Competitive intensity and quality maximizing seedings in knock-out tournaments

Dmitry Dagaev & Alex Suzdaltsev

Journal of Combinatorial Optimization

ISSN 1382-6905

J Comb Optim
DOI 10.1007/s10878-017-0164-7

Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media, LLC. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Author's personal copy

J Comb Optim
DOI 10.1007/s10878-017-0164-7

Competitive intensity and quality maximizing seedings in knock-out tournaments

Dmitry Dagaev[1] · **Alex Suzdaltsev**[2]

© Springer Science+Business Media, LLC 2017

Abstract Before a knock-out tournament starts, the participants are assigned to positions in the tournament bracket through a process known as seeding. There are many ways to seed a tournament. In this paper, we solve a discrete optimization problem of finding a seeding that maximizes spectator interest in a tournament when spectators are interested in matches with high competitive intensity (i.e., matches that involve teams comparable in strength) and high quality (i.e., matches that involve strong teams). We find a solution to the problem under two assumptions: the objective function is linear in quality and competitive intensity and a stronger team beats a weaker one with sufficiently high probability. Depending on parameters, only two special classes of seedings can be optimal. While one of the classes includes a seeding that is often used in practice, the seedings in the other class are very different. When we relax the assumption of linearity, we find that these classes of seedings are in fact optimal in a sizable number of cases. In contrast to existing literature on optimal seedings, our results are valid for an arbitrarily large number of participants in a tournament.

Keywords Knock-out tournament · Seeding · Combinatorial optimization · Operations research in sports

Dmitry Dagaev is supported by The National Research University—Higher School of Economics Academic Fund Program in 2014/2015 (research grant No 14-01-0007).

✉ Dmitry Dagaev
ddagaev@gmail.com

Alex Suzdaltsev
asuzdaltsev@gmail.com

1 National Research University Higher School of Economics, 20 Myasnitskaya street, Moscow, Russia 101000

2 Stanford Graduate School of Business, 655 Knight Way, Stanford, CA 94305, USA

Published online: 08 August 2017

Author's personal copy

1 Introduction

Knock-out tournament (also known as elimination tournament) is one of the most frequently used sports tournament formats. After each game the winner advances to the next round and the loser is out. Contrary to round-robin competitions (every team plays with every other), in a knock-out tournament teams play against only a very limited number of competitors. Usually, before the start of the competition a draw is held to fill in the tournament bracket. The role of the draw is crucial because even for a comparatively strong team an unlucky draw may lead to early elimination. In order to protect the best teams from meeting each other at the early stages, favorites—they are called seeded teams—are drawn at different parts of the tournament bracket. Such traditional design has its reasons because a loss of a strong and well-known team in the first rounds may reduce the spectator interest in the whole tournament.

We consider a standard knock-out tournament with 2^n teams, where $n \geq 1$ is the total number of rounds. Let the set of teams in the tournament be $\{1, 2, \dots, 2^n\}$. We assume that the teams are strictly ranked by their strength (rating). Let $s_i \in \mathbb{R}$ be the strength of the team i, $i = 1, 2, \dots, 2^n$, with $s_1 > s_2 > \dots > s_{2^n}$.

The knock-out tournament can be represented by a binary tree of height n with 2^n terminal nodes. Before the start of the tournament, the terminal nodes of the tree are labeled according to *a seeding*—a one-to-one function from the set of teams to the set of terminal nodes. Any two labeled nodes linked to another node stand for a match; the parent of the two nodes is then labeled with the winner of the match. The labeling proceeds until the root of the tree is labeled with the tournament's champion.

It is convenient to represent a seeding by a 2^n-tuple which is a permutation of the tuple $(1, 2, \dots, 2^n)$. Denote the set of permutations of $(1, 2, \dots, 2^n)$ by X_n, $n \geq 1$.

Empirical research shows (see Forrest and Simmons 2002) that the spectator demand for watching a football game depends on the two characteristics of a match—its *competitive intensity* and *quality*. The term "competitive intensity" (hereafter CI) refers to the degree of balancedness of a match. A match between two equally strong teams is said to have high CI, while a match between a strong team and a weak team is said to have low CI. The term "quality" accounts for the overall strength of teams playing in a match. In our model, $(s_i + s_j)$ is the quality while $-|s_i - s_j|$ is the competitive intensity of the match involving teams i and j.

We consider the problem of finding a seeding that maximizes an increasing function of both CI and quality, aggregated over all the matches of the tournament. Traditional seeding (depicted in Fig. 1 for a three-round tournament) does ensure high CI in late stages of a tournament at a price of having relatively low CI in the early stages. On the other hand, one can imagine seedings that let strong teams play with strong teams and weak teams play with weak teams already in the first round. In contrast to the traditional seeding, such seedings generate high CI in the beginning of the tournament while low CI in the end.

Our results are as follows. If the spectator interest function is linear in both CI and quality and it is true that in any match, a stronger team wins with sufficiently high probability, then, depending on the parameters, only two relatively small classes of seedings can be optimal. We call these classes *close* seedings and *distant* seedings. We postpone the formal definitions to Sect. 3. Informally, close seedings are such that,

Author's personal copy

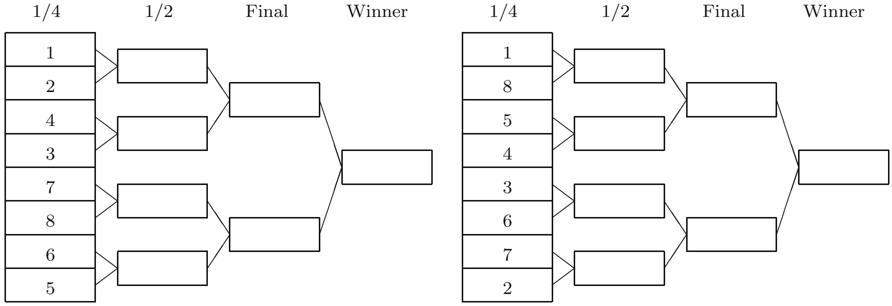

Fig. 1 The examples of close (*left*) and distant (*right*) seedings

if a stronger team always beats a weaker one, in every round of the tournament the strongest team out of the remaining participants faces the second strongest, the third strongest faces the fourth strongest, and so on. By contrast, distant seedings are such that at any stage of the tournament each of the top half participants meets one of the bottom half ones. The examples of close and distant seedings in a tournament with 8 participants are provided on Fig. 1.

We derive a simple condition which governs which seedings—close or distant—will be optimal in the linear case. It turns out that distant seedings are optimal whenever there is a sufficiently strong preference for quality (as reflected by its relative weight in the objective function) and if, other things being equal, the spectator interest increases at a sufficiently high rate as the tournament proceeds. Otherwise, close seedings are optimal. Our results hold for all values of teams' strength.

We then drop the linearity assumption and show that close and distant seedings remain optimal in a large number of cases. We provide sufficient conditions on the functional form of the objective function under which distant seedings are optimal as well as the sufficient conditions under which close seedings are optimal. In the case of close seedings, the sufficient conditions turn out to be "almost" necessary.

One of the distant seedings (the one depicted in Fig. 1) is the very seeding that we have referred to as "traditional" above; this seeding is widely used in practice and has been subject to much analysis in the literature (see, for example, Hwang 1982 or Schwenk 2000). By contrast, close seedings are, to the best of our knowledge, never employed by tournament organizers (although a "close" pairing may arise as an outcome of a draw if no seeding is used; this was the case in 2014–2015 English Football Association Challenge Cup where several Premier League clubs met already in the third round of the competition). Our results, then, suggest that for a certain set of parameters the existing practices may be far from optimal.

However, even though close seedings sometimes turn out to be optimal in our model, there are certain reasons for avoiding them that are beyond our framework. One of those reasons stems from the fact that seeding is usually determined on a basis of teams' historical rankings. If, under a certain seeding, a highly ranked team is sure to face another strong competitor already in the first round, it may have a perverse incentive to manipulate its rankings downwards by exerting less effort or even deliberately losing matches in previous competitions. Therefore, our analysis suggests that the provision

Author's personal copy

of the right incentives by means of the traditional seeding may have a cost in terms of the tournament's overall competitive intensity and quality. On the other hand, for the complementing set of parameters our model predicts that the traditional method of seeding is, in fact, optimal from the competitive intensity and quality point of view.

2 Literature review

A complete and up-to-date general survey of the operations research literature on the design of tournaments and sporting rules may be found in Wright (2014).

A particular dimension of tournament design is the problem of finding "good" seedings in knock-out tournaments. One approach is axiomatic—a seeding is considered "good" when it satisfies certain criteria. The most popular criterion here is "monotonicity": it is deemed desirable that the probability of winning the tournament is increasing in player rank since non-monotonicity may create perverse incentives for teams. Hwang (1982) shows that monotonicity may not hold under the traditional method of seeding while Baumann et al. (2010) find some statistical evidence for monotonicity violation using data from the NCAA March Madness basketball tournament. Hwang (1982) also proves that reseeding after each round restores monotonicity. Schwenk (2000) suggests another remedy for the problem. He shows that a certain randomization procedure (called "cohort randomized seeding") satisfies an axiom closely related to monotonicity ("sincerity rewarded", i.e. lack of the perverse incentives itself) as well as two other axioms.

Another approach attempts to find seedings that optimize certain quantities. A popular objective function is the probability of the highest-ranked player winning in the tournament (which is also called *predictive power*). Horen and Riezman (1985) show that in a 4-player knock-out tournament, the seeding (1,4,2,3) maximizes this probability. The authors find that matters are more complicated in an 8-player tournament where eight different seedings can be optimal depending on the matrix of winning probabilities. However, Ryvkin (2005) shows that, if winning probabilities depend on the ranks of players "smoothly", only one seeding can be optimal with 8 players. Horen and Riezman (1985) consider also other objectives such as the probability of a final between the two highest-ranked participants and the expected strength of the winning player. Glickman (2008) incorporated the incomplete information about the participant strengths in the standard model of knock-out tournament. He assumed that knowledge about player strengths is given by a multivariate normal distribution and formulated the corresponding general optimization problem.

The economics literature on tournament design has been traditionally concerned with total effort exerted by players. Thus, in a typical model the winning probabilities p_{ij} depend on strategic choices of effort and are endogenous. The variable of designer's choice is usually the prize structure of a tournament; the question of optimal seeding has been addressed to a lesser extent. As a supplementary result, Rosen (1986) finds that in a simple numerical example with two rounds a random seeding can yield higher total effort than the distant seedings. Groh et al. (2012) report that in Rosen's example, a close seeding yields even higher total effort. Also, they find that the seeding (1,4,2,3) maximizes a tournament's predictive power while the seeding (1,3,2,4) maximizes

Author's personal copy

both the total effort and the probability of the final between the two highest-ranked players.

Ely et al. (2015) build a novel framework with agents demanding noninstrumental information and find dynamic information policies that maximize suitably defined "suspense" and "surprise". This framework is directly applicable to modeling spectator interest in sporting events, with "suspense" being related to competitive intensity. In one of their examples, the authors consider the problem of finding suspense- and surprise-optimal seedings in a simple three-player knock-out tournament in which one of the players has a first-round "bye". They find that a seeding in which the two strongest teams play already in the first round and the weakest one has a bye generates most surprise and, frequently, most suspense. This result, though limited to the simplest example, is in line with our findings.

Unlike predictive power and total effort, the objective function that we consider has been studied little. Vu (2010) constructs a "revenue" function that includes two terms reflecting quality and competitive intensity, though the construction is different from ours. He then simulates the parameters of the model, including the matrix of winning probabilities; for every realization of parameters he does exhaustive search and finds the optimal seeding in a tournament with 8 players. The author finds that the traditional seeding is optimal in 23% of cases and achieves, on average, more than 99% of the optimal value. No seeding turns out to be optimal more often than the traditional one.

The number of non-trivially different seedings in a knock-out tournament with $N = 2^n$ participants is $N!/(2^{N-1})$. This quantity grows rapidly in N that makes both analytic work and exhaustive search generally hard to do (note that the majority of the results described above are for $N \leq 8$ only). In contrast, we provide results for an arbitrarily large number of participants.

3 Framework

3.1 The optimization problem

The organizers seek to maximize the overall spectator interest in watching the matches of the tournament. We assume that spectator interest in a single match depends positively on (i) the strength of the teams involved and (ii) the degree of balancedness of the match. Such a formulation conforms to intuition and is in accord with several previous treatments [see, for example, expressions for "demand" in Palomino and Rigotti (2000) and "revenue" in Vu (2010)]. It is also plausible that, other things being equal, a match later in the tournament attracts more attention than the same match in the beginning of the tournament.

Therefore, we model the spectator interest in a single match between teams i and j happening in round r (denoted by D^r_{ij}) as follows:

$$D^r_{ij} = \alpha^r f\left[\gamma \cdot (s_i + s_j) - |s_i - s_j|\right], \quad (1)$$

where $\alpha \geq 1$ is a coefficient that reflects how rapidly the attention to the tournament increases as it unfolds. The term $(s_i + s_j)$ is *the quality* of the match while the term

Author's personal copy

$-|s_i - s_j|$ is *the competitive intensity*. The relative importance of quality as compared to competitive intensity is given by the coefficient $\gamma \geq 0$. The function $f : \mathbb{R} \to \mathbb{R}$ captures possible nonlinearities in the effect of quality and competitive intensity on the spectator interest. We assume that f is strictly increasing and continuously differentiable.

Denote by M_r the set of matches played in round r. M_r depends both on the seeding x and the random outcomes of previous matches. Suppose that for any two teams i and j, there exists a fixed probability p_{ij} that i beats j. Let $P = (p_{ij})$. We assume that matrix P has the following properties: (i) $p_{ij} + p_{ji} = 1$ for all i, j; (ii) p_{ij} is nonincreasing in i and nondecreasing in j. Such probability matrices are sometimes called *doubly monotonic*.

To highlight the dependence of M_r on seeding x, denote it by $M_r(x)$. The probability distribution of $M_r(x)$ can be computed given P, x and the rules of the knock-out tournament. Then, the expression

$$\mathbb{E}D(x) = \mathbb{E}\sum_{r=1}^{n} \sum_{(ij)\in M_r(x)} D_{ij}^r$$

represents expected spectator interest of a whole tournament, where the notation $(ij) \in M_r(x)$ means that the match between teams i and j belongs to the random set $M_r(x)$. We study the following problem:

$$\max_x \mathbb{E}D(x). \tag{2}$$

The results in this paper are formulated under the following assumption on P:

Assumption 1 For any two teams i and j, $p_{ij} = 1$ whenever $i < j$, i.e. a stronger team always beats a weaker one.

Due to the finiteness of the set of seedings and the continuity of the expectation with respect to probabilities, all our results also hold when Assumption 1 holds "approximately" i.e. when the probabilities p_{ij} are sufficiently high for $i < j$. It is not difficult to provide counterexamples showing that our results do not hold for any doubly monotonic matrix P.

3.2 Close and distant seedings

Two sets of seedings play a special role in our analysis, as they turn out be the only possible solutions to the problem (2) when f is linear. We first discuss them informally.

As noted in the introduction, *close seedings* are seedings such that under Assumption 1, in every round any team faces an opponent closest to it in rank out of the participants remaining in the tournament. Hence, in the first round the team i where i is odd is paired with the team $i + 1$ and these pairs are placed within the bracket in such a way that in the second round the team 1 faces the team 3, the team 5 faces the team 7, and so on.

Author's personal copy

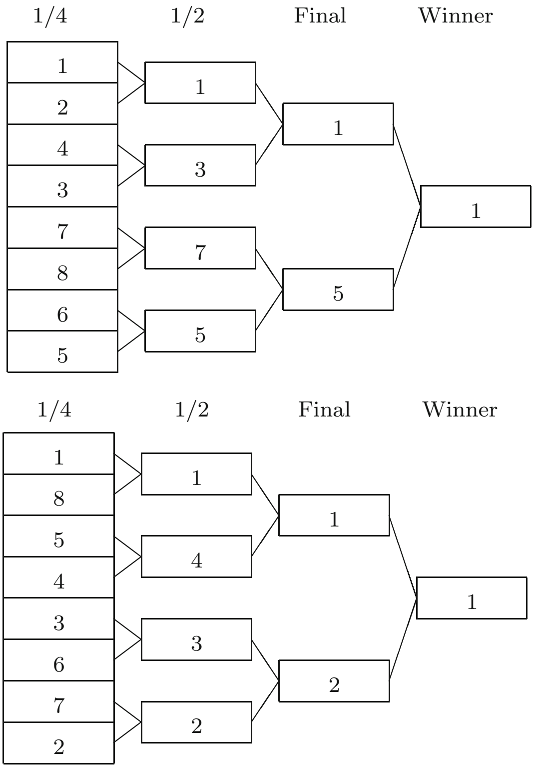

Fig. 2 The examples of close (*above*) and distant (*below*) seedings; the tournament bracket is filled in as anticipated under Assumption 1

In contrast, *distant seedings* are such seedings that under Assumption 1 a team from the strongest half of remaining teams meets a team from the weakest half of remaining teams in each match of the tournament. For example, for $n = 3$ the traditional seeding $(1, 8, 4, 5, 3, 6, 2, 7)$ is a distant seeding. Figure 2 shows again the examples of a close and a distant seeding, this time we fill the tournament bracket with the winners of each match under Assumption 1.

We now give formal definitions of close and distant seedings that do not make direct use of Assumption 1. We define “close tuples” and “distant tuples”. A close seeding is a seeding that can be represented by a close tuple, and a distant seeding is a seeding that can be represented by a distant tuple.

Let A and B be two finite subsets of $\mathbb{N}$. We say that A and B *do not overlap* if either the smallest element of A is greater than the largest element of B or the smallest element of B is greater than the largest element of A. Otherwise we say that A and B overlap.

We give formal definitions of close and distant tuples by induction over the number of rounds, n. In these definitions, we assume that tuples consist of natural numbers.

Definition 1 (Close 2^k-tuples)

(1) If $k = 1$, any 2^k-tuple is close;

Author's personal copy

(2) Suppose we defined close 2^k-tuples for some $k = l \geq 1$. A 2^{l+1}-tuple x is close if and only if two conditions hold:
   (a) The 2^l-tuples $(x_1, x_2, \dots, x_{2^l})$ and $(x_{2^l+1}, x_{2^l+2}, \dots, x_{2^{l+1}})$ are close;
   (b) The sets $\{x_1, x_2, \dots, x_{2^l}\}$ and $\{x_{2^l+1}, x_{2^l+2}, \dots, x_{2^{l+1}}\}$ do not overlap.

For a 2^n-tuple x, let $W(x)$ be the set $\{t \in \mathbb{N} : \exists k : t = \min\{x_{2k-1}, x_{2k}\}\}$. Analogously, let $L(x)$ be the set $\{t \in \mathbb{N} : \exists k : t = \max\{x_{2k-1}, x_{2k}\}\}$. That is, under Assumption 1 $W(x)$ would be the set of teams that win of the first round and $L(x)$ would be the set of teams that lose in first round if the seeding is x.

Definition 2 (Distant 2^k-tuples)

(1) If $k = 1$, any 2^k-tuple is distant;
(2) Suppose we defined distant 2^k-tuples for some $k = l$. A 2^{l+1}-tuple x is distant if and only if two conditions hold:
   (a) The 2^l-tuple $(\min\{x_1, x_2\}, \min\{x_3, x_4\}, \dots, \min\{x_{2^{l+1}-1}, x_{2^{l+1}}\})$ is distant;
   (b) The sets $W(x)$ and $L(x)$ do not overlap.

Note that in case of four participants, any seeding is either close or distant. However, as number of participants grows, the share of close and distant seedings in the set of all seedings converges to 0. One may compute that the share of close and distant seedings taken together is 15.56 % for $n = 3$, 0.3 % for $n = 4$ and $3.3 \cdot 10^{-7}$ for $n = 5$.[1] Note also that under Assumption 1 any close seeding leads to the same set of matches being played (and thus the same value of the objective function), whereas different distant seedings can involve different pairings of teams from the stronger half and the weaker half, and thus yield different values of the objective function. Under linearity of f, however, all distant seedings generate the same level of spectator interest.

4 Results: linear f

Denote the set of close seedings by C, and the set of distant seedings by D. Denote by $X^*(s, \alpha, \gamma)$ the set of optimal seedings as a function of the tuple of strengths s, the later-round preference parameter α and the quality preference parameter γ.

Theorem 1 *Suppose f is linear, and Assumption 1 holds. Then:*

(1) If $\alpha(\gamma + 1) < 2$, then $\forall s\ X^\star(s, \alpha, \gamma) = C$;
(2) If $\alpha(\gamma + 1) > 2$, then $\forall s\ X^\star(s, \alpha, \gamma) = D$;
(3) If $\alpha(\gamma + 1) = 2$, then any seeding is optimal.

Denote by P_r the set of participants of round r. Denote by W_r the set of winners of round r and by $L_r = P_r \setminus W_r$ the set of losers of round r. Consider special sets

$$W_r^* = \{1, 1 + 2^r, 1 + 2 \cdot 2^r, 1 + 3 \cdot 2^r, \dots, 2^n + 1 - 2^r\}$$
$$\text{and}\ \ W_r^{**} = \{1, 2, 3, \dots, 2^{n-r}\}.$$

[1] The share of close seedings in the set of all seedings is equal to $\frac{2^{2^n-1}}{(2^n)!}$ while the share of distant seedings is the same number multiplied by $\prod_{k=1}^{n-1}(2^k)!$.

Author's personal copy

Lemma 1 *For any round r, $\sum_{i\in W_r^*} s_i \le \sum_{i\in W_r} s_i \le \sum_{i\in W_r^{**}} s_i$.*

Proof of Lemma 1 The inequality $\sum_{i\in W_r} s_i \le \sum_{i\in W_r^{**}} s_i$ is obvious. As for the inequality $\sum_{i\in W_r^*} s_i \le \sum_{i\in W_r} s_i$, consider the weakest team that wins in round r. It is a winner of a sub-tournament with 2^r participants, so there are at least $2^r - 1$ teams weaker than it. Hence, its strength is at least $s_{2^n+1-2^r}$. Consider the second weakest team that wins in round r. It is a winner of another sub-tournament with 2^r participants, so there exist *another* $2^r - 1$ teams that are weaker than it. Overall, it must be stronger than $2^r - 1 + 2^r$ teams, so its strength is at least $s_{2^n+1-2\cdot 2^r}$. Proceeding in this fashion, we see that the strength of the i^{th} weakest winner of round r is at least $s_{2^n+1-i\cdot 2^r}$ which implies the result. □

Proof of Theorem 1 Without loss of generality, consider $f(t) = t$. Consider the quantities $u_r = \sum_{(ij)\in M_r(x)} (\gamma(s_i + s_j) - |s_i - s_j|)$. Note that for any r,

$$\begin{aligned} u_r &= \gamma \sum_{i\in P_r} s_i - \left(\sum_{i\in W_r} s_i - \sum_{i\in L_r} s_i\right) = \gamma \sum_{i\in P_r} s_i - \left(2\sum_{i\in W_r} s_i - \sum_{i\in P_r} s_i\right) \\ &= (\gamma+1)\sum_{i\in P_r} s_i - 2\sum_{i\in W_r} s_i. \end{aligned} \tag{3}$$

The objective function is equal to $\sum_{r=1}^{n} \alpha^r u_r$. Substituting u_r from (3) and using the fact that by the nature of a knock-out tournament $W_{r-1} = P_r$, one gets that

$$\begin{aligned} \sum_{r=1}^{n} \alpha^r u_r &= \alpha(\gamma+1)\sum_{i\in P_1} s_i + \left(\alpha^2(\gamma+1) - 2\alpha\right)\sum_{i\in W_1} s_i + \left(\alpha^3(\gamma+1) - 2\alpha^2\right) \\ &\quad \sum_{i\in W_2} s_i + \cdots + \left(\alpha^n(\gamma+1) - 2\alpha^{n-1}\right)\sum_{i\in W_{n-1}} s_i - 2\sum_{i\in W_n} s_i. \end{aligned} \tag{4}$$

Under Assumption 1, $2\sum_{i\in W_n} s_i = 2s_1$ and $\sum_{i\in P_1} s_i$ are just constants so eventually we should maximize the expression

$$(\alpha(\gamma+1) - 2)\sum_{r=1}^{n-1} \alpha^r \sum_{i\in W_r} s_i. \tag{5}$$

Then there are three cases.

Case 1 $\alpha(\gamma+1) < 2$, so we should minimize the expression $\sum_{r=1}^{n-1} \alpha^r \sum_{i\in W_r} s_i$. By Lemma 1, $\sum_{i\in W_r} s_i$ is minimized when $W_r = W_r^*$. The key point is that it is feasible

Author's personal copy

to set $W_r = W_r^*$ simultaneously for all r. It is evident that this happens if and only if the seeding is close.

Case 2 $\alpha(\gamma + 1) > 2$, so we should maximize the expression $\sum_{r=1}^{n-1} \alpha^r \sum_{i \in W_r} s_i$. By Lemma 1, $\sum_{i \in W_r} s_i$ is maximized when $W_r = W_r^{**}$. Again, it is feasible to set $W_r = W_r^{**}$ simultaneously for all r. This happens if and only if the seeding is distant.

Case 3 $\alpha(\gamma + 1) = 2$. The objective function is constant, so any seeding is optimal. □

Theorem 1 shows that if the effect of a match's quality and competitive intensity on spectator interest is linear, only two types of seedings—close seedings or distant seedings—can possibly maximize the objective function. It also elucidates the way the solution to Problem (2) depends on the parameters α and γ.

This relationship is intuitive. First, note that there is a trade-off between competitive intensity at early stages and late stages of the tournament; close seedings generate great intensity in first rounds, but a very unbalanced final, whereas distant seedings create unbalanced matches early in the tournament, but guarantee a final between the top two teams. As a result, close seedings are optimal when α is relatively low while distant seedings are optimal when α is relatively high. (Close seedings can be optimal even if $\alpha > 1$ because the sheer number of matches in the beginning of the tournament is greater than the number of matches at later stages.)

Second, notice that the longer strong teams are not eliminated from the tournament the higher its overall quality is (as high levels of strength are counted more times in (2)). Close seedings eliminate top teams quickly (except the strongest one) while distant seedings favor strong teams by pairing them with weak ones. As a result, close seedings are optimal when γ is relatively low while distant seedings are optimal when γ is relatively high. When the spectators care mostly about quality (γ is arbitrarily large), distant seedings are always optimal; by contrast, when the spectators care only about competitive intensity ($\gamma = 0$) and $\alpha < 2$, close seedings are optimal. Thus, when $\alpha < 2$ so that spectator preference for later-stage matches is not too strong, there also exists a trade-off between the tournament's overall quality and competitive intensity.

Finally, note that the set of optimal seedings does not depend on cardinal levels of the teams' strength even though they enter the objective function explicitly. Thus, in order to implement the solution stated in Theorem 1, the organizers would have to know only the relative ranking of the teams, and the value of the parameters α and γ.

5 Results: general f

To which extent do the results of the previous section generalize when f is not longer linear? In this section, we show that close and distant seedings can arise as a solution to Problem (2) for a nontrivial set of functions.

Author's personal copy

5.1 Optimality of close seedings

First, suppose that the spectators care only about competitive intensity, i.e. $\gamma = 0$. In this case we are able to give, for a fixed n and α, both necessary conditions and sufficient conditions on f for close seedings to be the only optimal seedings. These necessary and sufficient conditions differ only insignificantly; in a sense, what we provide is almost a charaterization of the set of functions f such that $\forall s\ X^*(s, \alpha, \gamma) = C$.

Note that for $\gamma = 0$ only negative arguments enter f in (1). Therefore, in this subsection f is understood as a function from the set of nonpositive real numbers to $\mathbb{R}$.

Recall that function f is called *subadditive* if $\forall\ u, v$ from the domain $u + v$ also belongs to the domain and $f(u + v) \leq f(u) + f(v)$.

Theorem 2 *(A sufficient condition for the optimality of close seedings)*
Suppose Assumption 1 holds and $\gamma = 0$.

(1) Fix $n = 2$ *and suppose that:*
  (a) f *is subadditive;*
  (b) $\inf f'(t) > (\alpha - 1) \sup f'(t)$.
  Then $\forall s\ X^*(s, \alpha, \gamma) = C$.
(2) Fix $n \geq 3$ *and suppose that:*
  (a) f *is subadditive;*
  (b) $\inf f'(t) > \frac{\alpha^2}{2+\alpha} \sup f'(t)$.
  Then $\forall s\ X^*(s, \alpha, \gamma) = C$.

The sufficient conditions stated in the Theorem show that close seedings are indeed optimal for any numerical levels of strength in a considerable number of cases. The conditions (b) ensure that the variation in the derivative of f is not too high. In a sense, they give a precise statement of the idea that f should not differ too much from a linear function. Note that when f is linear, $\inf f'(t) = \sup f'(t)$ and so both conditions (b) become just $\alpha \leq 2$, which is in accord with Theorem 1. Subadditivity ensures that in every two-round sub-tournament, the close structure $(1, 2, 3, 4)$ is weakly better than the structure $(1, 4, 2, 3)$.

To prove Theorem 2 for the case $n \geq 3$, we need the following lemma.

Lemma 2 *Suppose Assumption 1 holds, and a seeding* x_0 *is optimal. Then all the seedings induced by* x_0 *in all sub-tournaments are optimal in the corresponding sub-tournaments, i.e. they maximize spectator interest in the corresponding sub-tournaments given the sets of participants in the sub-tournaments.*

The proof of the lemma is evident so we omit it. □

Proof of Theorem 2 **Part 1** ($n = 2$).

It is sufficient to prove that the seeding (1,2,3,4) is strictly better than both seedings (1,4,2,3) and (1,3,2,4). As for the seeding (1,3,2,4), we should prove the inequality

$$f(-|s_1 - s_2|) + f(-|s_3 - s_4|) + \alpha f(-|s_1 - s_3|) > f(-|s_1 - s_3|) + f(-|s_2 - s_4|) + \alpha f(-|s_1 - s_2|). \quad (6)$$

Author's personal copy

This may be rewritten as

$$f(s_4 - s_3) - f(s_4 - s_2) > (\alpha - 1)[f(s_2 - s_1) - f(s_3 - s_1)]. \tag{7}$$

By Mean Value Theorem, it is true that for an increasing differentiable function f and any two points a and b, $a < b$, $\inf f'(t)(b-a) \leq f(b) - f(a) \leq \sup f'(t)(b-a)$.
So we have

$$\begin{aligned} f(s_4 - s_3) - f(s_4 - s_2) &\geq \inf f'(t)(s_2 - s_3) > (\alpha - 1) \sup f'(t)(s_2 - s_3) \geq \\ &\geq (\alpha - 1)[f(s_2 - s_1) - f(s_3 - s_1)], \end{aligned} \tag{8}$$

where the second inequality is by assumption.

As for the seeding (1,4,2,3), we should prove the inequality

$$\begin{aligned} f(-|s_1 - s_2|) + f(-|s_3 - s_4|) + \alpha f(-|s_1 - s_3|) > f(-|s_1 - s_4|) + f(-|s_2 - s_3|) \\ + \alpha f(-|s_1 - s_2|). \end{aligned} \tag{9}$$

This can be rewritten as

$$f(s_3 - s_1) + f(s_4 - s_3) > f(s_4 - s_1) + f(s_3 - s_2) + (\alpha - 1)[f(s_2 - s_1) - f(s_3 - s_1)]. \tag{10}$$

The inequality

$$f(s_3 - s_1) + f(s_4 - s_3) \geq f(s_4 - s_1). \tag{11}$$

is true by subadditivity. However, it is also true that

$$\begin{aligned} f(0) - f(s_3 - s_2) &\geq \inf f'(t)(s_2 - s_3) > (\alpha - 1) \sup f'(t)(s_2 - s_3) \geq \\ &\geq (\alpha - 1)[f(s_2 - s_1) - f(s_3 - s_1)], \end{aligned} \tag{12}$$

where the second inequality is again by assumption. Given that $f(0) = 0$, adding inequalities (11) and (12) proves the desired inequality (9).

Part 2 ($n \geq 3$).

The proof is by induction. We prove the statement not only for $n \geq 3$, but for all $n \geq 2$.

Induction base ($n = 2$) The inequality $\inf f'(t) > \frac{\alpha^2}{2+\alpha} \sup f'(t)$ implies that $1 \leq \alpha < 2$ as an infimum cannot be strictly greater than a supremum. But for $1 \leq \alpha < 2$ we have $\frac{\alpha^2}{2+\alpha} > (\alpha - 1)$ so the condition in the Part 1 is satisfied. Hence, the result of Part 1 applies.

Induction step Suppose close seedings are strictly optimal for any tournament (and sub-tournament) with $n = l \geq 2$ rounds. Consider a tournament with $n = l + 1$ rounds.

Take any optimal seeding $x^\star = (x_1, x_2, \ldots, x_{2^{l+1}})$. Suppose $x^\star$ is not close. By definition, there are three possibilities: (i) the tuple $(x_1, x_2, \ldots, x_{2^l})$ is not close;

Author's personal copy

(ii) the tuple $(x_{2^l+1}, x_{2^l+2}, \ldots, x_{2^{l+1}})$ is not close; (iii) the sets $\{x_1, x_2, \ldots, x_{2^l}\}$ and $\{x_{2^l+1}, x_{2^l+2}, \ldots, x_{2^{l+1}}\}$ overlap.

Note that the tuple $(x_1, x_2, \ldots, x_{2^l})$ is a seeding in the upper-bracket sub-tournament of the grand tournament. By Lemma 2, this seeding should be optimal in the sub-tournament. But there are 2^l participants in this sub-tournament, so by induction hypothesis, the tuple $(x_1, x_2, \ldots, x_{2^l})$ should be close. Analogously, the tuple $(x_{2^l+1}, x_{2^l+2}, \ldots, x_{2^{l+1}})$ should be close. This rules out the first two possibilities. We are left with the third one: suppose that the sets $A = \{x_1, x_2, \ldots, x_{2^l}\}$ and $B = \{x_{2^l+1}, x_{2^l+2}, \ldots, x_{2^{l+1}}\}$ overlap. Without loss of generality, assume that the strongest team belongs to A. There are again two cases.

Case 1 There are two or more teams in A which are weaker than the strongest team in B. Let u be the weakest team in A and v be the second weakest. As the tuple $(x_1, x_2, \ldots, x_{2^l})$ is close, these teams play against each other in the first round, and v wins. As v is weaker than the strongest team in B, the tuple $w = (\min\{x_1, x_2\}, \min\{x_3, x_4\}, \ldots, \min\{x_{2^{l+1}-1}, x_{2^{l+1}}\})$ is not close. However, w is a seeding in a sub-tournament of the grand tournament (this sub-tournament includes all the matches of the grand tournament except first-round matches). There are 2^l participants in this sub-tournament, so by Lemma 2 and the induction hypothesis, w should be close. Contradiction.

Case 2 There exists exactly one team in A which is weaker than the strongest team in B. Let this team be $u > 2^l$. The strongest team in B should be the team 2^l.

Now generate a new seeding $\hat{x}$ by switching the positions of u and the team 2^l in $x^\star$. We claim that $\hat{x}$ generates greater spectator interest than $x^\star$.

Subcase 1 $u = 2^l + 1$. The only differences in matches between x^* and $\hat{x}$ are as follows: in round 1, $\hat{x}$ assigns matches $(2^l - 1, 2^l)$ and $(2^l + 1, 2^l + 2)$ whereas x^* assigns matches $(2^l - 1, 2^l + 1)$ and $(2^l, 2^l + 2)$. In rounds 2 through $n - 1$, $\hat{x}$ assigns the team $2^l + 1$ to play against various weaker teams t_i, whereas x^* assigns the team 2^l to play against the very same teams. Finally, in round $n = l + 1$, $\hat{x}$ assigns team 1 to play with $2^l + 1$ whereas x^* assigns team 1 to play with 2^l.

So the objective function at $\hat{x}$ is strictly greater than at x^* if and only if

$$f(s_{2^l} - s_{2^l-1}) + f(s_{2^l+2} - s_{2^l+1}) + \sum_{i=2}^{l} \alpha^{i-1} f(s_{t_i} - s_{2^l+1}) + \alpha^l f(s_{2^l+1} - s_1) >$$
$$> f(s_{2^l+1} - s_{2^l-1}) + f(s_{2^l+2} - s_{2^l}) + \sum_{i=2}^{l} \alpha^{i-1} f(s_{t_i} - s_{2^l}) + \alpha^l f(s_{2^l} - s_1). \quad (13)$$

Rearranging terms, we get

$$[f(s_{2^l} - s_{2^l-1}) - f(s_{2^l+1} - s_{2^l-1})] + [f(s_{2^l+2} - s_{2^l+1}) - f(s_{2^l+2} - s_{2^l})] +$$
$$+ \sum_{i=2}^{l} \alpha^{i-1} [f(s_{t_i} - s_{2^l+1}) - f(s_{t_i} - s_{2^l})] > \alpha^l [f(s_{2^l} - s_1) - f(s_{2^l+1} - s_1)]. \quad (14)$$

Author's personal copy

Applying bounds given by mean value theorem (similar to that in Part 1) and dividing both parts by $(s_{2^l} - s_{2^l+1})$ we see that the inequality (14) is implied by the inequality

$$\left(2 + \sum_{i=1}^{l-1} \alpha^i\right) \inf f'(t) > \alpha^l \sup f'(t).$$

But this inequality follows directly from the assumption of the Theorem and the fact that for all $l \geq 2$ and $1 \leq \alpha < 2$

$$\frac{\alpha^l}{2 + \alpha + \alpha^2 + \cdots + \alpha^{l-1}} \leq \frac{\alpha^2}{2 + \alpha}.$$

Hence, $\hat{x}$ generates greater spectator interest, and x^* is not optimal. Contradiction.

Subcase 2 $u > 2^l + 1$. The only differences in matches between $\hat{x}$ and x^* are as follows: in round 1, $\hat{x}$ assigns matches $(2^l - 1, 2^l)$ and $(2^l + 1, u)$ whereas x^* assigns matches $(2^l - 1, u)$ and $(2^l, 2^l + 1)$. In rounds 2 through $n - 1$, $\hat{x}$ assigns the team $2^l + 1$ to play against various weaker teams t_i, whereas x^* assigns the team 2^l to play against the very same teams. Finally, in round $n = l + 1$, $\hat{x}$ assigns team 1 to play with $2^l + 1$ whereas x^* assigns team 1 to play with 2^l.

So the objective function at $\hat{x}$ is strictly greater than at x^* if and only if

$$\begin{aligned} f(s_{2^l} - s_{2^l-1}) + f(s_u - s_{2^l+1}) + \sum_{i=2}^{l} \alpha^{i-1} f(s_{t_i} - s_{2^l+1}) + \alpha^l f(s_{2^l+1} - s_1) > \\ > f(s_u - s_{2^l-1}) + f(s_{2^l+1} - s_{2^l}) + \sum_{i=2}^{l} \alpha^{i-1} f(s_{t_i} - s_{2^l}) + \alpha^l f(s_{2^l} - s_1). \end{aligned} \tag{15}$$

By subadditivity, $f(s_u - s_{2^l-1}) \leq f(s_{2^l+1} - s_{2^l-1}) + f(s_u - s_{2^l+1})$. Using this inequality and then applying the technique similar to those in Part 1 and Subcase 1 above, proves inequality (15).

Hence, again, x^* is not optimal. Contradiction.

This proves that any optimal seeding is close. To prove that any close seeding is optimal, note that an optimal seeding exists and all close seedings yield the same value of the objective function. □

It is remarkable that the conditions stated in Theorem 2 are not only sufficient, but also almost necessary for the optimality of close seedings. The necessary conditions differ from the sufficient conditions only in a knife-edge case when $\inf f'(t) = (\alpha - 1) \sup f'(t)$ for $n = 2$ and $\inf f'(t) = \frac{\alpha^2}{2+\alpha} \sup f'(t)$ for $n \geq 3$.

Theorem 3 *(A necessary condition for the optimality of close seedings.)*
Suppose Assumption 1 holds and $\gamma = 0$.

(1) Fix $n = 2$ *and suppose that* $\forall s\ X^*(s, \alpha, \gamma) = C$. *Then:*
- *(a)* f *is subadditive;*
- *(b)* $\inf f'(t) \geq (\alpha - 1) \sup f'(t)$.

Author's personal copy

(2) Fix $n \geq 3$ and suppose that $\forall s$ $X^*(s, \alpha, \gamma) = C$. Then:
   (a) f is subadditive;
   (b) $\inf f'(t) \geq \frac{\alpha^2}{2+\alpha} \sup f'(t)$.

Lemma 3 *Let f be a subadditive strictly increasing and continuously differentiable function $\mathbb{R}_- \to \mathbb{R}$ satisfying $f(0) = 0$. Then* $\inf f'(x) = f'_-(0)$, *where $f'_-(0)$ is the left derivative at zero.*

Proof By definition of a derivative we have

$$f'(t) = \lim_{\Delta t \to 0-} \frac{f(t + \Delta t) - f(t)}{\Delta t} \geq \lim_{\Delta t \to 0-} \frac{f(t) + f(\Delta t) - f(t)}{\Delta t} = f'_-(0),$$

where the inequality follows from subadditivity of function f and the fact that $\Delta t < 0$. □

Proof of Theorem 3 **Part 1** ($n = 2$).

A close seeding (1,2,3,4) should be strictly better than the seeding (1,4,2,3) (this is inequality (9) again:

$$f(-|s_1 - s_2|) + f(-|s_3 - s_4|) + \alpha f(-|s_1 - s_3|) > f(-|s_1 - s_4|) + f(-|s_2 - s_3|) + \alpha f(-|s_1 - s_2|). \tag{16}$$

So we again have

$$f(s_3 - s_1) + f(s_4 - s_3) > f(s_4 - s_1) + f(s_3 - s_2) + (\alpha - 1)[f(s_2 - s_1) - f(s_3 - s_1)]. \tag{17}$$

Fix s_1, s_3, s_4 and let $s_2 \to s_3$. As by assumption f is continuous and $f(0) = 0$, the term $f(s_3 - s_2) + (\alpha - 1)[f(s_2 - s_1) - f(s_3 - s_1)]$ can be made arbitrarily small so we should have

$$f(s_3 - s_1) + f(s_4 - s_3) \geq f(s_4 - s_1). \tag{18}$$

This is nothing but the subadditivity.

Also, (1,2,3,4) should be strictly better than the seeding (1,3,2,4). So we have inequality (6) again:

$$f(s_4 - s_3) - f(s_4 - s_2) > (\alpha - 1)[f(s_2 - s_1) - f(s_3 - s_1)]. \tag{19}$$

Fix all strengths except s_2 and let $s_2 \to s_3$. Then use first-order Taylor expansions around the limit points. One gets:

$$f'(s_4 - s_3)(s_2 - s_3) > (\alpha - 1) f'(s_3 - s_1)(s_2 - s_3) + o(s_2 - s_3). \tag{20}$$

Now divide both parts by $(s_2 - s_3)$ and note that the term $o(s_2 - s_3)/(s_2 - s_3)$ can be ignored if the strict inequality is replaced with a weak one. Let $s_4 \to s_3$. As by assumption the derivative of f is continuous, one gets

$$f'_-(0) \geq (\alpha - 1) f'(s_3 - s_1). \tag{21}$$

Author's personal copy

By subadditiviy and Lemma 3, $f'_-(0) = \inf f'(t)$. So, by varying s_1 we immediately get the required result:

$$\inf f'(t) \geq (\alpha - 1) \sup f'(t).$$

Part 2 ($n \geq 3$).

Fix α and suppose $\forall s\ X^\star(s, \alpha, 0) = C$, i.e. close seedings are strictly optimal. In particular, this means that the close seeding $(1, 2, 3, 4, 5, 6, \ldots, 2^n)$ is strictly better than the seeding $(1, 4, 2, 3, 5, 6, \ldots, 2^n)$. Note that the sets of matches generated by these two seedings differ only in a two-round sub-tournament won by the team 1. Hence, we should have that the seeding $(1, 2, 3, 4)$ in this two-round sub-tournament should be strictly better than $(1, 4, 2, 3)$. Applying the argument from Part 1 to this sub-tournament, we get that f is subadditive.

In order to prove that the inequality $\inf |f'(x)| \geq \frac{\alpha^2}{2+\alpha} \sup |f'(x)|$ holds, note that the close seeding $(1, 2, 3, 4, 5, 6, 7, 8, \ldots, 2^n)$ should be strictly better than $(1, 2, 3, 5, 4, 6, 7, 8, \ldots, 2^n)$. The sets of matches generated by these two seedings differ only in a three-round sub-tournament won by team 1. The corresponding inequality, then, boils down to

$$\begin{aligned}&[f(s_4 - s_3) - f(s_5 - s_3)] + [f(s_6 - s_5) - f(s_6 - s_4)]\\&\quad + \alpha[f(s_7 - s_5) - f(s_7 - s_4)] > \alpha^2[f(s_4 - s_1) - f(s_5 - s_1)].\end{aligned} \tag{22}$$

Fix all strengths except s_4 and let $s_4 \to s_5$. Then use first-order Taylor expansions around the limit points. One gets:

$$\begin{aligned}&f'(s_5 - s_3)(s_4 - s_5) + f'(s_6 - s_5)(s_4 - s_5) + \alpha f'(s_7 - s_5)|(s_4 - s_5)\\&\quad > \alpha^2 f'(s_5 - s_1)(s_4 - s_5) + o(s_4 - s_5).\end{aligned} \tag{23}$$

Now divide both parts by $(s_4 - s_5)$, ignore the term $o(s_4 - s_5)/(s_4 - s_5)$ and let all strengths except s_1 go to s_3. As by assumption the derivative of f is continuous, one gets

$$(2 + \alpha) f'_-(0) \geq \alpha^2 f'(s_3 - s_1). \tag{24}$$

By subadditivity and Lemma 3, $f'_-(0) = \inf f'(t)$. So, by varying s_1 we get the required result:

$$\inf f'(t) \geq \frac{\alpha^2}{2 + \alpha} \sup f'(t).$$

□

An infimum cannot be strictly greater than a supremum; this implies that $(\alpha - 1) \leq 1$ and $\frac{\alpha^2}{\alpha+2} \leq 1$. Hence, we obtain the following corollary.

Corollary 1 *Suppose the spectators care only about competitive intensity. If $\alpha > 2$ there does not exist a strictly increasing and continuously differentiable function f*

Author's personal copy

satisfying $f(0) = 0$ *such that close seedings maximize spectator interest for any tuple of strengths.* □

We know already from Theorem 1 that if f is linear, $\gamma = 0$, and $\alpha > 2$, close seedings are not optimal. Corollary 1 shows that if the preference for later-stage matches is sufficiently strong, no nonlinear effect of competitive intensity on spectator interest can restore the optimality of close seedings.

Ignoring the knife-edge cases, Theorems 2 and 3, taken together, provide a characterization of the set of functions $F(\alpha, n)$ such that close seedings are the only optimal seedings in the n-round tournament given the parameter α. Note that $F(\alpha, 3)$ is substantially smaller than $F(\alpha, 2)$ since the condition $\inf f'(t) \geq \frac{\alpha^2}{2+\alpha} \sup f'(t)$ is strictly more restrictive than the condition $\inf f'(t) \geq (\alpha - 1) \sup f'(t)$. This is understandable given the fact that a three-round tournament is a more complex structure than a two-round tournament and thus more conditions must hold for close seedings to be optimal. However, it is not the case that $F(\alpha, 4)$ is substantially smaller than $F(\alpha, 3)$. Indeed, for $n > 3$, sets $F(\alpha, n)$ differ from $F(\alpha, 3)$ in at most the knife-edge case when the condition (b) in part 2 of Theorem 3 holds as equality. This counterintuitive result suggests that, in a certain sense, there is a qualitative increase in the complexity of tournament structure when the number of rounds rises from 2 to 3 only; further increases in the number of rounds have a less significant effect.

5.2 Optimality of distant seedings

In this subsection, we provide the result analogous to Theorem 2 that deals with the optimality of distant seedings.

Unlike close seedings, different distant seedings in general result in different sets of matches being played and thus generate different levels of spectator interest. This complicates the analysis, and the results for distant seedings are true only in a weaker form than the results for close seedings. Namely, we have to replace the statement $X^\star(s, \alpha, \gamma) = C$ with $X^\star(s, \alpha, \gamma) \subset D$. Moreover, we state only a sufficient condition, but not a necessary condition, for the optimality of distant seedings. However, in the case of distant seedings we do not have to assume that $\gamma = 0$; we provide results for an arbitrary nonnegative value of γ.

Theorem 4 *(A sufficient condition for the optimality of distant seedings) Suppose Assumption 1 holds. Fix* $n \geq 1$ *and suppose* $\alpha(\gamma + 1) \inf f'(x) > 2 \sup f'(x)$. *Then* $\forall s\ X^\star(s, \alpha, \gamma) \subset D$.

Again, note that the above result is in total accord with Theorem 1 when f is linear.

Proof Theorem 4 The proof is by induction.

Induction base $(n = 1)$ is obvious.

Induction step Suppose that the statement has been proved for $n = l \geq 1$. Consider $n = l + 1$.

Take any optimal seeding $x^* = (x_1, x_2, \ldots, x_{2^{l+1}})$. Suppose x^* is not distant. By definition, there are two possibilities:

Author's personal copy

(i) the tuple $w = (\min\{x_1, x_2\}, \min\{x_3, x_4\}, \ldots, \min\{x_{2^{l+1}-1}, x_{2^{l+1}}\})$ is not distant;
(ii) the sets $W(x^*)$ and $L(x^*)$ do overlap.

Analogously to the proof of Theorem 2, note that w is the seeding in a sub-tournament with 2^l participants. So, by induction hypothesis and Lemma 2, w should be distant. Therefore, we are left with possibility (ii).

For any team u and seeding y, define the sets $\overline{W}(u|y) = \{w \in W(y) :$ w is stronger than $u\}$ and $\underline{W}(u|y) = \{w \in W(y) :\ w$ is weaker than $u\}$. As $W(x^*)$ and $L(x^*)$ overlap, there exists $u \in L(x^*)$ such that $\underline{W}(u|x^*)$ is not empty. Call the set of all such u's U^*. (As $U^* \subset L(x^*)$, for any $u \in U^*$, $\overline{W}(u|x^*)$ is not empty either).

Lemma 4 *For any $u \in U^*$ there exists $w \in \underline{W}(u|x^*)$ such that w loses in round 2 to a team from $\overline{W}(u|x^*)$.*

Proof of Lemma 4 Suppose there is no such w. Hence, all teams in $\underline{W}(u|x^*)$ that lose in round 2, lose to a team from $\underline{W}(u|x^*)$. Thus, there is an even number of teams in $\underline{W}(u|x^*)$ and half of them win in round 2 and half of them lose. However, as we have seen above, the seeding in round 2 (tuple w) is distant and any distant seeding has the following obvious property: if a team wins in a round, all stronger teams also win in that round. So it must be that all teams in $\overline{W}(u|x^*)$ also win in round 2. But this implies that there are more winning teams in round 2 than losing teams. Contradiction. □

Take any $u \in U^*$ and any $w \in \underline{W}(u|x^*)$ guaranteed by Lemma 4. Generate a new seeding $\hat{x}$ by switching the positions of u and w in x^*. We claim that $\hat{x}$ is strictly better than x^*.

As w loses to a team from $\overline{W}(u|x^*)$ in round two in the original seeding, u will also lose to it. Hence the switch will produce changes in matches played only in the first two rounds. Call the team that u loses to under y, a; the team that w wins in round 1 under y, b, and the team that w loses to in round 2 under y, c.

By simple bookkeeping, $\hat{x}$ is strictly better than x^* whenever

$$\begin{aligned}\alpha[f(\gamma(s_c+s_u)-(s_c-s_u))-f(\gamma(s_c+s_w)-(s_c-s_w))]\\ >[f(\gamma(s_a+s_u)-(s_a-s_u))-\\ -f(\gamma(s_a+s_w)-(s_a-s_w))]\\ +[f(\gamma(s_w+s_b)-(s_w-s_b))-f(\gamma(s_u+s_b)-(s_u-s_b))].\end{aligned} \tag{25}$$

However, (25) is true because the following three inequalities are true:

$$\begin{aligned}&\alpha[f(\gamma(s_c+s_u)-(s_c-s_u))-f(\gamma(s_c+s_w)-(s_c-s_w))]\\ &\quad\geq \alpha(\gamma+1)\inf f'(t)(s_w-s_u), \qquad (26)\\ &\alpha(\gamma+1)\inf f'(t)(s_w-s_u) > 2\sup f'(t)(s_w-s_u), \qquad (27)\\ &2\sup f'(t)(s_w-s_u) \geq [f(\gamma(s_a+s_u)-(s_a-s_u))-f(\gamma(s_a+s_w)\\ &\quad-(s_a-s_w))]+[f(\gamma(s_w+s_b)-(s_w-s_b))-f(\gamma(s_u+s_b)-(s_u-s_b))], \qquad (28)\end{aligned}$$

Author's personal copy

where the first and the third inequalities are our usual bounds guaranteed by Mean Value Theorem and the second inequality is by assumption. Hence, x^* is not optimal. Contradiction. $\square$

6 Concluding remarks

In this paper, we consider the problem of finding optimal seedings in knock-out tournaments where the objective function takes into account the competitive intensity and quality of every match, with the importance of every match increasing in the number of round. We prove that when the objective function is linear, only two classes of seeding can possibly be optimal (Theorem 1). We then identify sufficient (Theorems 2 and 4) and necessary (Theorem 3) conditions under which these classes are optimal in the general case.

Our results imply that the problem is computationally easy for sufficiently high winning probabilities. In contrast, Vu (2010) proves a number of hardness results for the problem of maximizing the winning probability of a given player by the reduction from a vertex cover problem. However, in the general case the computational complexity of the problem studied in our paper, as well as the complexity of other optimal seeding problems, is not known.

References

Baumann R, Matheson VA, Howe C (2010) Anomalies in tournament design: the madness of March madness. J Quant Anal Sports. doi:10.2202/1559-0410.1233

Ely J, Frankel A, Kamenica E (2015) Suspense and Surprise. J Polit Econ 123(1):215–260

Forrest D, Simmons R (2002) Outcome uncertainty and attendance demand in sport: the case of English soccer. J R Stat Soc Ser D Stat 51 Part 2:229–241

Glickman ME (2008) Bayesian locally optimal design of knockout tournaments. J Stat Plann Inference 138:2117–2127

Groh C, Moldovanu B, Sela A, Sunde U (2012) Optimal seedings in elimination tournaments. Econ Theory 49:59–80

Horen J, Riezman R (1985) Comparing draws for single elimination tournaments. Oper Res 33:249–262

Hwang F (1982) New concepts in seeding knockout tournaments. Am Math Mon 89:235–239

Palomino F, Rigotti L (2000) The sport league's dilemma: competitive intensity versus incentives to win. Working Paper, University of California, Department of Economics

Rosen S (1986) Prizes and incentives in elimination tournaments. Am Econ Rev 76(4):701–715

Ryvkin D (2005) The predictive power of noisy elimination tournaments. Working paper, Center for Economic Research and Graduate Education–Economic Institute, Prague

Schwenk A (2000) What is the correct way to seed a knockout tournament. Am Math Monthly 107:140–150

Vu TD (2010) Knockout tournament design: a computational approach. PhD dissertation, Stanford University, Department of Computer Science

Wright M (2014) OR analysis of sporting rules: a survey. Eur J Oper Res 232:1–8