FIT(T_p, \sigma) \leq (\frac{1}{2} + O(\frac{\log p}{p^{1/2}}))\binom{p}{2}.$$

By the argument of Section 3 this implies that the quality of any realization of T_p does not exceed $O(\log p/\sqrt{p})$. It would be interesting to get rid of the logarithmic factor and find an explicit example of tournaments on n vertices which admit no realization of quality better than $O(1/\sqrt{n})$.

Acknowledgment. I would like to thank Gil Kalai and Glenn Hurlbert for helpful discussions. Part of this work was carried out during a visit at Microsoft Research, Redmond, WA. I would like to thank my hosts at Microsoft for their hospitality.

References

[1] N. Alon, L. Babai and A. Itai, A fast and simple randomized parallel algorithm for the maximal independent set problem, J. Algorithms 7(1986), 567-583.

[2] N. Alon and J. H. Spencer, **The Probabilistic Method**, Second Edition, Wiley, New York, 2000.

[3] W. F. de la Vega, On the maximal cardinality of a consistent set of arcs in a random tournament, J. Combinatorial Theory, Ser. B 35 (1983), 328–332.

[4] P. Erdős and L. Moser, On the representation of directed graphs as unions of orderings, Magyar Tud. Akad. Mat. Kutató Int. Közl. 9 (1964), 125-132.

[5] P. Fishburn, **The Theory of Social Choice**, Princeton University Press, 1973.

[6] G. Hurlbert and H. Kierstead, in preparation.

[7] G. Kalai, Social choice and threshold phenomena, in preparation.

[8] R. Latała and K. Oleszkiewicz, On the best constant in the Khinchin-Kahana inequality, Studia Math. 109 (1994), 101-104.

[9] D. C. McGarvey, A theorem on the construction of voting paradoxes, Econometrica 21 (1953), 608-610.

[10] J. Spencer, Optimal ranking of tournaments, Networks 1 (1971/2), 135-138.

[11] R. Stearns, The voting problem, Amer. Math. Monthly 66 (1959), 761-763.

[12] S. J. Szarek, On the best constants in the Khinchin Inequality, Studia Math. 58 (1976), 197-208.