Probability in the Engineering and Informational Sciences, **33**, 2019, 500–510.
doi:10.1017/S0269964818000438

WINNER PLAYS STRUCTURE IN RANDOM KNOCKOUT TOURNAMENTS

YANG CAO AND SHELDON M. ROSS
Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90089, USA
E-mail: cao573@usc.edu; smross@usc.edu

Suppose there are n players, with player i having value $v_i > 0$, and suppose that a game between i and j is won by i with probability $v_i/(v_i + v_j)$. In the winner plays random knockout tournament, we suppose that the players are lined up in a random order; the first two play, and in each subsequent game the winner of the last game plays the next in line. Whoever wins the game involving the last player in line, is the tournament winner. We give bounds on players' tournament win probabilities and make some conjectures. We also discuss how simulation can be efficiently employed to estimate the win probabilities.

Keywords: negatively associated, random knockout tournament, winner plays

1. INTRODUCTION

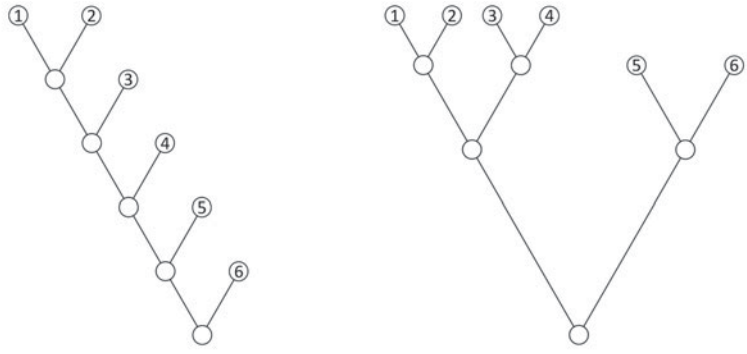
Consider a contest with players $1, 2, \ldots, n$. Suppose that each player has a value, with player i having value $v_i, i = 1, \ldots, n$. Suppose further that each game involves two players, and that a game played by i and j is won by i with probability $v_i/(v_i + v_j)$.

The type of contest we consider is a random knockout tournament, which supposes that the players are randomly assigned to positions $1, \ldots, n$. The players in positions 1 and 2 play a game in the first round, and in each of the following rounds $r = 2, \ldots, n-1$ the winner of the previous round plays with the player in position $r + 1$. The tournament ends after round $n - 1$ and the winner in that round is the winner of the tournament. We call this tournament format the winner plays structure (see Figure 1a e.g., with six players). It is a special case of what is known as a structured format, see [2,4,5], which is a tournament in which the structure is fixed and players are randomly assigned to positions of the structure. Figure 1b shows an example with six players, which in round 1 has a match between those in positions 1 and 2 and one between those in positions 3 and 4, with the winners then playing each other in round 2 and 5 playing 6 in round 2. The winners of round 2 then play a game in round 3.

The paper [2] proved that for all structured formats, a stronger player has a higher tournament win probability than a weaker player. That is, if $v_i > v_j$ then the win probability of i is greater than the win probability of j. The paper [1] studied non-structured random knockout tournaments where the number of games in each round is specified and the games are played by randomly selected players from all remaining players in that round. Bounds

© Cambridge University Press 2018 0269-9648/18 $25.00

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press



(a) Winner plays structure (b) Another structure

FIGURE **1.** Two structures with six players. (a) Winner plays structure (b) Another structure.

on players' win probabilities were given in [1]. The paper [6] provided efficient simulation algorithms for players' win probabilities in non-structured random knockout tournaments.

In Section 2, we obtain lower and upper bounds on the players' win probabilities. We also present two conjectures, which are verified in special cases. The first conjecture is that if $P(W_i)$ is the win probability of player i, then $v_i \geq v_j$ implies that $P(W_i)/v_i \geq P(W_j)/v_j$. The second conjecture is that the best player's win probability is less than or equal to what it would be if the values of the other $n-1$ players were all equal to the average of their actual values. In Section 3, we discuss how simulation can be efficiently employed to estimate the win probabilities.

2. BOUNDS ON THE PLAYERS' WIN PROBABILITIES AND CONJECTURES

Consider the winner plays random knockout tournament. Let W_i be the event that i wins the tournament, and let L_i denote the position of player i in the random permutation. For any set S of players, let $W(S)$ be the winner of a winner plays random knockout tournament whose contestants are those in S.

We first show that $P(W_i)$ is an increasing function of v_i.

PROPOSITION 1: *$P(W_i)$ is an increasing function of v_i.*

PROOF: Suppose there are two scenarios, the first with player i having strength v_i^1 and the second with i having strength $v_i^2 < v_i^1$, with the strengths of the others being the same in both scenarios. Couple these scenarios by using the same ordering permutation, and the same results for all games not involving i. For a game where i plays j in both scenarios, generate a random number U and let i win in scenario 1 if $U < v_i^1/(v_i^1 + v_j)$, and in scenario 2 if $U < v_i^2/(v_i^2 + v_j)$. Since $v_i^1/(v_i^1 + v_j) \geq v_i^2/(v_i^2 + v_j)$, a win for i in scenario 2 implies a win in scenario 1. Thus, if i wins the tournament in scenario 2 then it will win the tournament in scenario 1. Therefore, the win probability of i in scenario 1 is greater than or equal to the win probability of i in scenario 2, which yields the result. ■

The following theorem is a special case of the results in [2].

THEOREM 2: *If $v_i > v_j$ then $P(W_i) > P(W_j)$.*

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press

Theorem 2 immediately yields

COROLLARY 3: *If v_i decreases in i, then $P(W_1) + \cdots + P(W_k) \geq k/n,\ k = 1, \ldots, n$.*

To obtain bounds on $P(W_i)$, we need two lemmas.

LEMMA 4: *For a given set of players S, let $U(S)$ be equally likely to be any of the players in S. Then $v_{W(S)}$ is stochastically larger than $v_{U(S)}$.*

PROOF: Renumber the players in set S to be $1, 2, \ldots, |S|$, with values so that $v_1 \geq v_2 \geq \ldots \geq v_{|S|}$. Then $P(U(S) = j) = 1/|S|,\ j = 1, 2, \ldots, |S|$. Using the notation $X \geq_{\text{st}} Y$ to mean that X is stochastically larger than Y, it follows from Corollary 3 that

$$P(W(S) \leq k) \geq k/|S| = P(U(S) \leq k), \quad k = 1, 2, \ldots, |S|,$$

showing that $W(S) \leq_{\text{st}} U(S)$. In addition, because v_i is decreasing in i, it follows that $v_{W(S)} \geq_{\text{st}} v_{U(S)}$. ■

LEMMA 5: *For any set of players S and $i, i \notin S$, we have*

$$E\left[\frac{v_i}{v_i + v_{W(S)}}\right] \leq \frac{1}{|S|} \sum_{j \in S} \frac{v_i}{v_i + v_j}$$

PROOF: Using the previous lemma and the fact that $v_i/(v_i + x)$ is decreasing in x, it follows that $E[v_i/(v_i + v_{W(S)})] \leq E[v_i/(v_i + v_{U(S)})]$ which proves the lemma. ■

PROPOSITION 6: *For fixed i, let $S^i = \{1, \ldots, i-1, i+1, \ldots, n\}$, and let S^i_j be the set of all $\binom{n-1}{j}$ subsets of S^i of size j. Also, for $S \subset S^i$, let $v_{S^+} = \max\{v_k : k \in S\}$. Then*

$$P(W_i) \leq \frac{2}{n} \prod_{k \neq i} \frac{v_i}{v_i + v_k} + \frac{1}{n} \sum_{j=3}^{n} \frac{1}{\binom{n-1}{j-1}} \sum_{S \subset S^i_{j-1}} \frac{1}{|S|} \sum_{r \in S} \frac{v_i}{v_i + v_r} \prod_{k \in S^i - S} \frac{v_i}{v_i + v_k}$$

$$P(W_i) \geq \frac{2}{n} \prod_{k \neq i} \frac{v_i}{v_i + v_k} + \frac{1}{n} \sum_{j=3}^{n} \frac{1}{\binom{n-1}{j-1}} \sum_{S \subset S^i_{j-1}} \frac{v_i}{v_i + v_{S^+}} \prod_{k \in S^i - S} \frac{v_i}{v_i + v_k}$$

PROOF: Condition on L_i, the position of player i,

$$\begin{aligned} P(W_i) &= \frac{1}{n} \sum_{j=1}^{n} P(W_i | L_i = j) \\ &= \frac{2}{n} \prod_{k \neq i} \frac{v_i}{v_i + v_k} + \frac{1}{n} \sum_{j=3}^{n} \frac{1}{\binom{n-1}{j-1}} \sum_{S \subset S^i_{j-1}} E\left[\frac{v_i}{v_i + v_{W(S)}}\right] \prod_{k \in S^i - S} \frac{v_i}{v_i + v_k} \end{aligned}$$

where the final equality follows from conditioning on the set of players preceding i in the random permutation.

The upper bound now follows from Lemma 5, and the lower bound follows because $v_{W(S)} \leq v_{S^+}$. ■

The amount of computation in the preceding bounds grows exponentially as the number of players increases. The following gives a computationally simpler upper bound.

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press

PROPOSITION 7: *Suppose v_i decreases in i. With $a_i, b_{i,j}$ defined as*

$$a_i = \frac{1}{n-1}\sum_{k\neq i}\frac{v_i}{v_i+v_k},$$

$$b_{i,j} = \sum_{r=1}^{i-1}\frac{\binom{r-1}{j-2}}{\binom{n-1}{j-1}}\frac{v_i}{v_i+v_r} + \sum_{r=i+1}^{n}\frac{\binom{r-2}{j-2}}{\binom{n-1}{j-1}}\frac{v_i}{v_i+v_r},$$

we have that

$$P(W_i) \leq \frac{2}{n}\prod_{k\neq i}\frac{v_i}{v_i+v_k} + \frac{1}{n}\frac{v_i}{v_i+v_{i^*}} + \frac{1}{n}\sum_{j=3}^{n-1} b_{i,j}a_i^{n-j}$$

where $i^* = n - I\{i=n\}$ and we are using the convention that $\binom{s}{t} = 0$ if $t > s$.

PROOF: Fix i, and let I_k be the k^{th} one of players $1,\ldots,i-1,i+1,\ldots,n$ to appear in the random permutation. Then, with $\mathbf{I} = (I_1,\ldots,I_{n-1})$, we have

$$\begin{aligned}
P(W_i|\mathbf{I}) &= \frac{1}{n}\sum_{j=1}^{n} P(W_i|\mathbf{I}, L_i = j)\\
&= \frac{2}{n}\prod_{k\neq i}\frac{v_i}{v_i+v_k} + \frac{1}{n}\sum_{j=3}^{n-1} E\left[\frac{v_i}{v_i+v_{W(I_1,\ldots,I_{j-1})}}\prod_{k=j}^{n-1}\frac{v_i}{v_i+v_{I_k}}|\mathbf{I}\right]\\
&\quad + \frac{1}{n}E\left[\frac{v_i}{v_i+v_{W(I_1,\ldots,I_{n-1})}}|\mathbf{I}\right]
\end{aligned}$$

Hence,

$$\begin{aligned}
P(W_i) &= \frac{2}{n}\prod_{k\neq i}\frac{v_i}{v_i+v_k} + \frac{1}{n}\sum_{j=3}^{n-1} E\left[\frac{v_i}{v_i+v_{W(I_1,\ldots,I_{j-1})}}\prod_{k=j}^{n-1}\frac{v_i}{v_i+v_{I_k}}\right]\\
&\quad + \frac{1}{n}E\left[\frac{v_i}{v_i+v_{W(I_1,\ldots,I_{n-1})}}\right]\\
&\leq \frac{2}{n}\prod_{k\neq i}\frac{v_i}{v_i+v_k} + \frac{1}{n}\sum_{j=3}^{n-1} E\left[\frac{v_i}{v_i+v_{\max(I_1,\ldots,I_{j-1})}}\prod_{k=j}^{n-1}\frac{v_i}{v_i+v_{I_k}}\right]\\
&\quad + \frac{1}{n}E\left[\frac{v_i}{v_i+v_{\max(I_1,\ldots,I_{n-1})}}\right] \qquad (1)
\end{aligned}$$

where the inequality follows because v_i is monotone and $W(I_1,\ldots,I_{j-1}) \leq \max(I_1,\ldots,I_{j-1})$. Now, it has been shown in [3] that $I_1,\ldots,I_{n-1}$ are negatively associated, which means that for increasing functions h and k

$$E[h(I_1,\ldots,I_r)k(I_{r+1},\ldots,I_{n-1})] \leq E[h(I_1,\ldots,I_r)]E[k(I_{r+1},\ldots,I_{n-1})]$$

Because $v_i/(v_i+v_{\max(I_1,\ldots,I_{j-1})})$ is an increasing function of $I_1,\ldots,I_{j-1}$ and $\prod_{k=j}^{n-1} v_i/(v_i+v_{I_k})$ is an increasing function of $I_j,\ldots,I_{n-1}$, it follows from this negatively associated

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press

result that

$$E\left[\frac{v_i}{v_i+v_{\max(I_1,\ldots,I_{j-1})}}\prod_{k=j}^{n-1}\frac{v_i}{v_i+v_{I_k}}\right]\le E\left[\frac{v_i}{v_i+v_{\max(I_1,\ldots,I_{j-1})}}\right]E\left[\prod_{k=j}^{n-1}\frac{v_i}{v_i+v_{I_k}}\right] \quad (2)$$

Continually employing the negative association result yields that

$$\begin{aligned}E\left[\prod_{k=j}^{n-1}\frac{v_i}{v_i+v_{I_k}}\right]&\le E\left[\frac{v_i}{v_i+v_{I_j}}\right]E\left[\prod_{k=j+1}^{n-1}\frac{v_i}{v_i+v_{I_k}}\right]\\&\le E\left[\frac{v_i}{v_i+v_{I_j}}\right]E\left[\frac{v_i}{v_i+v_{I_{j+1}}}\right]E\left[\prod_{k=j+2}^{n-1}\frac{v_i}{v_i+v_{I_k}}\right]\\&\le\cdots\le\prod_{k=j}^{n-1}E\left[\frac{v_i}{v_i+v_{I_k}}\right].\end{aligned}$$

Thus, from equation (2),

$$E\left[\frac{v_i}{v_i+v_{\max(I_1,\ldots,I_{j-1})}}\prod_{k=j}^{n-1}\frac{v_i}{v_i+v_{I_k}}\right]\le E\left[\frac{v_i}{v_i+v_{\max(I_1,\ldots,I_{j-1})}}\right]\left(\frac{1}{n-1}\sum_{r\ne i}\frac{v_i}{v_i+v_r}\right)^{n-j}$$

Consequently, from equation (1) and with $i^*=n-I\{i=n\}$,

$$\begin{aligned}P(W_i)\le{}&\frac{2}{n}\prod_{k\ne i}\frac{v_i}{v_i+v_k}+\frac{1}{n}\sum_{j=3}^{n-1}E\left[\frac{v_i}{v_i+v_{\max(I_1,\ldots,I_{j-1})}}\right]\left(\frac{1}{n-1}\sum_{r\ne i}\frac{v_i}{v_i+v_r}\right)^{n-j}\\&+\frac{1}{n}\frac{v_i}{v_i+v_{i^*}}\end{aligned}$$

Now, using the convention that $\binom{s}{t}=0$ if $t>s$, we have

$$P(\max(I_1,\ldots,I_{j-1})=r)=\begin{cases}\binom{r-1}{j-2}/\binom{n-1}{j-1}, & \text{if } r<i\\ \binom{r-2}{j-2}/\binom{n-1}{j-1}, & \text{if } r>i\end{cases}$$

Hence, with

$$b_{i,j}\equiv E\left[\frac{v_i}{v_i+v_{\max(I_1,\ldots,I_{j-1})}}\right]=\sum_{r=1}^{i-1}\frac{\binom{r-1}{j-2}}{\binom{n-1}{j-1}}\frac{v_i}{v_i+v_r}+\sum_{r=i+1}^{n}\frac{\binom{r-2}{j-2}}{\binom{n-1}{j-1}}\frac{v_i}{v_i+v_r}$$

and

$$a_i\equiv\frac{1}{n-1}\sum_{k\ne i}\frac{v_i}{v_i+v_k}$$

we see that

$$P(W_i)\le\frac{2}{n}\prod_{k\ne i}\frac{v_i}{v_i+v_k}+\frac{1}{n}\frac{v_i}{v_i+v_{i^*}}+\frac{1}{n}\sum_{j=3}^{n-1}b_{i,j}a_i^{n-j}$$

■

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press

Remarks:

1. Numerical results that illustrate the bounds are given in Table 1 in Section 3.2.
2. It is unknown to us whether $P(W_i)$ is a decreasing function of v_j, $j \neq i$. Intuitively the statement seems reasonable, as increasing v_j makes player i less likely to beat player j in a game. On the contrary, we notice that by increasing v_j, player j may become more likely to beat other stronger potential opponents of the player i, which may increase $P(W_i)$. The paper [1] gave a counterexample to $P(W_i)$ being a decreasing function of v_j, $j \neq i$, for a structure with four players in which there are two games in the first round, with the two winners then playing each other in the second round. If the players' strengths are $v_1 = 2, v_2 = 1, v_3 = 1, v_4 = 1/100$, then $P(W_1) \approx 0.518858$. If changing v_4 to $1/99$, then $P(W_1) \approx 0.518861$. However, we have not found any counterexamples for the winner plays structure.

We now propose two conjectures regarding the players' win probabilities and provide evidence supporting these conjectures.

CONJECTURE 1: *If $v_i \geq v_j$ then $P(W_i)/v_i \geq P(W_j)/v_j$.*

The conjecture can be proven when $n \leq 6$ by a brute force approach that uses Mathematica. We now prove the result in a special case.

PROPOSITION 8: *If $v_1 \geq v_2 = v_3 = \ldots = v_n = v$, then $P(W_i)/v_i$ decreases in i.*

PROOF: Since $v_1 \geq v_2 = v_3 = \ldots = v_n = v$, the result is equivalent to

$$\frac{P(W_1)}{v_1} \geq \frac{P(W_2)}{v_2} = \frac{1 - P(W_1)}{(n-1)v},$$

which is then equivalent to

$$P(W_1) \geq \frac{v_1}{v_1 + (n-1)v}.$$

To prove this, we condition on the position of player 1 and obtain

$$P(W_1) = \frac{2}{n}\left(\frac{v_1}{v_1 + v}\right)^{n-1} + \frac{1}{n}\sum_{j=3}^{n}\left(\frac{v_1}{v_1 + v}\right)^{n-j+1}.$$

Thus,

$$\begin{aligned} P(W_1) - \frac{v_1}{v_1 + (n-1)v} &= \frac{2}{n}\left(\frac{v_1}{v_1 + v}\right)^{n-1} + \frac{1}{n}\sum_{j=3}^{n}\left(\frac{v_1}{v_1 + v}\right)^{n-j+1} - \frac{v_1}{v_1 + (n-1)v} \\ &= \frac{2}{n}\left(\frac{v_1}{v_1 + v}\right)^{n-1} + \frac{1}{n}\frac{v_1}{v}\left(1 - \left(\frac{v_1}{v_1 + v}\right)^{n-2}\right) - \frac{v_1}{v_1 + (n-1)v} \end{aligned}$$

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press

$$
\begin{aligned}
&= \frac{v_1}{nv} + \frac{v_1}{n}\left(\frac{v_1}{v_1+v}\right)^{n-2}\left(\frac{2}{v_1+v} - \frac{1}{v}\right) - \frac{v_1}{v_1+(n-1)v} \\
&= \frac{v_1}{nv} - \frac{v_1}{v_1+(n-1)v} + \frac{v-v_1}{nv}\left(\frac{v_1}{v_1+v}\right)^{n-1} \\
&= \frac{v_1-v}{nv}\frac{v_1}{v_1+(n-1)v} + \frac{v-v_1}{nv}\left(\frac{v_1}{v_1+v}\right)^{n-1} \\
&= \frac{v_1-v}{nv}\left[\frac{v_1}{v_1+(n-1)v} - \left(\frac{v_1}{v_1+v}\right)^{n-1}\right] \\
&= \frac{v_1-v}{nv}\frac{v_1}{(v_1+v)^{n-1}(v_1+(n-1)v)}[(v_1+v)^{n-1} - (v_1+(n-1)v)v_1^{n-2}] \\
&\geq 0
\end{aligned}
$$

■

If Conjecture 1 is true, it implies that if player i is the best player, that is, the player with the largest value, then $P(W_i) \geq v_i / \sum_{k=1}^{n} v_k$. (Because if $P(W_i) < v_i / \sum_{k=1}^{n} v_k$ then the conjecture implies that for all $j \neq i$, $P(W_j) \leq P(W_i)v_j/v_i < v_j / \sum_{k=1}^{n} v_k$ and thus $\sum_{j=1}^{n} P(W_j) < \sum_{j=1}^{n} v_j / \sum_{k=1}^{n} v_k = 1$.)

Second, we conjecture an upper bound for the best player's win probability. The conjecture follows from the intuition that the best player, say player i, would rather have all its opponents having the same strength $\bar{v}(i) = \sum_{j\neq i} v_j/(n-1)$.

Conjecture 2: *Under the winner plays structure, if* $v_i = \max_j v_j$, *then*

$$
P(W_i) \leq \frac{2}{n}\left(\frac{v_i}{v_i+\bar{v}(i)}\right)^{n-1} + \frac{1}{n}\sum_{j=3}^{n}\left(\frac{v_i}{v_i+\bar{v}(i)}\right)^{n-j+1}.
$$

When $n = 3$, Conjecture 2 is algebraically proven in the appendix. Numerical evidence for $n \leq 5$ also supports the conjecture. Conjecture 2 may only hold for the best player and need not hold for the other players. For a counterexample, suppose there are 3 players with respective strength $4, 1, 1$. Then the win probability of player 2 is $23/150 = 0.153$. On the other hand, if each of players 1, 3 has the same strength $(4+1)/2 = 5/2$, then the win probability of player 2 becomes $22/147 = 0.150$, which is smaller than its original win probability.

3. EFFICIENT USE OF SIMULATION TO ESTIMATE WIN PROBABILITIES

3.1. Recursive equations for computing conditional win probabilities

Consider a winner plays structure with players $1, 2, \ldots, n$. Let $\mathbf{I} = (I_1, I_2, \ldots, I_n)$ denote the random permutation of the players. Then in round 1 player I_1 plays with I_2, and in each of round $r = 2, \ldots, n-1$ the winner of the previous round plays with player I_{r+1}. The conditional win probabilities given the permutation $\mathbf{I}$ can be computed recursively as follows. Let E_j^k denote the event that the player in position j, that is, player I_j, advances to round k, $k = 1, 2, \ldots, n$, where advancing to round k means that the player has not lost in the first $k-1$ rounds, and advancing to round n means that the player has won the tournament. Since all players advance to round 1, and the player in position j automatically

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press

advances to round k for $1 < k < j \le n$, we have

$$P(E_j^1|\mathbf{I}) = 1, \quad j = 1, 2, \ldots, n, \tag{3}$$

$$P(E_j^k|\mathbf{I}) = 1, \quad 1 < k < j \le n. \tag{4}$$

Conditioning yields,

$$P(E_k^k|\mathbf{I}) = \sum_{j=1}^{k-1} P(E_j^{k-1}|\mathbf{I}) \frac{v_{I_k}}{v_{I_j} + v_{I_k}}, \quad k = 2, 3, \ldots, n. \tag{5}$$

$$P(E_j^k|\mathbf{I}) = P(E_j^{k-1}|\mathbf{I}) \frac{v_{I_j}}{v_{I_j} + v_{I_k}}, \quad 1 \le j < k \le n, \tag{6}$$

Thus, $P(E_j^k|\mathbf{I}), j = 1, \ldots, n, k = 1, \ldots, n$, can be computed recursively using equations (3), (4), (5) and (6). The conditional win probability of player i given the permutation $\mathbf{I}$ is $P(W_i|\mathbf{I}) = P(E_{L_i}^n|\mathbf{I})$, where L_i is the position assigned to player i (that is, $L_i = j$ if $I_j = i$).

3.2. Simulation of the win probabilities

The unconditional win probability of player i is $P(W_i) = E[P(W_i|\mathbf{I})]$, which is the average of conditional win probabilities over $n!$ permutations of the players. For each permutation, the time complexity for computing the conditional win probability is $O(n^2)$, so the total time complexity for computing $P(W_i)$ is $O(n!n^2)$. We now introduce two conditional expectation simulation algorithms for estimating the win probabilities, and then show how to further reduce the variance of these estimators by using control variates.

The first algorithm is as follows. At each simulation, we generate a random permutation $(i_1, i_2, \ldots, i_n)$ of the players. Then we apply the recursive equations (3), (4), (5) and (6) to compute the conditional win probabilities given that permutation. For each player i the conditional win probability $P(W_i|\mathbf{I})$ is an unbiased estimator for the unconditional win probability $P(W_i)$.

The preceding conditional expectation algorithm estimates the win probabilities for all players. However, if we want to estimate the win probability for a specific player, we can use the following algorithm which yields an estimator with smaller variance. The algorithm is the same for all players, so for notational convenience and without loss of generality, we estimate the win probability of player n. We generate the permutation of players $1, 2, \ldots, n$ in two steps. First, we generate $\mathbf{I}^*$, a random permutation of $1, 2, \ldots, n-1$ representing their relative ordering. Then we randomly insert player n to one of the n possible positions.

The conditional win probability $P(W_n|\mathbf{I}^*)$ is computed as follows. For a relative ordering $\mathbf{I}^*$, consider a tournament π^{n-1} whose contestants are players $1, 2, \ldots, n-1$ and the permutation is $\mathbf{I}^* = (I_1, I_2, \ldots, I_{n-1})$. Let E_j^k denote the event that the player in position j, that is, player I_j, advances to round k, $j = 1, 2, \ldots, n-1, k = 1, 2, \ldots, n-1$. The conditional probabilities $P(E_j^k|\mathbf{I}^*)$ can be computed by applying the recursive equations (3), (4), (5) and (6) in this tournament.

Now, we randomly insert player n to the relative ordering $\mathbf{I}^* = (I_1, \ldots, I_{n-1})$. Suppose L_n denotes the resulting position of player n. Then it plays with one of players $I_1, \ldots, I_{L_n-1}$ in round $L_n - 1$, and then plays with player I_r in round $r, r = L_n, L_n + 1, \ldots, n-1$. Let π^n denote the new tournament after player n is inserted. Then π^n is equivalent to a tournament with players $1, 2, \ldots, n$ and permutation $I_1, I_2, \ldots, I_{L_n-1}, n, I_{L_n}, \ldots, I_{n-1}$.

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press

The conditional win probability of player n given $\mathbf{I}^*$ is now obtained by conditioning on L_n and on the first opponent of player n. That is,

$$P(W_n|\mathbf{I}^*) = \sum_{k=1}^{n} P(W_n|\mathbf{I}^*, L_n = k)P(L_n = k)$$
$$= \frac{2}{n}\prod_{j=1}^{n-1}\frac{v_n}{v_n + v_j} + \frac{1}{n}\sum_{k=3}^{n}\sum_{j=1}^{k-1} P(E_j^{k-1}|\mathbf{I}^*)\frac{v_n}{v_n + v_{I_j}}\prod_{t=k}^{n-1}\frac{v_n}{v_n + v_{I_t}}. \quad (7)$$

With this equation, algorithm II works as follows. At each simulation, we generate $\mathbf{i}^* = (i_1, i_2, \ldots, i_{n-1})$, a random permutation of $1, 2, \ldots, n-1$ representing their relative ordering. We apply the recursive equations in this tournament to compute the conditional probabilities $P(E_j^k|\mathbf{I}^* = \mathbf{i}^*)$, $j = 1, 2, \ldots, n-1, k = 1, 2, \ldots, n-1$. Then we apply equation (7) and compute $P(W_n|\mathbf{I}^* = \mathbf{i}^*)$, the conditional win probability of player n given that the relative ordering of players $1, 2, \ldots, n-1$ is $i_1, i_2, \ldots, i_{n-1}$, as the estimator for $P(W_n)$. Moreover, since player n is an arbitrary player, we can apply this algorithm to estimate the win probability of any player.

Note that the estimator of $P(W_n)$ in the first algorithm is $P(W_n|\mathbf{I}^*, L_n)$ where L_n is the position of player n, and the estimator in the second algorithm is $P(W_n|\mathbf{I}^*)$. Since we condition on more information in the first algorithm than the second, the first algorithm will have a larger variance. (More formally, this is shown by noting that $P(W_n|\mathbf{I}^*) = E[P(W_n|\mathbf{I}^*, L_n)|\mathbf{I}^*]$, which shows that the second estimator is a conditional expectation of the first and thus has a smaller variance.)

Finally, we use control variates to further reduce the variance of the estimators. Recall that L_j is the position assigned to player j, $j = 1, 2, \ldots, n$. For the first algorithm, it is intuitive that L_n is an effective control variate for $P(W_n|\mathbf{I}^*, L_n)$, since they appear to be strongly positively correlated. For the second algorithm, we consider $Y = \sum_{j=1}^{n-1} v_j L_j$. If the strong opponents of player n are assigned to positions with small indices and the weak opponents are assigned to positions with large indices, then Y will be small and $P(W_n|\mathbf{I}^*)$ will intuitively be large. Similarly, if the strong opponents of player n are assigned to positions with large indices and the weak opponents are assigned to positions with small indices, then Y will be large and $P(W_n|\mathbf{I}^*)$ will intuitively be small. Thus, Y appears strongly negatively correlated with the estimator $P(W_n|\mathbf{I}^*)$, and is likely to be an effective control variate.

We compare the performance of the simulation algorithms with an experiment. Specifically, we consider the case where $v_i = n + 1 - i, i = 1, 2, \ldots, n$, and estimate the win probability of player 1. The following table shows the simulation results as well as the upper bound (denoted by UB1) and lower bound (denoted by LB1) defined in Proposition 6 and the upper bound (denoted by UB2) defined in Proposition 7.

The simulation results show that both algorithms perform much better than simulation without any variance reduction, whose variance is $P(W_1)(1 - P(W_1))$. In addition, the use of the control variates further reduces the estimation variance in both algorithms.

We also observe from the data that algorithm II significantly reduces the estimation variance compared with algorithm I. Therefore, it seems that when estimating the win probabilities for all players, applying algorithm II for every player may also be a better choice than applying algorithm I once. Moreover, we can generate a random permutation of players $1, 2, \ldots, n$, and when estimating each player's win probability using algorithm II, we can use the same permutation to determine the relative ordering of the other players.

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press

Table 1. Simulation results

Number of players	Algorithm	$P(W_1)$	Variance
$n = 5$	Algorithm I	0.3542	2.671×10^{-2}
	Algorithm I with control variate	0.3543	3.714×10^{-3}
	Algorithm II	0.3542	4.170×10^{-4}
	Algorithm II with control variate	0.3542	7.010×10^{-5}
	LB1	0.3266	
	UB1	0.3734	
	UB2	0.4296	
$n = 10$	Algorithm I	0.1922	3.681×10^{-2}
	Algorithm I with control variate	0.1930	7.928×10^{-3}
	Algorithm II	0.1925	6.987×10^{-4}
	Algorithm II with control variate	0.1925	1.002×10^{-4}
	LB1	0.1668	
	UB1	0.2116	
	UB2	0.2757	
$n = 50$	Algorithm I	0.03991	1.358×10^{-2}
	Algorithm I with control variate	0.03974	9.754×10^{-3}
	Algorithm II	0.03974	7.598×10^{-5}
	Algorithm II with control variate	0.03975	5.279×10^{-5}
	UB2	0.06360	
$n = 100$	Algorithm I	0.01982	7.098×10^{-3}
	Algorithm I with control variate	0.02000	6.108×10^{-3}
	Algorithm II	0.01993	2.088×10^{-5}
	Algorithm II with control variate	0.01993	1.748×10^{-5}
	UB2	0.03220	

Regarding the upper and lower bounds, we observe that Proposition 6 gives a tighter upper bound than Proposition 7. However, when the number of players is large, only the upper bound from Proposition 7 is computationally feasible.

Acknowledgments

This material is based upon work supported by, or in part by the National Science Foundation under contract/grant number CMMI1662442.

References

1. Adler, I., Cao, Y., Karp, R., Pekoz, E., & Ross, S.M. (2016). Random knockout tournaments. *arXiv preprint arXiv:1612.04448*, Accepted for publication on Operations Research.
2. Chung, F.R.K. & Hwang, F.K. (1978). Do stronger players win more knockout tournaments? *Journal of the American Statistical Association* 73(363): 593–596.
3. Joag-Dev, K. & Proschan, F. (1983). Negative association of random variables with applications. *The Annals of Statistics* 11(1): 286–295.
4. Marchand, É. (2002). On the comparison between standard and random knockout tournaments. *Journal of the Royal Statistical Society: Series D (The Statistician)* 51(2): 169–178.
5. Maurer, W. (1975). On most effective tournament plans with fewer games than competitors. *The Annals of Statistics* 3(3): 717–727.
6. Ross, S. & Ghamami, S. (2008). Efficient simulation of a random knockout tournament. *Journal Industrial and Systems Engineering* 2: 88–96.

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press

APPENDIX

Proof of Conjecture 2 when $n = 3$: Suppose $n = 3$ and, without loss of generality, that $v_1 = \max_j v_j$. We need to prove that

$$P(W_1) \le \frac{2}{3}\left(\frac{v_1}{v_1 + (v_2+v_3)/2}\right)^2 + \frac{1}{3}\frac{v_1}{v_1 + (v_2+v_3)/2}$$

Using that

$$P(W_1) = \frac{2}{3}\frac{v_1^2}{(v_1+v_2)(v_1+v_3)} + \frac{1}{3}\left[\frac{v_2v_1}{(v_2+v_3)(v_1+v_2)} + \frac{v_3v_1}{(v_2+v_3)(v_1+v_3)}\right]$$

the result follows as

$$\begin{aligned}
&\frac{2}{3}\left(\frac{v_1}{v_1 + (v_2+v_3)/2}\right)^2 + \frac{1}{3}\frac{v_1}{v_1 + (v_2+v_3)/2} - P(W_1) \\
&\quad = \frac{v_1^2(2v_1 - v_2 - v_3)(v_2 - v_3)^2}{3(v_1+v_2)(v_1+v_3)(v_2+v_3)(2v_1+v_2+v_3)^2} \\
&\quad \ge 0
\end{aligned}$$

■

https://doi.org/10.1017/S0269964818000438 Published online by Cambridge University Press