Combinatorics, Probability and Computing (1997) **6**, 1–15
Copyright © 1997 Cambridge University Press

Imperfections in Random Tournaments

A. D. BARBOUR[1]†, ANANT P. GODBOLE[2]‡ and JINGHUA QIAN[3]

[1] Abteilung für Angewandte Mathematik, Universität Zürich-Irchel, Winterthurerstrasse 190, CH 8057 Zürich, Switzerland

(e-mail: adb@amath.unizh.ch)

[2] Department of Mathematical Sciences, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931-1295, USA

(e-mail: anant@mtu.edu)

[3] Department of Mathematics, Tufts University, Medford, MA 02155, USA

(e-mail: jqian@diamond.tufts.edu)

Received 31 May 1996

A *tournament* T on a set V of n players is an orientation of the edges of the complete graph K_n on V; T will be called a *random tournament* if the directions of these edges are determined by a sequence $\{Y_j : j = 1, \ldots, \binom{n}{2}\}$ of independent coin flips. If (y, x) is an edge in a (random) tournament, we say that y *beats* x. A set $A \subset V$, $|A| = k$, is said to be *beaten* if there exists a player $y \notin A$ such that y beats x for each $x \in A$. If such a y does not exist, we say that A is *unbeaten*. A (random) tournament on V is said to have property S_k if each k-element subset of V is beaten. In this paper, we use the Stein–Chen method to show that the probability distribution of the number W_0 of unbeaten k-subsets of V can be well-approximated by that of a Poisson random variable with the same mean; an improved condition for the existence of tournaments with property S_k is derived as a corollary. A multivariate version of this result is proved next: with W_j representing the number of k-subsets that are beaten by precisely j external vertices, $j = 0, 1, \ldots, b$, it is shown that the joint distribution of $(W_0, W_1, \ldots, W_b)$ can be approximated by a multidimensional Poisson vector with independent components, provided that b is not too large.

1. Introduction and statement of results

Let $V = \{1, 2, \ldots, n\}$ be a finite set. A *tournament* T on V is an orientation of the $\binom{n}{2}$ edges of the complete graph K_n on V. If (x, y) is an oriented edge in a tournament, we say that x *beats* y. A *random tournament* is a tournament on V obtained by choosing independently, for each $1 \leqslant x < y \leqslant n$, one of the edges (x, y) or (y, x) with equal probability. Moon [7] describes applications of the combinatorics of tournaments to areas as diverse as the method of paired comparisons, voting schemes, choice functions and dominance relations in sociometric groups.

T is said to have property S_k (written $T \in S_k$) if for each set A, $|A| = k \geqslant 1$, there exists

† Supported in part by Schweizerischer Nationalfonds Projekte Nr. 20-31262.91 and 21-37354.93.
‡ Supported in part by NSF Grant DMS-9200409. A portion of this work was done during a visit to Universität Zürich in January–February 1993; the travel support derived from Swiss NF Grant Nr. 20-31262.91 is acknowledged.

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

$y \notin A$ such that (y, x) is an oriented edge for each $x \in A$. In other words, $T \in S_k$ if each member of each group $A = \{a_1, \ldots, a_k\}$ of k players is beaten by a player $y = y_A \notin A$. If such a y exists, we shall say that the set A is *beaten*, or *dominated*; if not, it will be said to be *unbeaten*. It is clear that when k is fixed and n is sufficiently large, a *random* tournament is very likely to have property S_k: Erdős [5] showed that if $\binom{n}{k}(1-2^{-k})^{n-k} < 1$, then there is a tournament on n vertices with property S_k; a proof of this fact may be found in Alon, Spencer and Erdős [1, p. 5]. If we define $W_0 = \sum_{j=1}^{\binom{n}{k}} I_j$ as the number of unbeaten k-sets in a random tournament, where I_j is the indicator of the event that the jth set of size k is unbeaten, then it is clear that $\mathbf{E}(W_0) = \binom{n}{k}(1-2^{-k})^{n-k}$, and Erdős' result may be reformulated as the elementary (but far-reaching) observation that

$$\lambda = \mathbf{E}(W_0) < 1 \Rightarrow \mathbf{P}(W_0 = 0) > 0. \tag{1}$$

Throughout this paper, j will represent a k-subset as well as its index; we trust that this dual use of notation will not be confusing.

Let $f(k)$ denote the minimum possible number of vertices of a tournament $T \in S_k$. It can readily be verified that $f(1) = 3$ and $f(2) = 7$; a general lower bound due to Szekeres is $f(k) \geqslant ck2^k$ (see Moon [7]). On the other hand, it is an easy exercise to deduce from Erdős's result that $f(k) \leqslant k^2 2^k (\log 2)(1 + o(1))$, $k \to \infty$: in fact, for $n = k^2 2^k \log 2$, one has $\binom{n}{k}(1-2^{-k})^{n-k} \sim k^k (\log 2)^k e^k \to \infty$, while for $n = k^2 2^k$, $\binom{n}{k}(1-2^{-k})^{n-k} \sim k^k e^k (2/e)^{k^2} \to 0$, showing that $f(k) \leqslant k^2 2^k$. A more careful analysis yields

$$f(k) \leqslant k^2 2^k \log 2 + (1+\epsilon) k 2^k \log k = k^2 2^k \log 2 (1 + o(1)), \tag{2}$$

whatever the value of $\epsilon > 0$, if k is large enough. The above observations reveal how rapidly decreasing the function $\phi(n) = \binom{n}{k}(1-2^{-k})^{n-k}$ is, and provide a natural reason why the elementary result of Erdős has not been bettered (thus far).

Property S_k is undesirable, in a very natural sense; a tournament $T \in S_k$ can, consequently, be thought of as being rather imperfect. We have seen, moreover, that a random tournament with an excessively large number of players has the potential to have the blemish S_k, i.e. satisfies the condition $\mathbf{P}(W_0 = 0) > 0$. The random variable W_0 may, in fact, be used as a measure of the *degree* of imperfection of a random tournament T on n vertices. What, then, can one assert about the probability distribution of W_0? Exact expressions are clearly quite intractable, but, since $W_0 = \sum_{j=1}^{\binom{n}{k}} I_j$, where $\mathbf{P}(I_j = 1) = (1-2^{-k})^{n-k}$ is small if n is substantially larger than 2^k, one might expect that the distribution of W_0 is close to Poisson for such values of n. On the other hand, the dependence between the I_j's is global, which might lead to the failure of the Poisson paradigm unless n is *even larger* than indicated above. Our first result, proved in Section 2, makes the above statements precise:

Theorem 1. *Let W_0 denote the number of unbeaten k-sets in a random tournament T on n vertices, and let* $\mathrm{Po}(\lambda)$ *denote the Poisson distribution with mean $\lambda = \binom{n}{k}(1-2^{-k})^{n-k}$. Assume that $2 \leqslant k \leqslant 8$, or that $k \geqslant 9$ and $n \geqslant 2k^2$. Then*

$$\begin{aligned} d_{\mathrm{TV}}(\mathscr{L}(W_0), \mathrm{Po}(\lambda)) &= \sup_{A \subseteq \mathbf{Z}^+} |\mathbf{P}(W_0 \in A) - \mathrm{Po}(\lambda)\{A\}| \\ &\leqslant \left(1 - \frac{1}{2^k}\right)^{n-k} + \frac{\lambda k^2}{4^k} A_k \{A_k + 2B_{nk}\}, \end{aligned}$$

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

where $A_k = (1-2^{-k})^{-k} = 1+O(k2^{-k})$ $(k\to\infty)$, *and* B_{nk} *is defined by*

$$B_{nk} = \left(1+\frac{2}{2^k(2^k-1)}\right)^n + k\exp\varsigma(k,n),$$

with

$$\varsigma(k,n) = \max\left\{\log\left(\frac{k^2}{n}\right)+\frac{4n}{2^k(2^k-1)}, \log\left(\frac{1}{(k-1)!}\left(\frac{2k^2}{n}\right)^{k-2}\right)+\frac{n}{2(2^k-1)}\right\}.$$

The proof of Theorem 1 relies on the Stein–Chen method of Poisson approximation (see Barbour, Holst and Janson [4] for theoretical aspects (and applications) of the *coupling* approach to Stein–Chen approximation that we employ here). We shall see, moreover, that Theorem 1 yields, as a corollary, a significant improvement over Erdős's result (1).

We next consider a more complicated question that is motivated by the notion of tournaments with 'property $S(k,m)$', i.e. one for which each k-set is beaten by at least m players [see [7] for more details]: Given a random tournament on n vertices, what can be asserted about the overall structure of the numbers of groups of k players that are each beaten by exactly j others ($j=0,1,\ldots,b$)? Specifically, how does one approximate the *joint* distribution of the random vector $(W_0, W_1, \ldots, W_b)$, where W_j denotes the number of k-subsets of $\{1,2,\ldots,n\}$ that are beaten by precisely j external vertices? Our second result is in the spirit of other multivariate Poisson approximations obtained, e.g. by Arratia and Tavaré [3], Arratia, Barbour and Tavaré [2] and Godbole, Skipper and Sunley [6], in the context of random permutations, the Ewens sampling formula, and random edge colourings of K_n respectively; given the nature of this result, moreover, we must evidently restrict ourselves to values of b that are smaller, by an order of magnitude, than the expected number $(n-k)/2^k$ of vertices that beat any particular k-set. Letting $\lambda_j = \mathbf{E}(W_j) = \binom{n}{k}\binom{n-k}{j}2^{-kj}(1-2^{-k})^{n-k-j}$ denote the expected number of k-subsets that are dominated by j players, and setting $\Lambda_b = \sum_{j=0}^{b}\lambda_j$, we have:

Theorem 2. *Let* W_j and λ_j *be as in the preceding paragraph. Assume that* $2\leqslant k\leqslant 8$, *or that* $k\geqslant 9$ *and* $n\geqslant 2k^2$. *Then the joint distribution of* $(W_0, W_1, \ldots, W_b)$ *can be approximated by a multivariate Poisson distribution* $(\mathrm{Po}(\lambda_0), \mathrm{Po}(\lambda_1), \ldots, \mathrm{Po}(\lambda_b))$ *with independent components, with total variation error given by*

$$d_{\mathrm{TV}}\left(\mathscr{L}(\{W_j\}_{j=0}^b), \prod_{j=0}^{b}\mathrm{Po}(\lambda_j)\right)$$
$$\leqslant \Lambda_b\left\{2R_b+\frac{\Lambda_b k^2}{4^k}A_k^2\left(1+\frac{2^k b}{n-k-b}\right)\left(1+4B_{nk}+\frac{2k^2}{n-k}+\frac{R_b 2^k b}{n-k-b}\right)\right\},$$

where $A_k = (1-2^{-k})^{-k}$ *and* B_{nk} *are as in Theorem* 1, *and* $R_b = \Lambda_b/\binom{n}{k}$.

For fixed k and n, the bound in Theorem 2 increases with b, as is to be expected. Clearly,

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

if the bound is to be small, we must have $n \geqslant 2^k$, and the quantity $\Lambda_b k 2^{-k}$ must be small. Since Λ_b can be expressed in terms of binomial probabilities as

$$\Lambda_b = \binom{n}{k} \mathrm{Bi}(n-k, 2^{-k})\{[0, b]\},$$

this implies that b must certainly be smaller than $n2^{-k}$. Under these circumstances, the bound is of order $\Lambda_b^2 k^2 4^{-k}$, and takes the value $\Lambda_b^2 k^2 4^{-k}\{1+o(1)\}$ if $b = o(n2^{-k})$; good Poisson approximation is obtained whenever $\Lambda_b \ll k^{-1} 2^k$. The order of the error bound is the same as that of Theorem 1, but for the presence of a Λ_b^2 multiplier in place of the λ. The square is due to the fact that the univariate Theorem 2.C of Barbour, Holst and Janson [4] has the 'magic' factor $(1-e^{-\lambda})/\lambda$, which is not available in their multivariate Theorem 10.J. Partial reinstatement of the magic factor is possible, however, as in Theorem 10.K in Barbour, Holst and Janson [4], although it may be verified that our results do not improve significantly as a result.

To summarize, Theorem 1 reveals that the distribution of the number of unbeaten k-sets is approximately Poisson, and yields an improved condition for the existence of tournaments with property S_k. Theorem 2, on the other hand, provides an overall view of the numbers of k-sets that are beaten by various numbers of external vertices, showing that the 'big picture' can be modeled by a Poisson process with independent components. Given these results, it is not surprising that the variables W_j can, for $j \geqslant 1$, be *individually* approximated by Poisson random variables with an even higher degree of precision than that given by Theorem 2. Results along these lines have been proved in Qian [8], but are not included in this paper.

2. Proof of Theorem 1

By Formula (1.2) of Section 2.1 in Barbour, Holst and Janson [4],

$$\begin{aligned} d_{\mathrm{TV}}(\mathscr{L}(W_0), \mathrm{Po}(\lambda)) &\leqslant \frac{1-e^{-\lambda}}{\lambda} \sum_{j=1}^{\binom{n}{k}} \left(\pi_j^2 + \pi_j \sum_{i \neq j} \mathbf{P}(I_i \neq J_{ij}) \right) \\ &\leqslant \pi_1 + \sum_{i \neq 1} (P_{1i} + P_{2i}), \end{aligned} \tag{3}$$

where $\pi_j = \mathbf{P}(I_j = 1) = (1-2^{-k})^{n-k}$, the variables $\{J_i\}_{i=1}^{\binom{n}{k}} = \{J_{ij}\}_{i=1}^{\binom{n}{k}}$ satisfy

$$\mathscr{L}\left(J_{ij} : 1 \leqslant i \leqslant \binom{n}{k} \right) = \mathscr{L}\left(I_i : 1 \leqslant i \leqslant \binom{n}{k} \middle| I_j = 1 \right) \tag{4}$$

and P_{1i} and P_{2i} are defined by

$$P_{1i} = \mathbf{P}(I_i = 1, J_{i1} = 0)$$

and

$$P_{2i} = \mathbf{P}(I_i = 0, J_{i1} = 1)$$

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

respectively. The coupled random variables $\{J_{i1}\}$ are defined as follows: If $I_1 = 1$ (i.e. if the 1st set is unbeaten), we 'do nothing', letting $J_{i1} = I_i$ for each i. If $I_1 = 0$, on the other hand, we consider the set $C(\neq \varnothing)$ of players who beat the 1st set. For each $c \in C$, we independently choose, with equal probability, one of the $2^k - 1$ orientations of the edges between c and members of the first set that lead to c *not* dominating the set 1, and redirect the edges accordingly, setting $J_{i1} = 1$ if the ith k-set is unbeaten after the adjustment ($J_{i1} = 0$ otherwise). It is evident that the J_{i1}'s satisfy (4), providing an accurate model for the global behaviour of the indicators, conditional on the first k-set being unbeaten. We begin by estimating P_{1i}. It is clear that the event $\{I_i = 1, J_{i1} = 0\} = \{I_i = 1, I_1 = 0, J_{i1} = 0\}$ occurs only if there is at least one player in i that beats 1 and, in the process of 'redirecting edges' (i.e. achieving the coupling), the set i is beaten by a member of 1. It follows, since there are $\binom{k}{r}\binom{n-k}{k-r}$ sets i such that $|i \cap 1| = r$, that

$$\begin{aligned}
\sum_{i \neq 1} P_{1i} &\leqslant \sum_{r=0}^{k-1} \binom{k}{r}\binom{n-k}{k-r}\left(1-\frac{1}{2^k}\right)^{n-2k+r} \sum_{s=1}^{k-r} \binom{k-r}{s}\frac{1}{s^k}\left(1-\frac{1}{s^k}\right)^{k-r-s} \frac{k-r}{2^{k-2}}\left(\frac{2^{k-1}}{2^k-1}\right)^s \\
&= \sum_{r=0}^{k-1} \binom{k}{r}\binom{n-k}{k-r}\left(1-\frac{1}{2^k}\right)^{n-k} \frac{k-r}{2^k}\left\{\left(1+\frac{2^k}{(2^k-1)^2}\right)^{k-r} - 1\right\} \\
&\leqslant \lambda \sum_{r=0}^{k-1} \frac{\binom{k}{r}\binom{n-k}{k-r}}{\binom{n}{k}}\left(\frac{k}{2^k-1}\right)^2\left(1+\frac{2^k}{(2^k-1)^2}\right)^{k-1} \\
&\leqslant \frac{\lambda k^2}{4^k} A_k^2, \qquad (5)
\end{aligned}$$

with $A_k = (1-2^{-k})^{-k}$. The first inequality above follows since the $s (1 \leqslant s \leqslant k-r)$ elements of i that beat 1 originally must now be beaten by some element of $1 \backslash i$ (there are $k-r$ of these). The probability of this event is $(2^{k-1}/(2^k-1))^s$. Also, the other $k-s$ elements of i must have been beaten by the above element of $1 \backslash i$ *before* the arrow reversal.

Notice next that the event $\{I_i = 0, J_{i1} = 1\} = \{I_i = 0, I_1 = 0, J_{i1} = 1\}$ can only occur provided that (a) the sets 1 and i have a non-empty intersection, (b) there is at least one player who beats both 1 and i, and, (c) when the coupling is implemented, these players do not beat 1 or i any more. Finally, observe that $\mathbf{P}(I_i = 0, J_{i1} = 1) = 0$ if i is beaten by any point which does not beat 1. It follows that

$$\begin{aligned}
\sum_{i \neq 1} P_{2i} &\leqslant \sum_{r=1}^{k-1} \binom{k}{r}\binom{n-k}{k-r}\frac{n}{2^{2k-r}}\left\{\frac{1}{2^{2k-r}} + \left(1-\frac{1}{2^k}\right)\right\}^{n-2k+r-1} \\
&\leqslant \frac{n\lambda}{4^k} \sum_{r=1}^{k-1} \frac{\binom{k}{r}\binom{n-k}{k-r}}{\binom{n}{k}} 2^r \left\{1+\frac{2^r}{2^k(2^k-1)}\right\}^n A_k, \qquad (6)
\end{aligned}$$

where A_k is as in (5). Now $\binom{k}{r}\binom{n-k}{k-r}/\binom{n}{k} \leqslant k^{2r}/\{n^r r!\}$, so that we thus have

$$\sum_{i \neq 1} P_{2i} \leqslant \frac{2k^2\lambda}{4^k} A_k \sum_{r=1}^{k-1} \phi(r, n), \qquad (7)$$

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

where

$$\phi(r,n) = \frac{1}{r!}\left(\frac{2k^2}{n}\right)^{r-1}\left\{1+\frac{2^r}{2^k(2^k-1)}\right\}^n.$$

Now, taking logarithms, we can write

$$\log\phi(r,n) \leqslant b_n(r) := (r-1)\log\left(\frac{2k^2}{n}\right)+\frac{n2^r}{2^k(2^k-1)}-\log(r!).$$

The function b_n has derivative

$$b_n'(r) = \log\left(\frac{2k^2}{n}\right)+\frac{n2^r\log 2}{2^k(2^k-1)}-\psi(r+1),$$

where $\psi(x)$ denotes $\frac{d}{dx}\log\Gamma(x)$. The representation

$$\psi(x) = \Gamma'(x)/\Gamma(x) = -\gamma-\frac{1}{x}+\sum_{n=1}^{\infty}\frac{x}{n(x+n)}, \tag{8}$$

where γ is Euler's constant, shows that b_n' is a convex function in $r \geqslant 0$ satisfying

$$b_n'(1) = \log\left(\frac{2k^2}{n}\right)+\frac{2n\log 2}{2^k(2^k-1)}-\psi(2) \leqslant \frac{2n\log 2}{2^k(2^k-1)}-(1-\gamma) \leqslant 0$$

if

$$2k^2 \leqslant n \leqslant \frac{2^k(2^k-1)(1-\gamma)}{2\log 2}. \tag{9}$$

Thus for such n and k,

$$\max_{2\leqslant r\leqslant k-1}\log\phi(r,n) \leqslant \max\{b_n(2), b_n(k-1)\},$$

so that

$$\begin{aligned}\sum_{i\neq 1}P_{2i} &\leqslant \frac{2k^2\lambda}{4^k}A_k\{\phi(1,n)+(k-2)\,e^{b_n(2)\vee b_n(k-1)}\}\\ &\leqslant \frac{2k^2\lambda}{4^k}A_k B_{nk},\end{aligned} \tag{10}$$

with $B_{nk} = \phi(1,n)+k\,e^{b_n(2)\vee b_n(k-1)}$. If (9) is *not* satisfied, then it is not immediately clear whether the bound in (10) still holds. Suppose, first, that $n \geqslant 2^k(2^k-1)(1-\gamma)/2\log 2$. In this case, (8) yields

$$b_n''(2) = \frac{4n\log^2 2}{2^k(2^k-1)}-\frac{1}{9}-\sum_{m=1}^{\infty}\frac{1}{(3+m)^2} \geqslant \frac{4(1-\gamma)\log^2 2}{2\log 2}-\frac{\pi^2}{6}+1.25 > 0,$$

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

implying, by the convexity of b_n', that b_n'' is positive for each $r \geqslant 2$, and thus that $\max_{2 \leqslant r \leqslant k-1} b_n(r) = b_n(2) \vee b_n(k-1)$. It follows that (10) continues to provide an upper bound.

If $n \leqslant 2k^2$, on the other hand, it is no longer the case that the maximum of b_n, over the range $2, 3, \ldots, k-1$, is attained at an endpoint. A simple *Mathematica*™ program revealed, however, that for $k \leqslant 8$, $n \leqslant 2k^2$, the series $\sum_{r=2}^{k} b_n(r)$ *is* dominated by k times the contribution of the larger of $b_n(2)$ and $b_n(k-1)$, and, moreover, that the quantity $(k-2)\exp\{b_n(2) \vee b_n(k-1)\}$ is too small to provide an upper bound on $\sum \phi(r,n)$ for *all* the cases under consideration. This observation led us, in fact, to define B_{nk} as we did in (10) above. Theorem 1 now follows on combining (3), (5) and (10) with the above remarks; it is quite conceivable that more extensive numerical calculations will lead to a widening of its domain of applicability. □

Corollary. *Asymptotically as* $k \to \infty$, W_0 *can be well-approximated by a Poisson r.v. provided that* $2^k \leqslant n \leqslant k^2 2^k \log 4$, $\lambda = o(4^k/k^2)$. *A tournament with property* S_k *exists, moreover, if* $\lambda \leqslant (\log 4 - o(1))\, k$ $(k \to \infty)$.

Proof. We begin by using the function

$$c_k(n) = b_n(2) - b_n(k-1) = (k-3)\log\left(\frac{n}{2k^2}\right) + \log((k-1)!) - \log 2 - \frac{1}{2}\left(\frac{n}{2^k}\right)\left(\frac{2^k-8}{2^k-1}\right)$$

to exhibit one of several alternative bounds that may be extracted from the proof of Theorem 1. Now c_k is a concave function of n $(n > 0)$, so that

$$\min_{n_1 \leqslant n \leqslant n_2} c_k(n) = \min\{c_k(n_1), c_k(n_2)\}, \tag{11}$$

and hence

$$\sum_{r=2}^{k-1} \phi(r,n) \leqslant (k-2)\, e^{b_n(2)} \tag{12}$$

for all $n_1 \leqslant n \leqslant n_2$, if $\min\{c_k(n_1), c_k(n_2)\} > 0$ and (9) is satisfied. Note that (9) holds if, for example, $k \geqslant 9$ and $2^k \leqslant n \leqslant k^2 2^k \log 4$. For $n_1 = 2^k$, we have $c_k(2^k) \sim k^2 \log 2$ $(k \to \infty)$, and calculation shows that in fact $c_k(2^k) > 0$ for all $k \geqslant 4$. For $n_2 = k^2 2^k \log 2$, $c_k(k^2 2^k \log 2) \sim \frac{1}{2} k^2 \log 2$ $(k \to \infty)$, and calculation shows that $c_k(k^2 2^k \log 2) > 0$ for all $k \geqslant 5$. For the critical $n_2 = k^2 2^k \log 4$, $c_k(k^2 2^k \log 4) \sim k \log k$ $(k \to \infty)$ and $c_k(k^2 2^k \log 4) > 0$ for all $k \geqslant 32$. Thus, collecting (7), (12) and the above, we have the estimate

$$\sum_{i=1} P_{2i} \leqslant \frac{2k^2 \lambda}{4^k} A_k\{\phi(1,n) + (k-2)\, e^{b_n(2)}\}$$

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

for $k \geqslant 9$ and $2^k \leqslant n \leqslant k^2 2^k \log 2$ or $k \geqslant 32$ and $2^k \leqslant n \leqslant k^2 2^k \log 4$. This in turn implies that, in these ranges of n,

$$\begin{aligned}\sum_{i+1} P_{2i} &\leqslant \frac{2k^2\lambda}{4^k} A_k \left\{ \left(1 + \frac{2}{2^k(2^k-1)}\right)^n + \frac{(k-2)k^2}{n} e^{\frac{4n}{2^k(2^k-1)}} \right\} \\ &\leqslant \frac{2k^2\lambda}{4^k} A_k \left\{ 1 + \left(\frac{2n}{2^k(2^k-1)} + \frac{(k-2)\,k^2}{n} \right) e^{\frac{4n}{2^k(2^k-1)}} \right\} \\ &= \frac{2k^2\lambda}{4^k} A_k\, C_{nk},\end{aligned}$$

with $C_{nk} = 1 + O(k^3 2^{-k})\,(k \to \infty)$ for $2^k \leqslant n \leqslant k^2 2^k \log 4$. We thus obtain the bound

$$d_{\mathrm{TV}}(\mathscr{L}(W_0), \mathrm{Po}(\lambda)) \leqslant \left(1 - \frac{1}{2^k}\right)^{n-k} + \frac{\lambda k^2}{4^k} A_k \{A_k + 2C_{nk}\},$$

where $A_k, C_{nk} \to 1$ $(k \to \infty)$ for $2^k \leqslant n \leqslant k^2 2^k \log 4$. The first assertion follows. In particular, we have

$$\mathbf{P}(W_0 = 0) \geqslant e^{-\lambda} - \frac{3\lambda k^2}{4^k}(1 + o(1))$$

uniformly in n, $2^k \leqslant n \leqslant k^2 2^k \log 4$, so that a tournament with property S_k exists if

$$\frac{3\lambda\, e^{\lambda} k^2}{4^k}(1 + o(1)) < 1,$$

that is, if $\lambda \leqslant (\log 4 - o(1))\, k$, or, if $\eta(k) \leqslant n \leqslant k^2 2^k \log 4$, where $\eta(k)$ is certainly smaller than the upper bound on $f(k)$ stated in (2). If $n \geqslant k^2 2^k \log 4$, a tournament $T \in S_k$ exists by monotonicity. This proves the corollary. □

Example. Consider a random round-robin tournament between 6693 equally capable players. The expected number of unbeaten 7-subsets is $\binom{6693}{7}(1 - 1/2^7)^{6686} \approx 2$. Erdős's result is not applicable, since $\lambda > 1$. It may be verified that $c_7(6693) < 0$, where $c_k(n)$ is as defined in the proof of the above corollary, so that (10) yields

$$\sum_{i+1} P_{2i} \leqslant \frac{2k^2\lambda}{4^k} A_k \{\phi(1, n) + (k-2)\, e^{b_n(k-1)}\}.$$

Computation of the quantities in question reveals that $A_k \leqslant 1.0565$, $\phi(1, n) \leqslant 2.279$, and $e^{b_n(k-1)} = e^{b_n(6)} \leqslant 0.260$ and hence, using (3) and (5), $|\mathbf{P}(W_0 = 0) - e^{-\lambda}| \leqslant 0.05192$. Since also $e^{-\lambda} \approx e^{-2} \approx 0.135$, it follows that $P(W_0 = 0) > 0$. Of particular significance is the fact that Theorem 1 provides a total variation error bound that enables one to estimate more than just point probabilities: for example, the probability that there are between 2 and 5 unbeaten 7-subsets satisfies

$$0.52548 < P(W_0 \in \{2, 3, 4, 5\}) \leqslant 0.62932.$$

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

3. Proof of Theorem 2

By Theorem 10.J in Barbour, Holst and Janson [4],

$$d_{\mathrm{TV}}\left(\mathscr{L}(\{W_j\}_{j=0}^b), \prod_{j=0}^b \mathrm{Po}(\lambda_j)\right) \leqslant \sum_{\alpha\in\Gamma}\left\{\pi_\alpha^2 + \sum_{\beta\neq\alpha}\pi_\alpha \mathbf{P}(J_{\beta\alpha}\neq I_\beta)\right\} \tag{13}$$

where

$$\Gamma = \left\{(i,j), 1\leqslant i\leqslant \binom{n}{k}, 0\leqslant j\leqslant b\right\};$$

$$I_{i,j} = \begin{cases} 1 & \text{if the } i\text{th } k\text{-subset is beaten by exactly } j \text{ external vertices,} \\ 0 & \text{otherwise;} \end{cases}$$

$$\pi_{i,j} = P(I_{i,j}=1) = \binom{n-k}{j}2^{-kj}(1-2^{-k})^{n-k-j} := \rho_j;$$

$$W_j = \sum_{i=1}^{\binom{n}{k}} I_{i,j}; \quad \lambda_j = E(W_j) = \binom{n}{k}\rho_j,$$

and where, for each (i,j), the coupled indicator variables $J_{i',j'} = J_{(i',j')(i,j)}$ satisfy the condition

$$\mathscr{L}\{(J_{(i',j')(i,j)}:(i',j')\in\Gamma)\} = \mathscr{L}\{(I_{i',j'},(i',j')\in\Gamma)\,|\,I_{i,j}=1\}. \tag{14}$$

The coupling is constructed as follows. If $I_{i,j}=1$, we do nothing, letting $J_{(i',j')(i,j)} = I_{i',j'}$ for each (i',j'). If, on the other hand, $I_{i,j}=0$, then i is either beaten by the s vertices in a set $A := \{a_1, a_2, \ldots, a_s\}$ where $s<j$, or by the t vertices in a set $B := \{b_1, b_2, \ldots, b_t\}$ where $t>j$. In the former case, we choose $j-s$ vertices randomly from the $n-k-s$ vertices in $\{1,2,\ldots,n\}\backslash(i\cup A)$ (so that the probability that any one is picked is $(j-s)/(n-k-s)$) and 'redirect arrows' so that each of these $j-s$ vertices beats the set i. In the latter case, we randomly select j of the t vertices in B, and continue to have the selected vertices beat the set i. For each of the other $t-j$ vertices in B, we realign the arrows to one of the 2^k-1 possibilities as in the proof of Theorem 1, chosen at random. Finally, we let $J_{(i',j')(i,j)}=1$ if the set i' is beaten by exactly j' vertices after the above-described coupling is implemented, with $J_{(i',j')(i,j)}=0$ otherwise. A little reflection reveals that (14) is indeed satisfied. Furthermore, since the intersection patterns of each of the $\binom{n}{k}$ k-subsets of V are identical, and since, for a fixed j, $\pi_{i,j}=\rho_j$ is the same for each i, (13) reduces to

$$\begin{aligned} d_{\mathrm{TV}}\left(\mathscr{L}(\{W_j\}_{j=0}^b), \prod_{j=0}^b \mathrm{Po}(\lambda_j)\right) &= \binom{n}{k}\sum_{j=0}^b \rho_j^2 + \binom{n}{k}\sum_{j=0}^b \sum_{(i',j')\neq(1,j)} \rho_j \mathbf{P}(I_{i',j'}\neq J_{i',j'}) \\ &\leqslant \Lambda_b R_b + \sum_{j=0}^b \lambda_j \sum_{(i',j')\neq(1,j)} \mathbf{P}(I_{i',j'}\neq J_{i',j'}), \end{aligned} \tag{15}$$

where $J_{i',j'} = J_{(i',j')(1,j)}$.

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

Consider the second term on the right side of (15). Clearly,

$$\begin{aligned}\sum_{(i',j')\neq(1,j)} \mathbf{P}(I_{i',j'} \neq J_{i',j'}) &= \sum_{i'=1,j'\neq j} \mathbf{P}(I_{i',j'} \neq J_{i',j'}) + \sum_{i'\neq 1,0\leqslant j'\leqslant b} \mathbf{P}(I_{i',j'} \neq J_{i',j'}) \\ &= \sum_{\substack{j'=0\\ j'\neq j}}^{b} \mathbf{P}(I_{1,j'} = 1) + \sum_{i'\neq 1,0\leqslant j'\leqslant b} \mathbf{P}(I_{i',j'} \neq J_{i',j'}) \\ &\leqslant R_b + \Sigma_1 + \Sigma_2, \end{aligned} \tag{16}$$

where

$$\Sigma_1 = \sum_{|i'\cap 1|=0,0\leqslant j'\leqslant b} \mathbf{P}(I_{i',j'} \neq J_{i',j'})$$

and

$$\Sigma_2 = \sum_{i'\neq 1,|i'\cap 1|\neq 0,0\leqslant j'\leqslant b} \mathbf{P}(I_{i',j'} \neq J_{i',j'})$$

represent, respectively, the contributions to (16) of k-sets that are disjoint from the set 1 or that intersect it. We begin by analysing Σ_1. Notice that our coupling has been constructed in a way that guarantees [when $i' \cap 1 = \varnothing$] that its effect on i' is one-directional. Specifically, we find, that after the coupling is implemented, i' is either beaten by fewer (or the same number of) vertices, or by more (or the same number of vertices) according as 1 was dominated by fewer than, or more than, j players. We thus have

$$\Sigma_1 = \Sigma_{11} + \Sigma_{12}, \tag{17}$$

where Σ_{11} and Σ_{12} correspond to 1 being dominated by fewer than or more than j vertices, respectively. So, if $N(1)$ denotes the number of vertices which initially dominate 1, we can write

$$\begin{aligned}\Sigma_{11} &= \sum_{|i'\cap 1|=0,0\leqslant j'\leqslant b} \mathbf{P}(I_{i',j'} = 0, J_{i',j'} = 1, N(1) < j) \\ &\quad + \sum_{|i'\cap 1|=0,0\leqslant j'\leqslant b} \mathbf{P}(I_{i',j'} = 1, J_{i',j'} = 0, N(1) < j) \\ &= \binom{n-k}{k} \sum_{j'=0}^{b} [\mathbf{P}(A_{j,j'}) + \mathbf{P}(B_{j,j'})]. \end{aligned} \tag{18}$$

In (18), $A_{j,j'}$ represents the event that 1 is beaten by fewer than j, and i' by more than j' vertices initially, and that exactly j' players best i' after the coupling is implemented. Likewise, $B_{j,j'}$ is defined as the event that 1 is beaten by fewer than j players, and that i' is beaten by j' (less than j') individuals before (after) the coupling.

Notice that $A_{j,j'}$ occurs only if (i) 1 is initially dominated by fewer than j vertices, none of which are in the set i'; (ii) i' is initially beaten by j' vertices outside 1 and at least one from

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

within 1; and, finally, (iii) at least one vertex in i' is chosen to bring the number beating 1 up to j. It follows that

$$\mathbf{P}(A_{j,j'}) \leqslant \sum_{s=0}^{j-1} \binom{n-2k}{s} \frac{1}{2^{ks}} \left(1-\frac{1}{2^k}\right)^{n-2k-s} \binom{n-2k}{j'} \frac{1}{2^{kj'}} \left(1-\frac{1}{2^k}\right)^{n-2k-j'} \left(1-\left(1-\frac{1}{2^k}\right)^k\right) \left(\frac{k(j-s)}{n-k-s}\right)$$

$$\leqslant A_k^2 R_b \rho_{j'} \frac{k}{2^k} \frac{kb}{n-k-b}; \tag{19}$$

in (19), $A_k = (1-2^{-k})^{-k}$ as before, while $(1-(1-2^{-k})^k)$ represents the probability that i' is beaten by at least one member of 1.

In a similar fashion, one may see that a necessary condition for $B_{j,j'}$ to occur is that 1 is beaten by fewer than j players, none of whom are from i'; i' is initially beaten by exactly j' vertices, of which at most $j'-1$ are from outside 1; and that at least one vertex in i' is chosen to beat 1. It follows, under the above set of conditions, that no member of 1 dominates i' any more, and that i' is beaten by at most $j'-1$ vertices. We thus have

$$\begin{aligned}
\mathbf{P}(B_{j,j'}) &\leqslant \sum_{s=0}^{j-1} \binom{n-2k}{s} \frac{1}{2^{ks}} \left(1-\frac{1}{2^k}\right)^{n-2k-s} \sum_{r=0}^{j'-1} \binom{n-2k}{r} \frac{1}{2^{kr}} \left(1-\frac{1}{2^k}\right)^{n-2k-r} \\
&\quad \times \binom{k}{j'-r} \frac{1}{2^{k(j'-r)}} \left(1-\frac{1}{2^k}\right)^{k-(j'-r)} \left(\frac{k(j-s)}{n-k-s}\right) \\
&\leqslant A_k \left(\sum_{s=0}^{j-1} \rho_s\right) \binom{n-k}{j'} \sum_{r=0}^{j'-1} \frac{\binom{n-2k}{r}\binom{k}{j'-r}}{\binom{n-k}{j'}} \frac{1}{2^{kj'}} \left(1-\frac{1}{2^k}\right)^{n-k-j'} \left(\frac{kb}{n-k-b}\right) \\
&\leqslant A_k \left(\frac{kb}{n-k-b}\right) \rho_{j'} R_b \left(\frac{kb}{n-k}\right).
\end{aligned} \tag{20}$$

Inequalities (18) through (20) now yield

$$\begin{aligned}
\Sigma_{11} &\leqslant \binom{n-k}{k} A_k^2 \sum_{j'=0}^{b} \frac{R_b k^2 b}{2^k (n-k-b)} \rho_{j'} \left\{1+\frac{2^k b}{n-k}\right\} \\
&\leqslant A_k^2 \frac{\Lambda_b R_b k^2 b}{2^k (n-k-b)} \left\{1+\frac{2^k b}{n-k}\right\}.
\end{aligned} \tag{21}$$

For Σ_{12}, recall that if 1 is dominated by more than j vertices, i' cannot be beaten by fewer players than before. The contribution of Σ_{12} to (17) is thus given by

$$\begin{aligned}
\Sigma_{12} &= \sum_{\substack{|i' \cap 1|=0 \\ 0 \leqslant j' \leqslant b}} \mathbf{P}(I_{i',j'} = 1, J_{i',j'} = 0, N(1) > j) \\
&\quad + \sum_{\substack{|i' \cap 1|=0 \\ 0 \leqslant j' \leqslant b}} \mathbf{P}(I_{i',j'} = 0, J_{i',j'} = 1, N(1) > j) \\
&= \binom{n-k}{k} \sum_{j'=0}^{b} [\mathbf{P}(C_{j,j'}) + \mathbf{P}(D_{j,j'})],
\end{aligned} \tag{22}$$

where $C_{j,j'}$ (resp. $D_{j,j'}$) is the event that 1 is beaten by more than j vertices, and that i' is dominated by exactly j' (fewer than j') players *initially*, but by more than j' (exactly j')

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

vertices *after* the coupling is carried out. Notice that for $C_{j,j'}$ to occur requires that 1 is beaten by at least one vertex from i'; that i' is initially dominated by exactly j' players, none of whom are members of 1; and that redirection leads to at least one vertex in 1 beating i'. It follows that

$$\begin{aligned}
\mathbf{P}(C_{j,j'}) &\leqslant \sum_{l=1}^{k} \binom{k}{l} \frac{1}{2^{kl}} \left(1-\frac{1}{2^k}\right)^{k-l} \binom{n-2k}{j'} \frac{1}{2^{kj'}} \left(1-\frac{1}{2^k}\right)^{n-2k-j'} \frac{k}{2^{k-l}} \left(\frac{2^{k-1}}{2^k-1}\right)^l \\
&\leqslant A_k^2 \sum_{l=1}^{k} \binom{k}{l} \frac{1}{2^{kl}} \left(1-\frac{1}{2^k}\right)^{k-l} \rho_{j'} \frac{k}{2^k} \\
&= A_k^2 \left(1-\left(1-\frac{1}{2^k}\right)^k\right) \frac{k}{2^k} \rho_{j'} \\
&\leqslant \frac{k^2 A_k^2}{4^k} \rho_{j'}. \qquad (23)
\end{aligned}$$

Finally, a necessary condition for $D_{j,j'}$ to occur is as follows. Initially, set 1 must be beaten by more than j vertices, at least one of which must be from i'. The set i', in turn, is to be dominated by $r < j'$ vertices, which must *all* be outside the set 1. Finally, in the process of redirecting arrows, exactly $j'-r$ vertices in 1 end up beating i', so that $J_{i'j'} = 1$. We thus have

$$\begin{aligned}
\mathbf{P}(D_{j,j'}) &\leqslant \sum_{l=1}^{k} \binom{k}{l} \frac{1}{2^{kl}} 1-\frac{1}{2^k}\Big)^{k-l} \sum_{r=0}^{j'-1} \binom{n-2k}{r} \frac{1}{2^{kr}} \left(1-\frac{1}{2^k}\right)^{n-2k-r} \\
&\quad \times \binom{k}{j'-r} \left(\frac{1}{2^{k-l}}\right)^{j'-r} \left\{\frac{2^{k-(j'-r)}}{2^k-1}\right\}^l \\
&\leqslant A_k \sum_{l=1}^{k} \binom{k}{l} \frac{1}{2^{kl}} \left(1-\frac{1}{2^k}\right)^{k-l} \sum_{r=0}^{j'-1} \binom{n-2k}{r} \frac{1}{2^{kr}} \left(1-\frac{1}{2^k}\right)^{n-2k-r} \binom{k}{j'-r} \left(\frac{1}{2^k}\right)^{j'-r} \\
&\leqslant A_k \frac{k}{2^k} \binom{n-k}{j'} \frac{1}{2^{kj'}} \sum_{r=0}^{j'-1} \left(1-\frac{1}{2^k}\right)^{n-2k-r} \frac{\binom{n-2k}{r}\binom{k}{j'-r}}{\binom{n-k}{j'}} \\
&\leqslant A_k \frac{k}{2^k} \binom{n-k}{j'} \frac{1}{2^{kj'}} \left(1-\frac{1}{2^k}\right)^{n-2k-j'} \left(\frac{kj'}{n-k}\right) \\
&\leqslant \frac{A_k^2 k^2 b}{2^k (n-k)} \rho_{j'}. \qquad (24)
\end{aligned}$$

Inequalities (22) through (24) thus reveal that

$$\begin{aligned}
\Sigma_{12} &= \binom{n-k}{k} \sum_{j'=0}^{b} (\mathbf{P}(C_{j,j'}) + \mathbf{P}(D_{j,j'})) \\
&\leqslant \binom{n-k}{k} A_k^2 \sum_{j'=0}^{b} \rho_{j'} \frac{k^2}{4^k} \left\{1+\frac{2^k b}{n-k}\right\} \\
&\leqslant A_k^2 \frac{\Lambda_b k^2}{4^k} \left\{1+\frac{2^k b}{n-k}\right\}, \qquad (25)
\end{aligned}$$

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

so that the total contribution of the non-overlapping sets to (16) is given, on using (21) and (25), by

$$\Sigma_1 = \Sigma_{11} + \Sigma_{12} \leqslant A_k^2 \frac{\Lambda_b k^2}{4^k} \left\{ 1 + \frac{R_b 2^k b}{n-k-b} \right\} \left\{ 1 + \frac{2^k b}{n-k} \right\}. \tag{26}$$

We now turn to the case when i' and 1 have a non-empty overlap, and bound the contribution $\Sigma_2 = \sum_{i' \neq 1, i' \cap 1 \neq \varnothing, 0 \leqslant j' \leqslant b} \mathbf{P}(I_{i',j'} \neq J_{i',j'})$ to the discrepancy (16). In this case, the situation is far more complicated than in the non-overlapping case; in particular, the 'one-directional' effect exploited systematically in the argument leading to (26) no longer holds. We bypass this difficulty, however, by bounding $\sum_{j'} \mathbf{P}(I_{i',j'} \neq J_{i',j'})$ by twice the probability that i' gets beaten by *a different number of vertices than before*. This latter quantity is far easier to control, as we shall see. Accordingly, we define $E_{j,i'}$ (resp. $F_{j,i'}$) as the event that i' is beaten by a different number of players after the coupling, on the event that 1 is beaten by more than (resp. less than) j vertices initially, and set

$$\mathbf{P}(E_{j,i'}) = \mathbf{P}(G_{j,i'}) + \mathbf{P}(H_{j,i'}) \quad \text{and} \quad \mathbf{P}(F_{j,i'}) = \mathbf{P}(K_{j,i'}) + \mathbf{P}(L_{j,i'}) \tag{27}$$

where $G_{j,i'}$ represents the event that 1 is dominated by more than j vertices, and that i' is subsequently beaten by *more players than before*, and where $H_{j,i'}$, $K_{j,i'}$ and $L_{j,i'}$ have similar interpretations. For $G_{j,i'}$, we dissect according to the size r of the overlap between 1 and i' and the number s of vertices in i' initially beating 1, and observe that there must be at least one vertex in 1 which afterwards dominates i'. This yields

$$\begin{aligned}
\sum_{i' \neq 1, i' \cap 1 \neq \varnothing} \mathbf{P}(G_{j,i'}) &\leqslant \sum_{r=1}^{k-1} \binom{k}{r} \binom{n-k}{k-r} \sum_{s=1}^{k-r} \binom{k-r}{s} \frac{1}{2^{ks}} \left(1 - \frac{1}{2^k}\right)^{k-r-s} \\
&\quad \times (k-r) \left(\frac{2^{k-1}}{2^k - 1}\right)^k \sum_{l=0}^{b} \binom{n-2k+r}{l} \frac{1}{2^{kl}} \left(1 - \frac{1}{2^k}\right)^{n-2k+r-l} \\
&\leqslant A_k^2 \binom{n}{k} \sum_{l=0}^{b} \rho_l \sum_{r=1}^{k-1} \frac{\binom{k}{r}\binom{n-k}{k-r}}{\binom{n}{k}} \frac{(k-r)^2}{4^k} \\
&\leqslant A_k^2 \frac{k^4}{(n-k)\,4^k} \Lambda_b.
\end{aligned} \tag{28}$$

The quantity $\sum_{l=0}^{b} \binom{n-2k+r}{l} (1/2^{kl})(1-2^{-k})^{n-2k+r-l}$ in the above chain of inequalities represents the probability that i' is initially beaten by fewer than b vertices from outside $i' \cup 1$, which must be the case, since otherwise both $I_{i',j'}$ and $J_{i',j'}$ are zero.

For $H_{j,i'}$, we similarly obtain

$$\begin{aligned}
&\sum_{i' \neq 1, i' \cap 1 \neq \varnothing} \mathbf{P}(H_{j,i'}) \\
&\quad \leqslant \sum_{r=1}^{k-1} \binom{k}{r} \binom{n-k}{k-r} \frac{n}{2^{2k-r}} \sum_{l=0}^{b} \binom{n-2k+r-1}{l} \left(\frac{1}{2^k} - \frac{1}{2^{2k-r}}\right)^l \left(1 - \frac{1}{2^k} + \frac{1}{2^{2k-r}}\right)^{n-2k+r-1-l} \\
&\leqslant \frac{n\Lambda_b}{4^k} A_k \sum_{r=1}^{k-1} \frac{\binom{k}{r}\binom{n-k}{k-r}}{\binom{n}{k}} 2^r \left[1 + \frac{2^r}{2^k(2^k-1)}\right]^n,
\end{aligned} \tag{29}$$

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

because at least one vertex must initially beat both 1 and i', and at most b points outside $i' \cup 1$ may initially beat i' but not 1, since otherwise i' would afterwards still be dominated by more than b vertices, forcing both $I_{i',j'}$ and $J_{i',j'}$ to be zero. Comparing (29) with (6), we see that

$$\sum_{i' \neq 1, i' \cap 1 \neq \varnothing} \mathbf{P}(H_{j,i'}) \leqslant \frac{2\Lambda_b k^2}{4^k} A_k B_{nk}, \tag{30}$$

where B_{nk} is as in Theorem 1.

We next turn to the computation of $\mathbf{P}(K_{j,i'})$ and $\mathbf{P}(L_{j,i'})$, which represent, respectively, the probabilities that i' is afterwards beaten by more or fewer vertices, both on the event that fewer than j players beat 1 initially. For $K_{j,i'}$, note that i' must initially be dominated by at most b vertices outside $1 \cup i'$ which do not dominate 1, and that at least one of the additional points chosen in the coupling to dominate 1 must also dominate i'. Hence

$$\begin{aligned}
&\sum_{i' \neq 1, i' \cap 1 \neq \varnothing} \mathbf{P}(K_{j,i'}) \\
&\quad \leqslant \sum_{r=1}^{k-1} \binom{k}{r}\binom{n-k}{k-r} \sum_{l=0}^{b} \binom{n-2k+r}{l} \left[\frac{1}{2^k}\left(1-\frac{1}{2^{k-r}}\right)\right]^l \left(1-\frac{1}{2^k}+\frac{1}{2^{2k-r}}\right)^{n-2k+r-l} \frac{j}{2^{k-r}} \\
&\quad \leqslant 2^{-k} b A_k \binom{n}{k} \sum_{r=1}^{k-1} \frac{\binom{k}{r}\binom{n-k}{k-r}}{\binom{n}{k}} 2^r \sum_{l=0}^{b} \rho_l \left[1+\frac{2^r}{2^k(2^k-1)}\right]^n \\
&\quad \leqslant \frac{2\Lambda_b k^2 b}{2^k n} A_k B_{nk},
\end{aligned} \tag{31}$$

again using the argument following (6).

For $L_{j,i'}$, dissect according to the number s of vertices outside i' which initially dominate 1, and note that some vertex of 1 must initially dominate i', whereas some vertex of i' must afterwards dominate 1. This gives

$$\begin{aligned}
\sum_{i' \neq 1, i' \cap 1 \neq \varnothing} \mathbf{P}(L_{j,i'}) &\leqslant \sum_{r=1}^{k-1} \binom{k}{r}\binom{n-k}{k-r} \sum_{s=0}^{j-1} \binom{n-2k+r}{s} \frac{1}{2^{ks}} \left(1-\frac{1}{2^k}\right)^{n-2k+r-s} \\
&\quad \times \left\{1-\left(1-\frac{1}{2^k}\right)^{k-r}\right\} \frac{(k-r)\,b}{n-k-b} \\
&\leqslant A_k \binom{n}{k} \sum_{r=1}^{k-1} \frac{\binom{k}{r}\binom{n-k}{k-r}}{\binom{n}{k}} \sum_{s=0}^{b} \rho_s \frac{k}{2^k} \frac{kb}{n-k-b} \\
&\leqslant A_k \Lambda_b \frac{k^2 b}{2^k(n-k-b)} \frac{k^2}{n-k}.
\end{aligned} \tag{32}$$

Inequalities (27), (28), (30), (31) and (32) reveal that

$$\Sigma_2 \leqslant \frac{2\Lambda_b k^2}{4^k} A_k^2 \left\{1+\frac{2^k b}{n-k-b}\right\}\left\{2B_{nk}+\frac{k^2}{n-k}\right\}. \tag{33}$$

Equation (33), together with (15), (16) and (26), now proves Theorem 2. □

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press

Example. Consider a random tournament on $n = 9.4545719 \times 10^{33}$ vertices; suppose that $k = 100$ and that we wish to approximate the joint distribution of the vector (W_0, W_1, W_2). Then

$$\lambda_0 = \binom{n}{k}\left(1-\frac{1}{2^k}\right)^{n-k} \approx 3,$$

$$\lambda_1 = \binom{n}{k}\binom{n-k}{1}\frac{1}{2^k}\left(1-\frac{1}{2^k}\right)^{n-k-1} \approx 22387,$$

and

$$\lambda_2 = \binom{n}{k}\binom{n-k}{2}\frac{1}{2^{2k}}\left(1-\frac{1}{2^k}\right)^{n-k-2} \approx 8.34832 \times 10^7.$$

In this case, (9) holds and $c_{100}(9.45\ldots \times 10^{33}) > 0$, while the error bound given by Theorem 2 may be checked to be less than 2.2×10^{-40}.

References

[1] Alon, N., Spencer, J. and Erdős, P. (1992) *The Probabilistic Method*. John Wiley, New York.

[2] Arratia, R., Barbour, A. D. and Tavaré, S. (1992) Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Probab.* **2** 519–535.

[3] Arratia, R. and Tavaré, S. (1992) The cycle structure of random permutations. *Ann. Probab.* **20** 1567–1591.

[4] Barbour, A. D., Holst, L. and Janson, S. (1992) *Poisson Approximation*. Oxford University Press.

[5] Erdős, P. (1963) On a problem of graph theory. *Math. Gaz.* **47** 220–223.

[6] Godbole, A. P., Skipper, D. E. and Sunley, R. A. (1995) The asymptotic lower bound on the classical Ramsey numbers: A closer look, in D. Aldous, P. Diaconis, J. Spencer and J. M. Steele (editors), *Discrete Probability and Algorithms*, *IMA Volumes in Mathematics and its Applications* **72** 81–94. Springer-Verlag.

[7] Moon, A. (1968) *Topics on Tournaments*. Holt, Reinhart and Winston, New York.

[8] Qian, J. (1993) A probabilistic analysis of tournaments. MS Thesis, Michigan Technological University.

https://doi.org/10.1017/S0963548396002829 Published online by Cambridge University Press