Condorcet-Consistent and Approximately Strategyproof Tournament Rules

Jon Schneider[1], Ariel Schvartzman[2], and S. Matthew Weinberg[3]

1   Department of Computer Science, Princeton University, Princeton, USA
     js44@cs.princeton.edu
2   Department of Computer Science, Princeton University, Princeton, USA
     acohenca@cs.princeton.edu
3   Department of Computer Science, Princeton University, Princeton, USA
     smweinberg@princeton.edu

───── **Abstract** ─────

We consider the manipulability of tournament rules for round-robin tournaments of n competitors. Specifically, n competitors are competing for a prize, and a tournament rule r maps the result of all $\binom{n}{2}$ pairwise matches (called a *tournament*, T) to a distribution over winners. Rule r is *Condorcet-consistent* if whenever i wins all $n-1$ of her matches, r selects i with probability 1.

We consider strategic manipulation of tournaments where player j might throw their match to player i in order to increase the likelihood that one of them wins the tournament. Regardless of the reason why j chooses to do this, the potential for manipulation exists as long as $\Pr[r(T) = i]$ increases by more than $\Pr[r(T) = j]$ decreases. Unfortunately, it is known that every Condorcet-consistent rule is manipulable [1]. In this work, we address the question of *how manipulable* Condorcet-consistent rules must necessarily be - by trying to minimize the difference between the increase in $\Pr[r(T) = i]$ and decrease in $\Pr[r(T) = j]$ for any potential manipulating pair.

We show that every Condorcet-consistent rule is in fact 1/3-manipulable, and that selecting a winner according to a random single elimination bracket is not α-manipulable for any $\alpha > 1/3$. We also show that many previously studied tournament formats are all 1/2-manipulable, and the popular class of Copeland rules (any rule that selects a player with the most wins) are all in fact 1-manipulable, the worst possible. Finally, we consider extensions to match-fixing among sets of more than two players.

1   Introduction

In recent years, numerous scandals have unfolded surrounding match fixing and throwing at the highest levels of competitive sports (e.g. Olympic Badminton [11], Professional Tennis [6], European Football [22], and even eSports [24]). In some instances, the motivation behind these scandals was gambling profits, and no amount of clever tournament design can possibly mitigate this. In others, however, the surprising motivation was an improved performance *at that same tournament.* For instance, four Badminton teams (eight players) were disqualified from the London 2012 Olympics for throwing matches. Interestingly, the reason teams wanted to lose their matches was in order to *improve* their probability of winning an Olympic medal. Olympic Badminton (like many other sports) conducts a two-phase tournament. In the first stage, groups of four play a round-robin tournament, with the top two teams advancing. In

the second stage, the advancing teams participate in a single elimination tournament, seeded according to their performance in the group stage. An upset in one group left one of the world's top teams with a low seed, so many teams actually preferred to receive a *lower* seed coming out of the group stage to face the tougher opponent as late as possible.

While much of the world blames the teams for their poor sportsmanship, researchers in voting theory have instead critiqued the poor tournament design that punished teams for trying to maximize their chances of winning a medal. Specifically, the two-phase tournament lacks the basic property of *monotonicity*, where no competitor can unilaterally improve their chances of winning by throwing a match that they otherwise could have won. Thus, recent work has addressed the question of whether tournament structures exist that are both fair, in that they select some notion of a qualified winner, and strategyproof, in that teams have no incentive to do anything but play their best in each match.

One minimal notion of fairness studied is *Condorcet-consistence*, which just guarantees that whenever one competitor wins *all* of their matches (and is what's called a *Condorcet winner*), they win the event with probability 1. Designing Condorcet-consistent, monotone rules is simple: any single elimination bracket suffices. Popular voting rules such as the Copeland Rule or the Random Condorcet Removal Rule are also Condorcet-consistent and monotone, but two-phase tournaments with an initial group play aren't [17].

Still, monotonicity only guarantees that no team wishes to unilaterally throw a match to improve their chances of winning, whereas one might also hope to guarantee that no two teams could fix the outcome of their match in order to improve the probability that one of them wins. While we have to go back further in history to find a clear instance of this kind of match-fixing, it did indeed result in a historical scandal. In the 1982 FIFA World Cup (again a two-stage tournament), Austria, West Germany, and Algeria were in the same group of four where two would advance. Algeria had already won two matches and lost one, Austria was 2-0, West Germany was 1-1, and the only remaining game was Austria vs. West Germany. Due to tie-breakers and the specific outcomes of previous matches, Austria would have been eliminated by a large West German victory, and West Germany would have been eliminated by a loss or draw. Once West Germany scored an early goal, *both* teams essentially threw the rest of the match, allowing both of them to advance at Algeria's expense [25]. While the incident was never formally investigated, many fans were confident the teams had colluded beforehand, and the event is remembered as the "disgrace of Gijón." Before being eliminated, Algeria had become the first African team to beat a European team at the World Cup, and also the first to win two games. West Germany went on to become the runners-up of the tournament.

Motivated by events like this, it is important also to design tournaments where no two teams can fix the outcome of their match and improve the probability that one of them wins. Altman and Kleinberg terms this property 2-Strongly Nonmanipulable (2-SNM), and showed that no tournament rule is both Condorcet-consistent and 2-SNM [1] (it was previously shown by Altman et. al. that no *deterministic* rule is both Condorcet-consistent and 2-SNM [2]).

In light of this, both works relax the notion of Condorcet-consistency and design tournament rules that are at least *non-imposing* (could possibly select each competitor as a winner) and 2-SNM [2], or α-Condorcet-consistent (if there is a Condorcet winner, she wins with probability at least α) and 2-SNM. While these relaxations are well-motivated for settings where pair-wise comparisons are only *implicitly* made, and not even necessarily learned in the end (e.g. elections), it is hard to imagine a successful sports competition format where a competitor could win all their matches and still leave empty handed. This happened during the 2008 NCAA Football Season. Utah went undefeated (#2, 13-0) in their region but

were not invited to the bowl game because critics deemed their schedule weak. They were eventually ranked second nation-wide and beat Alabama (#6, 12-2) in the Sugar Bowl, while Florida (#1, 13-1) beat Oklahoma (#5, 12-2) for the National Championship. This event prompted organizers to reconsider the process by which teams are invited to the National Championship game.

Motivated by match-based applications such as sporting events, where the outcome of pair-wise matches is *explicitly* learned and used to select a winner, we consider instead the design of tournament rules that are exactly Condorcet-consistent, but only approximately 2-SNM. Specifically, we say that a tournament rule is 2-SNM-α if it is *never* possible for two teams i and j to fix their match such that the probability that the winner is in $\{i, j\}$ improves by at least α. The idea behind this relaxation is that whatever motivates j to throw the match (perhaps j and i are teammates, perhaps i is paying j some monetary bribe, etc.), the potential gains scale with α. So it is easier to disincentivize manipulation (either through investigations and punishments, reputation, or just feeling morally lousy) in tournaments that are less manipulable.

1.1 Our Results

Our main result is a matching upper and lower bound of $1/3$ on attainable values of α for Condorcet-consistent 2-SNM-α tournament rules. The optimal rule that attains this upper bound is actually quite simple: a random single elimination bracket. Specifically, each competitor is randomly placed into one of $2^{\lceil \log_2 n \rceil}$ seeds, along with $2^{\lceil \log_2 n \rceil} - n$ byes, and then a single elimination tournament is played.

Proving a lower bound of $1/3$ is straight-forward: imagine a tournament with three players, A, B and C, where A beats B, B beats C, and C beats A. Then some pair must win with combined probability at most $2/3$. Yet, any pair could create a Condorcet winner by colluding, who necessarily wins with probability 1 in any Condorcet-consistent rule. Embedding this within examples for arbitrary n is also easy: just have A, B, and C each beat all of the remaining $n - 3$ competitors[1].

On the other hand, proving that a random single elimination bracket is optimal is tricky, but our proof is still rather clean. For any i, j in any tournament, we directly show that i can improve her probability of winning by at most $1/3$ when j throws their match using a coupling argument. For every deterministic single elimination bracket where i and j could potentially gain from manipulation (because i would be the champion if i beat j, but j would *not* be the champion even if j beat i), we construct *two* deterministic single elimination brackets where no potential exists (possibly because one of them will lose before facing each other, or because the winner would be in $\{i, j\}$ no matter the outcome of their match). For our coupling to be valid, we not only need each mapping to be invertible, but also for their images to be disjoint. Our coupling is necessarily somewhat involved in order to obtain this property, but otherwise we believe our proof is likely as simple as possible. Because the probability that j wins cannot possibly go up by throwing a match to i, this immediately proves that a random single elimination bracket is 2-SNM-1/3.

We also show that the Copeland rule, a popular rule that chooses the team with the most wins, is asymptotically 2-SNM-1, the *worst* possible. Essentially, the problem is that if all teams have the same number of wins, then any two can collude to guarantee that one of

---

[1] Interestingly, this lower-bound example is far from pathological and occurs at even the highest levels of professional sports (see [18], for instance).

them wins, no matter the tie-breaking rule. We further show that numerous other formats, (the Random Voting Caterpillar, the Iterative Condorcet Rule, and the Top Cycle Rule) are all at best 2-SNM-1/2. The same example is bad for all three formats: there is one superman who beats $n - 2$ of the remaining players, and one kryptonite, who beats only the superman (but loses to the other $n - 2$ players).

Our results extend to settings where the winner of each pairwise match is not determinist-ically known, but randomized (i.e. all partipants know that i will beat j with probability p_{ij}). Specifically, we show that any rule that is 2-SNM-α when all $p_{ij} \in \{0, 1\}$ is also 2-SNM-α for arbitrary p_{ij}. Clearly, any lower bound using integral p_{ij} also provides a lower bound for arbitrary p_{ij}, so as far as upper/lower bounds are concerned the models are equivalent. Of course, the randomized model is much more realistic, so it is convenient that we can prove theorems in this setting by only studying the deterministic setting, which is mathematically much simpler.

Finally, we consider manipulations among coalitions of $k > 2$ participants. We say that a rule is k-SNM-α if no set S of size $\leq k$ can *ever* manipulate the outcomes of matches between players in S to improve the probability that the winner is in S by more than α. We prove a simple lower bound of $\alpha = \frac{k-1}{2k-1}$ on all Condorcet-consistent rules, and conjecture that this is tight.

1.2   Related Works

The mathematical study of tournament design has a rich literature, ranging from social choice theory to psychology. The overarching goal in these works is to design tournament rules that satisfy various properties a designer might find desirable. Examples of such properties might be that all players are treated equally, that a winner is chosen without a tiebreaking procedure, or that a "most qualified" winner is selected [8, 20, 7, 19, 28, 15, 23]. See [14] for a good review of this literature and its connections to other fields as well.

Most related to our work are properties involving *strategic manipulation*. In the more general field of Voting Theory, there is a rich literature on the design of strategyproof mechanisms dating back to Arrow's Impossibility Theorem [3] and the Gibbard-Satterthwaite Theorem [9, 21, 10]. While tournaments are a very special case (voters are indifferent among outcomes where they do not win, voters can only "lie" in specific ways, etc.), tournament design indeed seems to inherit much of the impossibility associated with strategyproof voting procedures [1], [2].

Specifically, Altman et. al. proved that no deterministic tournament rule is 2-SNM and Condorcet-consistent, and Altman and Kleinberg proved that no randomized tournament rule is 2-SNM and Condorcet-consistent either [2, 1]. More recently, Pauly studied the specific two-stage tournament rule used by the World Cup (and Olympic Badminton, etc.) [17]. There, it is shown essentially that the problem lies in the first round group stage: no changes to the second phase can possibly result in a strategyproof [2] tournament.

To cope with their impossibility results, Altman et. al. propose a relaxation of Condorcet-consistence called *non-imposing*. A rule r is non-imposing if for all i, there exists a T such that player i wins with probability 1. They design a clever recursive rule that is non-imposing and 2-SNM for all $n \neq 3$. Interestingly, they also show that for $n = 3$ no such rule exists. Altman and Kleinberg consider a different relaxation called α-*Condorcet-consistent*. A rule r is α-Condorcet-consistent if whenever i is a Condorcet winner in T, we have their probability of winning T is at least α. They design a rule that is $2/n$-Condorcet-consistent and 2-SNM (in fact it is also k-SNM for all k), but conjecture that much better is attainable.

---

[2] See [17] for the specific notion of strategyproofness studied.

The two works above are most similar to ours in spirit: motivated by the non-existence of Condorcet-consistent and 2-SNM tournament rules, we relax one of the notions. These previous works relax Condorcet-consistency while maintaining 2-SNM exactly, and are most appropriate in settings where pairwise comparisons of players are only learned *implicitly* (or perhaps not at all) through the outcome and not *explicitly* as the result of matches. Instead, we relax the notion of 2-SNM and maintain the notion of Condorcet-consistency exactly. In settings like sports competitions where pairwise comparisons of players are learned explicitly through matches played, Condorcet-consistency is a non-negotiable desideratum. Therefore, we believe our approach is more natural in such settings.

Another line of work introduced by [4] considers a different kind of strategyproofness: how much control does the designer of a single-elimination tournament have over the winner? Can the designer efficiently find a bracket in such a way to maximize the likelihood that a player of their choice wins the tournament? The models in this area assume that the designer is given the probabilities p_{ij} that team i beats team j and the problem is known in the literature as *agenda control* when p_{ij} are real numbers and Tournament Fixing Problem (TFP) when all probabilities are 0 or 1.

On the negative side, it is known that for n-player tournaments it is NP-hard to decide whether or not there exists a seeding such that the probability of team k winning is at least δ, given k, δ, even if $\mathsf{p}_{ij} \in \{0, 0.5, 1\}$ for all i, j [27]. [26] show that the hardness results persist even for the TFP when the given team k is a king (for every team j, either k beats j or k beats a team that beats j) with at least $n/4$ wins, or a $3-$king (is at most 3 "wins" away from every team) that wins at least half of their games. Follow up work [13] shows that in the case of balanced single elimination brackets, it is still NP-hard to find a bracket that favors team k when the designer is allowed to bribe at most $(1 - \varepsilon) \log n$ of the teams to throw their respective matches.

On the positive side, there exist structural results that dictate when it is computationally efficient to find a tournament that favors a given team. [26] show conditions under which, for large enough tournaments, any sufficiently good team can be favored by the tournament seeding. Other results [13, 12] show conditions under which $3-$kings can be made into winners of single-elimination tournaments.

A large body of literature exists regarding manipulation and bribery in the context of voting rules. For an introduction, we recommend the reader consult chapter 7 of the handbook [16].

1.3 Conclusions and Future Work

Our work contributes to a recent literature on incentive compatible tournament design. While most previous works insist on strong incentive properties and relaxed fairness properties, such rules are inadequate for sporting events. Instead, we insist at least that events maintain Condorcet-consistency, and aim to relax strategyproofness as minimally as possible.

At a high level, our work suggests (similar to previous works), that single elimination brackets are desirable whenever incentive issues come into play. However, previous desiderata (such as those considered in [1]) don't necessarily rule out other tournament formats, like the Copeland rule, which is ubiquitous in tournaments (both as a complete format and as subtournaments in a two-phase format). In comparison, our work identifies single elimination brackets (2-SNM-1/3) as having significantly better strategic properties versus the Copeland rule (2-SNM-1).

Our work also identifies two practical suggestions when match-fixing is a concern that aren't explained by prior benchmarks. First, when hosting a single elimination tournament,

it might be desirable to release the exact bracket as late as possible. The idea is that as soon as the exact bracket is known, competitors have greater incentive to fix matches (in our model, up to three times as much), which presumably takes some time and organization. Obviously, there are more tradeoffs at play: a later release inconveniences athletes and fans, and (perhaps more importantly to the designers) could negatively impact ticket sales. But our work does at least identify match-fixing as a part of this tradeoff. Note that some Olympic events (such as Taekwondo) contest the entire competition in a single day at a single venue, so a delayed release may indeed be practical. We also note that a similar "fix" was applied after the 1982 World Cup: the last two matches in each group are now played at the same time to minimize the amount of information teams have when making potentially strategic decisions.

Additionally, our work suggests that even in the optimal tournament, hefty punishments for cheaters might be necessary in order to discourage match-fixing (even without taking gambling into consideration). In many sports, winning an Olympic gold can make a career. Unfortunately, our work suggests that punishments roughly on this order might be necessary in order to properly deter match-fixing.

Finally, we propose two directions for future work. First, while we obtain tight results for Condorcet-consistent 2-SNM-α rules, we only prove a lower bound of k-SNM-$\frac{k-1}{2k-1}$ for Condorcet-consistent rules and $k > 2$. We conjecture that this is tight, but unfortunately simulations indicate that all of the formats studied in our work do *not* achieve this bound. So it is an interesting open question to design a rule that does. Even partial results (of the form identified below) would require a new tournament format than those considered in this work.

▶ **Open Question 1.** *Does there exist a tournament rule that is Condorcet-consistent and k-SNM-$\frac{k-1}{2k-1}$ for all k? What about a family of rules $\mathcal{F}$ such that for all k, F_k is k-SNM-$\frac{k-1}{2k-1}$? What about a rule that is k-SNM-1/2 for all k?*[3]

It is also important to study what bounds are attainable in restricted versions of our probabilistic model (e.g. if for all i, j, the probability that i beats j lies in $[\epsilon, 1-\epsilon)$). Realistic instances at least have *some* non-zero probability of an upset in every match, but our lower bounds don't hold in this model. So it is interesting to see if better formats are possible.

▶ **Open Question 2.** *Is a random single elimination bracket still optimal among Condorcet-consistent rules (w.r.t. 2-SNM-α) if for all i, j, the probability that i beats j lies in $[\epsilon, 1-\epsilon]$? How does the optimal attainable α for Condorcet-consistent, 2-SNM-α tournament formats change as a function of ϵ?*

2   Preliminaries and Notation

In this section, we present notation used throughout the remainder of the paper. Where possible, we adopt notation from [1].

▶ **Definition 1.** A (round-robin) *tournament* T on n players is the set of outcomes of the $\binom{n}{2}$ games played between all pairs of distinct players. We write $T_{ij} = 1$ if player i beats player j and $T_{ij} = -1$ otherwise. We also let $\mathcal{T}_n$ denote the set of tournaments on n players.

▶ **Definition 2.** For a subset $S \subseteq [n]$ of players, two tournaments T and T' are *S-adjacent* if they only differ on the outcomes of some subset of games played between members of S.

---

[3] Note that $\frac{k-1}{2k-1} \to 1/2$ as $k \to \infty$.

In particular, two tournaments T and T' are $\{i, j\}$ adjacent if they only differ in the result of the game played between player i and player j.

▶ **Definition 3.** A *tournament rule* (or *winner determination rule*) $r : \mathcal{T}_n \rightarrow \Delta([n])$ is a mapping from the set of tournaments on n players to probability distributions over these n players (representing the probability we choose a given player to be the winner). We will write $r_i(T) = \Pr[r(T) = i]$ to denote the probability that player i wins tournament T under rule r.

Many tournament rules, while valid by the above definition, would be ill-suited for running an actual tournament; for example, the tournament rule which always crowns player 1 the winner. In an attempt to restrict ourselves to 'reasonable' tournament rules, we consider tournaments that obey the following two criteria.

▶ **Definition 4.** Player i is a *Condorcet winner* in tournament T if player i wins their match against all the other $n - 1$ players. A tournament rule r is *Condorcet-consistent* if $r_i(T) = 1$ whenever i is a Condorcet winner in T.

▶ **Definition 5.** A tournament rule r is *monotone* if, for all i, $r_i(T)$ does not increase when i loses a game it wins in T. That is, if i beats j in T and T and T' are $\{i, j\}$ adjacent, then if r is monotone, $r_i(T) \geq r_i(T')$.

Intuitively, this first criterion requires us to award the prize to the winner in the case of a clear winner (hence making the tournament a contest of skill), and the second criterion makes it so that players have an incentive to win their games. There are various other criteria one might wish a tournament rule to satisfy; many can be found in [1].

In this paper, we consider the scenario where certain coalitions of players attempt to increase the overall chance of one of them winning by manipulating the outcomes of matches within players of the coalition. The simplest case of this is in the case of coalitions of size 2, where player j might throw their match to player i. If T is the original tournament and T' is the manipulated tournament where j loses to i, then player i gains $r_i(T') - r_i(T)$ from the manipulation, and player j loses $r_j(T) - r_j(T')$ (in terms of probability of winning). Therefore, as long as $r_i(T') - r_i(T) > r_j(T) - r_j(T')$, it will be in the players' interest to manipulate. Equivalently, if $r_i(T') + r_j(T') > r_i(T) + r_j(T)$ (i.e., the probability either player i or j wins increases upon throwing the match), there is incentive for i and j to manipulate.

Ideally, we would like to choose a tournament rule so that, regardless of the tournament, there will be no incentive to perform manipulations of the above sort. This is encapsulated in the following definition from [1].

▶ **Definition 6.** A tournament rule r is *2-strongly non-manipulable (2-SNM)* if, for all pairs of $\{i, j\}$-adjacent tournaments T and T', $r_i(T) + r_j(T) = r_i(T') + r_j(T')$.

Unfortunately, no tournament rules exist that are simultaneously Condorcet-consistent and 2-strongly non-manipulable (this is shown in [1] and also follows from our lower bound in Section 3.1). As tournament designers, one way around this obstacle is to discourage manipulation. This discouragement can take many forms, both explicit (if players are caught fixing matches, they are disqualified/fined) and implicit (it is logistically hard to fix matches, it is unsportsmanlike). So the focus of this paper is to quantify *how manipulable* certain tournament formats are (i.e. how much can teams possibly gain by fixing matches), the idea being that it is easier to discourage manipulation in tournaments that are less manipulable.

▶ **Definition 7.** A tournament rule r is *2-strongly non-manipulable at probability α (2-SNM-α)* if, for all i and j and pairs of $\{i, j\}$-adjacent tournaments T and T', $r_i(T') + r_j(T') - r_i(T) - r_j(T) \leq \alpha$.

It is straightforward to generalize this definition to larger coalitions of colluding players.

▶ **Definition 8.** A tournament rule r is k-*strongly non-manipulable at probability* α (k-*SNM-*α*)* if, for all subsets S of players of size at most k, for all pairs of S-adjacent tournaments T and T', $\sum_{i \in S} r_i(T') - \sum_{i \in S} r_i(T) \leq \alpha$.

2.1   The Random Single-Elimination Bracket Rule

Our main result concerns a specific tournament rule we call the *random single-elimination bracket rule*. This rule can be defined formally as follows.

▶ **Definition 9.** A *single-elimination bracket* (or *bracket*, for short) B on $n = 2^h$ players is a complete binary tree of height h whose leaves are labelled with some permutation of the n players. The outcome of a bracket B under a tournament T is the labelling of internal nodes of B where each node is labelled by the winner of its two children under T. The winner of B under T is the label of the root of B under this labelling.

▶ **Definition 10.** The *random single-elimination bracket rule* r is a tournament rule on $n = 2^h$ players where $r_i(T)$ is the probability player i is the winner of B under T when B is chosen uniformly at random from the set of $n!$ possible brackets.

If n is not a power of 2, we define the random single-elimination bracket rule on n players by introducing $2^{\lceil \log_2 n \rceil} - n$ dummy players who lose to all of the existing n players.

It is straightforward to check that the random single-elimination bracket rule is both Condorcet-consistent and monotone. Our main result (Theorem 13) shows that in addition to these properties, the random single-elimination bracket rule is 2-SNM-1/3 (which is the best possible, by Theorem 11).

We give some examples of other common tournament rules in Section 3.4. While many of these rules are both Condorcet-consistent and monotone, we do not know of any which are additionally 2-SNM-1/3.

3   Main Result

3.1   Lower bounds for k-SNM-α

We begin by showing that no tournament rule is 2-SNM-α for $\alpha < 1/3$. A similar theorem appears as Proposition 17 in [1] (which states that $\alpha = 0$ is impossible).

▶ **Theorem 11.** *There is no Condorcet-consistent tournament rule on n players (for $n \geq 3$) that is 2-SNM-α for $\alpha < \frac{1}{3}$.*

Proof. Consider the tournament T on three players A, B, and C where A beats B, B beats C, and C beats A (illustrated in Figure 1). . Note that, while this tournament has no Condorcet winner, changing the result of any of the three games results in a Condorcet winner. For example, if A bribes C to lose to A, then A becomes the Condorcet winner.

If we have a tournament rule r that is 2-SNM-α, then combining this with the above fact gives rise to the following three inequalities.

$$
\begin{aligned}
r_A(T) + r_B(T) &\geq 1 - \alpha \\
r_B(T) + r_C(T) &\geq 1 - \alpha \\
r_C(T) + r_A(T) &\geq 1 - \alpha
\end{aligned}
$$
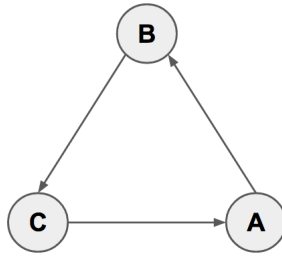
Figure 1 A tournament which attains the lower bound of $\alpha = 1/3$ for all tournament rules.

Together these imply $r_A(T) + r_B(T) + r_C(T) \geq \frac{3}{2}(1-\alpha)$. But $r_A(T) + r_B(T) + r_C(T) = 1$; it follows that $\alpha \geq \frac{1}{3}$, as desired.

We can extend this counterexample to $n > 3$ players by introducing $n - 3$ dummy players who all lose to A, B, and C; the argument above continues to hold.   ◄

We can use similar logic to prove lower bounds for the more general case of k-SNM-α.

▶ **Theorem 12.** *There is no Condorcet-consistent tournament rule on n players (for $n \geq 2k - 1$) that is k-SNM-α for $\alpha < \frac{k-1}{2k-1}$.*

Proof. Consider the following tournament T on the $2k - 1$ players labelled 1 through $2k - 1$. Each player i wins their match versus the $k - 1$ players $i + 1, i + 2, \ldots, i + (k-1)$, and loses their match versus the $k - 1$ players $i - 1, i - 2, \ldots, i - (k-1)$ (indices taken modulo $2k - 1$). Note that the coalition of players $S_i = \{i, i - 1, \ldots, i - (k-1)\}$ of size k can cause i to become a Condorcet winner if all players in the coalition agree to lose their games with i. If we have a tournament rule r that is k-SNM-α, then this implies the following $2k - 1$ inequalities (one for each $i \in [2k - 1]$):

$$\sum_{j \in S_i} r_j(T) \geq 1 - \alpha \tag{1}$$

Summing these $2k - 1$ inequalities, we obtain

$$k \sum_{j=1}^{2k-1} r_j(T) \geq (2k-1)(1-\alpha) \tag{2}$$

Since $\sum_{j=1}^{2k-1} r_j(T) \leq 1$, this implies that $\alpha \geq \frac{k-1}{2k-1}$, as desired. Again, it is possible to extend this example to any number of players $n \geq 2k - 1$ by introducing dummy players who lose to all $2k - 1$ of the above players.   ◄

3.2   Random single elimination brackets are 2-SNM-1/3

We now show that the random single elimination bracket rule is optimal against coalitions of size 2. The proof idea is simple; for every bracket B that contributes to the incentive to manipulate $r_i(T') + r_j(T') - r_i(T) - r_j(T)$ we will show that there are two that do not (in other words, for every scenario where team i benefits from the manipulation, there exist two other scenarios where the manipulation does not benefit either team).

▶ **Theorem 13.** *The random single elimination bracket rule is 2-SNM-1/3.*

Proof. Let $\mathcal{B}$ be the set of $n!$ different possible brackets amongst the n players. For a given tournament T and a given player i, write $\mathbb{1}(B, T, i)$ to represent the indicator variable which is 1 if i wins bracket B under the outcomes in T and 0 otherwise. Then we can write

$$r_i(T) = \frac{1}{|\mathcal{B}|} \sum_{B \in \mathcal{B}} \mathbb{1}(B, T, i).$$

Assume i loses to j in T. Then, if we let T' be the tournament that is $\{i, j\}$ adjacent to T, we can write the increase in utility resulting from j throwing to i

$$\frac{1}{|\mathcal{B}|} \sum_{B \in \mathcal{B}} \left(\mathbb{1}(B, T', i) + \mathbb{1}(B, T', j) - \mathbb{1}(B, T, i) - \mathbb{1}(B, T, j) \right). \tag{3}$$

Our goal is to show that this sum is at most $1/3$. Now, note that if i does not end up playing j in bracket B under T, i also does not play j in B under T' (and vice versa). In these brackets, $\mathbb{1}(B, T', i) = \mathbb{1}(B, T, i)$ and $\mathbb{1}(B, T', j) = \mathbb{1}(B, T, j)$, so these brackets contribute nothing to the sum in Equation 3. On the other hand, in a bracket B where i does play j, we are guaranteed that $\mathbb{1}(B, T, i) = 0$ and $\mathbb{1}(B, T', j) = 0$ (since i loses to j in T and j loses to i in T'). Therefore, letting $\mathcal{B}_{ij}$ be the subset of $\mathcal{B}$ of brackets where i meets j, we can rewrite Equation 3 as

$$\frac{1}{|\mathcal{B}|} \sum_{B \in \mathcal{B}_{ij}} \left(\mathbb{1}(B, T', i) - \mathbb{1}(B, T, j) \right).$$

Since $\mathbb{1}(B, T', i) \leq 1$, this is at most

$$\frac{1}{|\mathcal{B}|} \sum_{B \in \mathcal{B}_{ij}} \left(1 - \mathbb{1}(B, T, j) \right).$$

This final sum counts exactly the number of brackets B where i and j meet (under T, so j beats i) but j does not win the tournament. Call such brackets *bad*, and call the remaining brackets *good*. We will exhibit two injective mappings σ_i and σ_j from bad brackets to good brackets such that the ranges of σ_i and σ_j are disjoint. This implies that there are at least twice as many good brackets as bad brackets, and thus that the sum above is at most $1/3$, completing the proof.

For both mappings, we will need the following terminology. Consider a bad bracket B, and consider the path from j up to the root of this tree. The nodes of this path are labelled by players that j would face if they got that far. More specifically, j has some opponent in the first round. Should j win, j would face some opponent in the second round, then the third round, etc. all the way to the finals, and these opponents do not depend on the outcomes of any of j's matches. Then since B is a bad bracket, j does not win, and at least one of the players on this path can beat j. Choose the **latest** such player (i.e. the closest to the root) and call this player k. Note that k might *not* be the player that knocks j out of the tournament (that is the *first* player along this path who would beat j).

Suppose that i and j meet at height h of the bracket (i.e. in the hth round). Let B_i, B_j, B_k be the subtrees of height h that contain i, j, and k respectively. An example is shown in Figure 2.

We first describe the simpler of the two maps, σ_i. Define $\sigma_i(B)$ by swapping the subtrees B_i and B_k as shown in Figure 3. In this bracket j will lose to k before ever meeting i, so $\sigma_i(B)$ is good. Moreover σ_i is injective since we can construct its inverse. In $\sigma_i(B)$, j certainly would lose to k at height h before reaching i. Furthermore, because we didn't
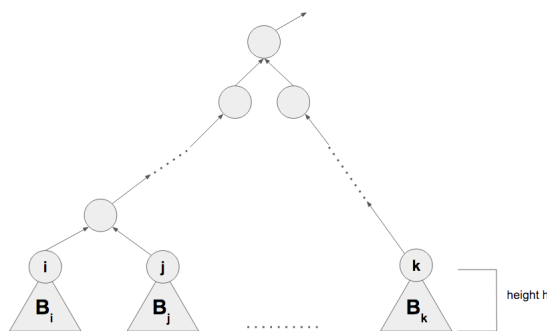
Figure 2 An example of a bad bracket B.
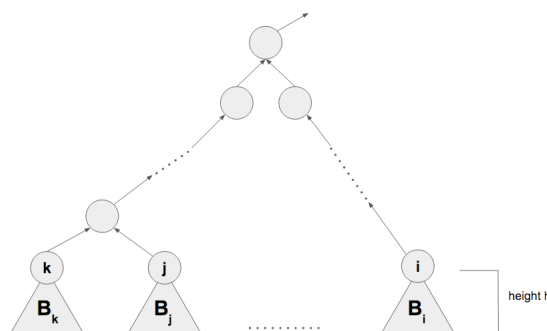
Figure 3 $\sigma_i(B)$.

change B_j at all, j still wins all of its first $h-1$ matches and makes it to k (because we started from a B where j makes it to i at height h). So we can identify k as the first player who beats j in $\sigma_i(B)$, learn the height h, and undo the swap of B_k and B_i.

We now describe the second map, σ_j. To construct $\sigma_j(B)$, begin by swapping the subtrees B_j and B_k (see Figure 4). Note that the bracket formed in this way is good; since we chose k to be the latest player on j's path to victory that can beat j, if j meets i, j will also beat all subsequent players and win the tournament (note that it is of course possible that j doesn't even make it to i, in which case $\sigma_j(B)$ is still good. But it is clear that *if j meets i*, then j will win the tournament, so $\sigma_j(B)$ is good in either case). Unfortunately, this map as stated is not injective; in particular, we cannot recover the height h to undo the swap as in the previous case.

The only reason we cannot uniquely identify k in the same way as when we invert σ_i is that i might meet some player k' at height $h' < h$ in B_i who also could beat j. So, intuitively, we would like to swap such players out with players who lose to j. Since j beats all of its opponents in B_j, B_j is an ample source of such players. We will therefore perform some additional 'subswap' operations, swapping subtrees of B_j and B_k so as to uniquely identify k as the first player i meets in $\sigma_j(B)$ who can beat j.

Specifically, for $0 \le h' < h$, let $a(h')$ be the opponent i plays at height h' in B_i, and let $B_i(h')$ be the subtree of B_i with root $a(h')$ (note that the player that i meets at height h' is the root of a subtree of height $h' - 1$, and that all these subtrees are disjoint). Similarly, let $b(h')$ be the opponent j plays at height h' in B_j, and let $B_j(h')$ be the subtree of B_j with root $b(h')$. To construct $\sigma_j(B)$ from B, first swap B_j and B_k. Then for each $h' \in [0, h)$ such that $a(h')$ would beat j, swap the subtrees $B_i(h')$ and $B_j(h')$. See Figure 5 for an illustration of a subswap operation.
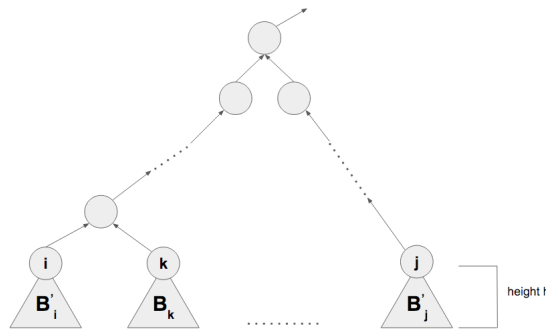
Figure 4 $\sigma_j(B)$.
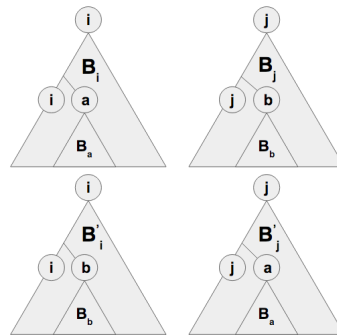
Figure 5 Subswap operation for σ_j.

Note that $\sigma_j(B)$ is still good; it is still the case that if j meets i, j will beat all subsequent players (all we have done in that part of the bracket is perhaps alter whether or not j will indeed meet i). On the other hand, since j makes it to height h in B_j, j can beat player $b(h')$ for all h', so k is now the first player i would encounter in $\sigma_j(B)$ who can beat j. From this, we can recover k and thus h, and undo the swap of B_i and B_j. To undo the subswaps, observe that because we started with a bad bracket B, that j must have beaten all opponents it faces in the first h rounds. Since all opponents on j's path who beat j at height less than h were necessarily put there by our subswap operations, we can just find all such opponents and swap them back out. This process inverts σ_j, thus proving that σ_j is injective.

Finally, note that in $\sigma_i(B)$, k must play j before either plays i, whereas in $\sigma_j(B)$, k must play i before either plays j. Therefore the ranges of σ_i and σ_j are disjoint, and this completes the proof.

For the reader aiming to understand our coupling argument better, Appendix A contains some specific examples.   ◀

3.3 Extension to randomized outcomes

Thus far we have been assuming that all match results are deterministic and known to the players in advance. Of course, this is not true in general; in real life, the outcomes of games are inherently unpredictable. It is perhaps imaginable that this unpredictability could increase the incentive to manipulate. In this section we show that this is not the case; a simple application of linearity of expectation shows that results about deterministic tournaments still hold for their randomized counterparts. We begin by defining a randomized tournament as follows.

▶ **Definition 14.** A *randomized tournament* $\mathcal{T}$ is a random variable whose values range over (deterministic) tournaments T. As shorthand, we will write $\mathbb{P}_{\mathcal{T}}(T)$ to represent the probability that $\mathcal{T} = T$.

Note that this definition accounts for the most straightforward generalization of tournament outcomes from deterministic to randomized, where for each match between players i and j we assign a probability p_{ij} to the probability that i beats j. This definition further allows for the possibility of correlation between matches (e.g., with some probability player i has a good day and wins all his matches, and with some probability he has a bad day and loses all his matches).

Manipulations in this randomized model are similar to manipulations in the deterministic model in that they effectively force the result of a match to a win or a loss. Formally, let $\sigma_{ij}(T)$ for a (deterministic) tournament T be the tournament formed by T but where i beats j (if i beats j in T, then $\sigma_{ij}(T) = T$). A tournament rule r is 2-SNM-α if for all i and j,

$$\mathbb{E}_{\mathcal{T}}\left[r_i(\sigma_{ij}(\mathcal{T})) + r_j(\sigma_{ij}(\mathcal{T})) - r_i(\mathcal{T}) - r_j(\mathcal{T})\right] \leq \alpha \qquad (4)$$

We then have the following theorem:

▶ **Theorem 15.** *If a rule r is 2-SNM-α in the deterministic tournament model, it is also 2-SNM-α in the randomized tournament model.*

Proof. Note that we can write the expectation in Equation 4 as

$$\sum_T \mathbb{P}_{\mathcal{T}}(T)\left(r_i(\sigma_{ij}(T)) + r_j(\sigma_{ij}(T)) - r_i(T) - r_j(T)\right)$$

If r is 2-SNM-α for deterministic tournaments, then each term in this sum is at most $\mathbb{P}_{\mathcal{T}}(T)\alpha$. It follows that this sum is at most α, and therefore r is also 2-SNM-α for randomized tournaments. ◀

It is straightforward to generalize the above definitions and result to the case of k-SNM-α.

3.4 Other tournament formats

Finally, there are many other tournament formats that are either used in practice or have been previously studied. In this section we show that many of these formats are more susceptible to manipulation than the random single elimination bracket rule; in particular, all of the following formats are at best 2-SNM-1/2.

By far the most common tournament rule for round robin tournaments is some variant of a 'scoring' rule, where the winner is the player who has won the most games (with ties broken in some fashion if multiple players have won the same maximum number of games). In voting theory, this rule is often called Copeland's rule, or Copeland's method [5].

▶ **Definition 16.** A tournament rule r is a *Copeland rule* if the winner is always selected from the set of players with the maximum number of wins.

We begin by showing that no Copeland rule can be 2-SNM-α for any $\alpha < 1$ (regardless of how the rule breaks ties).

▶ **Theorem 17.** *There is no Copeland rule on n players that is 2-SNM-α for $\alpha < 1 - \frac{2}{n-1}$.*

Proof. Assume to begin that $n = 2k + 1$ is odd, and let r be a Copeland rule on n players. Let T be the tournament where each player i beats the k players $\{i + 1, i + 2, \ldots, i + k\}$ but loses to the k players $\{i - 1, i - 2, \ldots, i - k\}$, with indices taken modulo n (similar to the tournament in the proof of Theorem 12).

Since $\sum_{i=1}^{n} r_i(T) = 1$, there must be some i such that $r_{i-1}(T) + r_i(T) \leq \frac{2}{n}$. On the other hand, if player $i - 1$ throws their match to player i, then player i becomes the unique Copeland winner (winning $k + 1$ games) and $r_i(T') = 1$. It follows that, for such a rule, if r is 2-SNM-α, then $\alpha \geq 1 - \frac{2}{n}$.

If n is even, then we can embed the above example for $n - 1$ by assigning one player to be a dummy player that loses to all teams. This immediately implies $\alpha \geq 1 - \frac{2}{n-1}$ in this case.   ◀

In [1], Altman and Kleinberg provide three examples of tournament rules that are Condorcet-consistent and monotone: the top cycle rule, the iterative Condorcet rule, and the randomized voting caterpillar rule. We prove lower bounds on α for each of these in turn. Interestingly, the same tournament provides all three lower bounds.

▶ **Definition 18.** The *superman-kryptonite* tournament on n players has i beat j whenever $i < j$, except that player n beats player 1. That is, player 1 beats everyone except for player n, who loses to everyone except for player 1.

Now we show that the superman-kryptonite tournament provides lower bounds against the tournament rules considered in [1].

▶ **Definition 19.** The *top cycle* of a tournament T is the minimal set of players who never lose to any other player. The *top cycle rule* is a tournament rule which assigns the winner to be a uniformly random element of this set.

▶ **Theorem 20.** *The top cycle rule on n players is not 2-SNM-α for any $\alpha < 1 - \frac{2}{n}$.*

Proof. Let T be the superman-kryptonite tournament on n players. The top cycle in T contains all the players, so $r_1(T) + r_n(T) = \frac{2}{n}$. However, if player n throws their match to player 1, player 1 becomes a Condorcet winner and $r_1(T') = 1$. It follows that $\alpha \geq 1 - \frac{2}{n}$. ◀

▶ **Definition 21.** The *iterative Condorcet rule* is a tournament rule that uniformly removes players at random until there is a Condorcet winner, and then assigns that player to be the winner.

▶ **Theorem 22.** *The iterative Condorcet rule on n players is not 2-SNM-α for any $\alpha < \frac{1}{2} - \frac{1}{n(n-1)}$.*

Proof. Let T be the superman-kryptonite tournament on n players. Note that no Condorcet winner will appear until either player 1 is removed, player n is removed, or all other $n - 2$ players are removed. If all the other $n - 2$ players are removed before players 1 or n (which occurs with probability $\frac{2}{n(n-1)}$), then player n wins. If this does not happen and player n is removed before player 1 (which occurs with probability $\frac{1}{2}\left(1 - \frac{2}{n(n-1)}\right) = \frac{1}{2} - \frac{1}{n(n-1)}$), then player 1 becomes the Condorcet winner and wins. Otherwise, player 1 will be removed before player n, while some players in 2 through $n - 1$ remain, and one of them will become the Condorcet winner (the remaining player in $\{2, \ldots, n - 1\}$ with lowest index). It follows that $r_1(T) = \frac{1}{2} - \frac{1}{n(n-1)}$ and $r_n(T) = \frac{2}{n(n-1)}$, so $r_1(T) + r_n(T) = \frac{1}{2} + \frac{1}{n(n-1)}$.

On the other hand, if player n throws their match to player 1, then again player 1 becomes a Condorcet winner and $r_1(T') = 1$. It follows that $\alpha \geq \frac{1}{2} - \frac{1}{n(n-1)}$. ◀

▶ **Definition 23.** The *randomized voting caterpillar rule* is a tournament rule which chooses a winner as follows. Choose a random permutation π of $[n]$. Start by matching $\pi(1)$ and $\pi(2)$, and choose a winner according to T. Then for all $i \geq 3$ match $\pi(i)$ with the winner of the most recent match. The player that wins the last match (against $\pi(n)$) is declared the winner.

▶ **Theorem 24.** *The randomized voting caterpillar rule on n players is not 2-SNM-α for any $\alpha < \frac{1}{2} - \frac{n-3}{n(n-1)}$.*

Proof. Let T be the superman-kryptonite tournament on n players. The only way player 1 loses is if either player n occurs later in π than player 1 (which happens with probability $\frac{1}{2}$) or if $\pi(n) = 1$ and $\pi(1) = 2$ and they play in the first round (which happens with probability $\frac{1}{n(n-1)}$). The only way player n can win is if $\pi(n) = n$ (i.e., they only play the very last game), in which case they will play player 1 and win (this happens with probability $\frac{1}{n}$). It follows that $r_1(T) = \frac{1}{2} - \frac{1}{n(n-1)}$ and $r_n(T) = \frac{1}{n}$, so $r_1(T) + r_n(T) = \frac{1}{2} + \frac{n-2}{n(n-1)}$.

On the other hand, if player n throws their match to player 1, then again player 1 becomes a Condorcet winner and $r_1(T') = 1$. It follows that $\alpha \geq \frac{1}{2} - \frac{n-2}{n(n-1)}$. ◀

─── **References** ───

1   Alon Altman and Robert Kleinberg. Nonmanipulable randomized tournament selections. In Maria Fox and David Poole, editors, *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, July 11-15, 2010.* AAAI Press, 2010. URL: `http://www.aaai.org/ocs/index.php/AAAI/AAAI10/paper/view/1703`.

2   Alon Altman, Ariel D. Procaccia, and Moshe Tennenholtz. Nonmanipulable selections from a tournament. In *Proceedings of the 21st International Jont Conference on Artifical Intelligence*, IJCAI'09, pages 27–32, San Francisco, CA, USA, 2009. Morgan Kaufmann Publishers Inc. URL: `http://dl.acm.org/citation.cfm?id=1661445.1661451`.

3   Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346, 1950. URL: `http://www.jstor.org/stable/1828886`.

4   John J. Bartholdi, Craig A. Tovey, and Michael A. Trick. How hard is it to control an election? *Mathematical and Computer Modelling*, 16(8):27 – 40, 1992. `doi:http://dx.doi.org/10.1016/0895-7177(92)90085-Y`.

5   A.H. Copeland. A 'reasonable' social welfare function. *Seminar on Mathematics in Social Sciences*, 1951.

6   S Cox. Tennis match fixing: Evidence of suspected match-fixing revealed, January 2016. http://www.bbc.com/sport/tennis/35319202.

7   Bhaskar Dutta. Covering sets and a new condorcet choice correspondence. *Journal of Economic Theory*, 44(1):63 – 80, 1988. `doi:10.1016/0022-0531(88)90096-8`.

8   Peter C. Fishburn. Condorcet social choice functions. *SIAM Journal on Applied Mathematics*, 33(3):469–489, 1977. `doi:10.1137/0133030`.

9   Allan Gibbard. Manipulation of voting schemes: a general result. *Econometrica*, 41(4):587–601, 1973.

10 Allan Gibbard. Manipulation of schemes that mix voting with chance. *Econometrica*, 45(3):665–681, 1977. URL: `http://www.jstor.org/stable/1911681`.

11 P Kelso. Badminton pairs expelled from london 2012 olympics after 'match-fixing' scandal, August 2012. .

12 Michael P. Kim, Warut Suksompong, and Virginia Vassilevska Williams. Who can win a single-elimination tournament? In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA.*, pages 516–522, 2016. URL: `http://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12194`.

13   Michael P. Kim and Virginia Vassilevska Williams. Fixing tournaments for kings, chokers, and more. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*, pages 561–567, 2015. URL: `http://ijcai.org/Abstract/15/085`.

14   Jean-Francois Laslier. *Tournament solutions and majority voting.* Number 7 in Studies in Economic Theory. Springer Verlag, 1997.

15   H. Moulin. Choosing from a tournament. *Social Choice and Welfare*, 3(4):271–291, 1986. URL: `http://www.jstor.org/stable/41105842`.

16   Hervé Moulin, Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. *Handbook of Computational Social Choice.* Cambridge University Press, 2016.

17   Marc Pauly. Can strategizing in round-robin subtournaments be avoided? *Social Choice and Welfare*, 43(1):29–46, 2014. `doi:10.1007/s00355-013-0767-6`.

18   B Phillips. The tennis triangle, July 2011. http://grantland.com/features/the-tennis-triangle/.

19   Ronald L. Rivest and Emily Shen. An optimal single-winner preferential voting system based on game theory, 2010.

20   Ariel Rubinstein. Ranking the participants in a tournament. *SIAM Journal on Applied Mathematics*, 38(1):108–111, 1980. URL: `http://www.jstor.org/stable/2100804`.

21   Mark Allen Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, 1975.

22   S Scherer. Italy breaks up soccer match-fixing network involving mafia, May 2015. http://www.bbc.com/news/world-europe-32793892.

23   T. Schwartz. Cyclic tournaments and cooperative majority voting: A solution. *Social Choice and Welfare*, 7(1):19–29, 1990. URL: `http://www.jstor.org/stable/41105932`.

24   B Sinclair. 12 arrested in esports match fixing scandal - report, October 2015. http://www.gamesindustry.biz/articles/2015-10-19-12-arrested-in-esports-match-fixing-scandal-report.

25   R Smyth. World cup: 25 stunning moments ... no3: West germany 1-0 austria in 1982, February 2014. URL: `http://www.theguardian.com/football/blog/2014/feb/25/world-cup-25-stunning-moments-no3-germany-austria-1982-rob-smyth`.

26   Isabelle Stanton and Virginia Vassilevska Williams. Rigging tournament brackets for weaker players. In *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, pages 357–364, 2011. `doi:10.5591/978-1-57735-516-8/IJCAI11-069`.

27   Thuc Vu, Alon Altman, and Yoav Shoham. On the complexity of schedule control problems for knockout tournaments. In *8th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009), Budapest, Hungary, May 10-15, 2009, Volume 1*, pages 225–232, 2009. `doi:10.1145/1558013.1558044`.

28   H. P. Young. Social choice scoring functions. *SIAM Journal on Applied Mathematics*, 28(4):824–838, 1975. URL: `http://www.jstor.org/stable/2100365`.

A   More Details on our Coupling Argument

In this appendix we present examples of bracket transformations. Recall that our transformations took as input any "bad" bracket, where player i eventually meets player j, *and* player j will lose to some player k in the future if she advances past i (and k is the latest such player). The players benefit from manipulating these brackets. We transformed them into "good" brackets, where either player j is eliminated before even meeting player i, or where

player j would be the champion conditioned on getting past i. The players have no incentive to manipulate these brackets.

We designed two injective transformations with disjoint images, σ_i and σ_j. σ_i was more straight-forward, but we include an example below anyway. σ_j was more complex. We include below an example showing that the complexity is necessary, and then an example of σ_j. All figures are at the end.

A.1   Example of the transformation $\sigma_i(B)$

Recall that σ_i essentially swaps the sub-brackets rooted at i and k. See Section 3.2 for a formal description.

Consider the partial bracket B_1 shown in Figure 6. Then, applying the transformation $\sigma_i(B_1)$ as described in our paper will yield the bracket B_1' shown in Figure 7. Note that this mapping is injective: by examining $\sigma_i(B)$, we see exactly where j is eliminated, and conclude that this must be where i met j in the original B.

A.2   Counterexample to a naive $\sigma_j(B)$

We could try using the same ideas in σ_i for σ_j: simply swap the subtrees rooted at k and j. Unfortunately, this mapping is not injective.

Consider the two brackets B_3, B_4 shown in Figure 8. Then applying this naive transformation will map these brackets to the same bracket (see Figure 9), showing that the mapping may not be injective. This motivates the need for the more involved transformation σ_j from Section 3.2.

Specifically, observe that in B_3, i meets j in round 2, so the depth-2 subtree rooted at k would get swapped with the depth-2 subtree rooted at j. In B_4, i meets j in round 1, so the single node i_1 would get swapped with the single node j. It is easy, but tedious, to complete this into a full tournament/bracket.

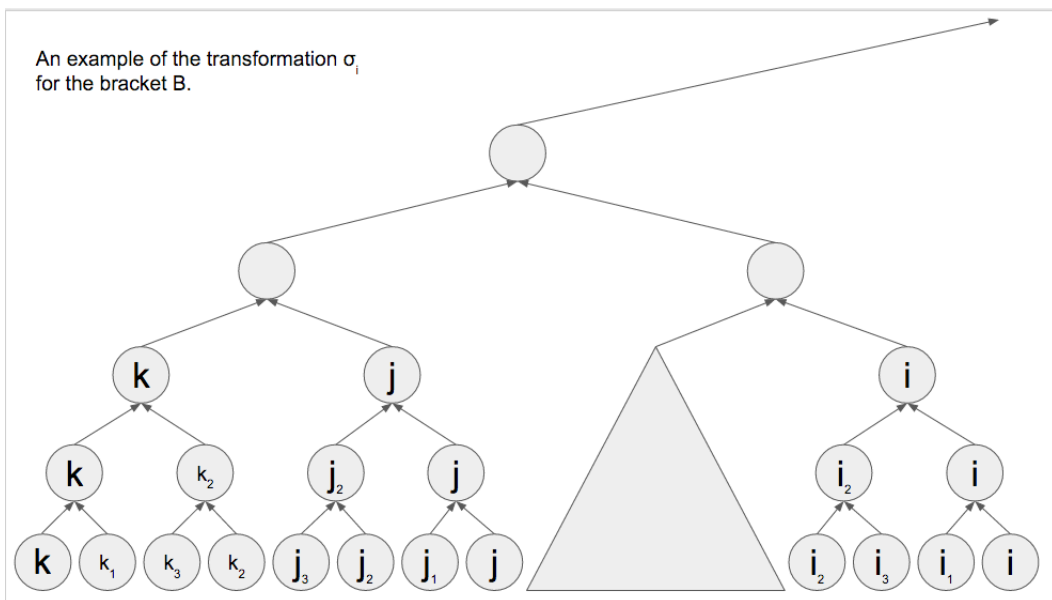## A.3   Example of the transformation $\sigma_j(B)$

Essentially, the problem with the naive transformation is that it's hard to recover where i met j in the original B just from the naive $\sigma_j(B)$. This is because maybe on its path to j, i met many other competitors who also would have beaten j, in addition to the k we swap in from the mapping. Our more involved transformation fixes this by additionally swapping all such competitors out of the subtree below i, so we can again recover where i met j in the original B.

Consider the partial bracket B_2 shown in Figure 10 and assume that in the tournament in case i_2 would beat j. Then, applying the transformation $\sigma_j(B_2)$ as described in our paper will yield the bracket B_2' shown in 11.

Note that this mapping is injective! First, we can recover where i met j in the original B by looking at where i first encounters someone who would beat j in $\sigma_j(B)$. Once we learn this, we also know that in the original B, j actually advanced this far in the tournament to meet i, so we know exactly which subtrees we need to un-swap with subtrees of i.
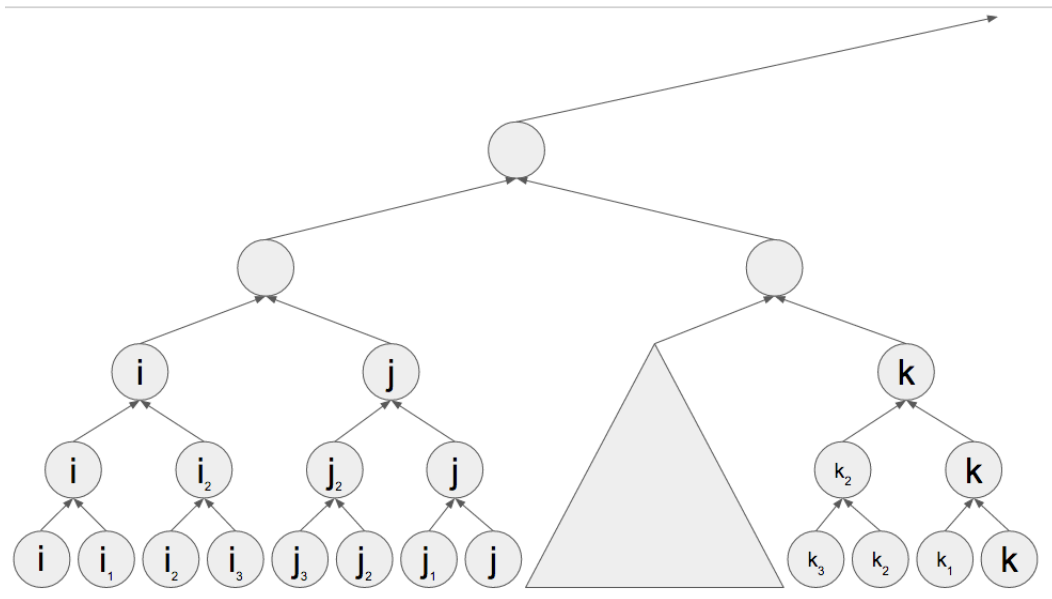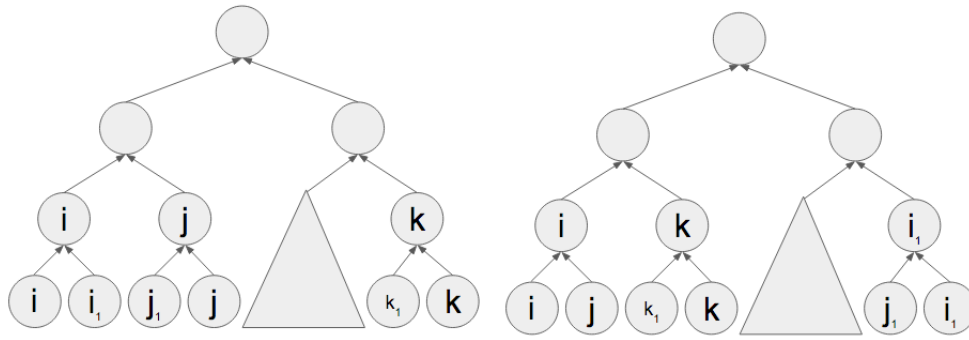
Figure 6 A partial bracket B_1.

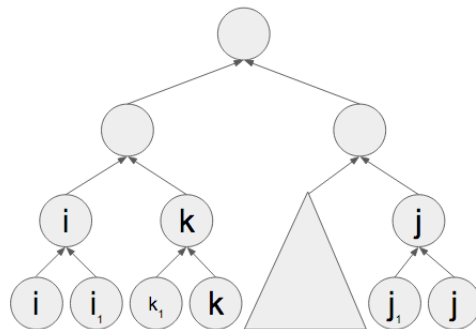**Figure 7** $\sigma_i(B_1)$.

An example of how the naive transformation can be non-injective.
Consider the following brackets B, B'. Applying the naive
transformation (i.e. swapping the trees or j and k) will produce the
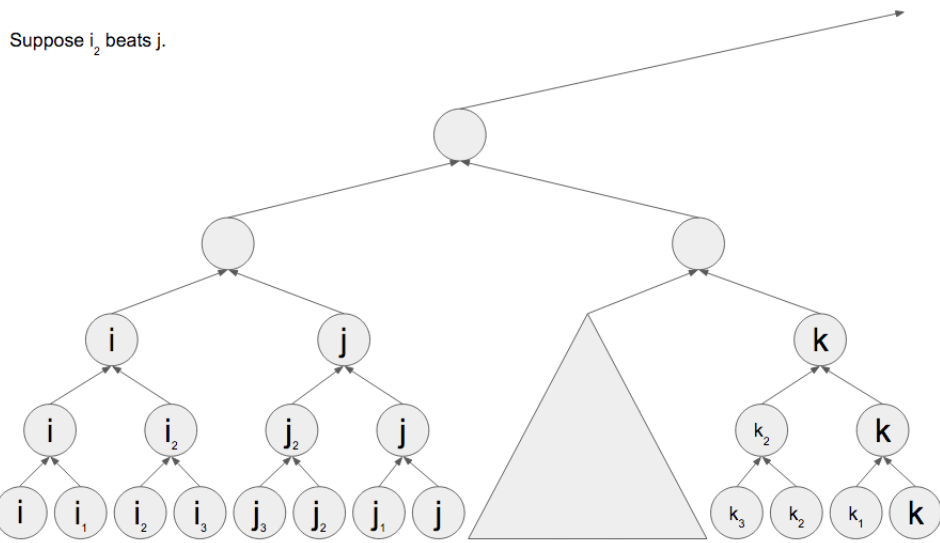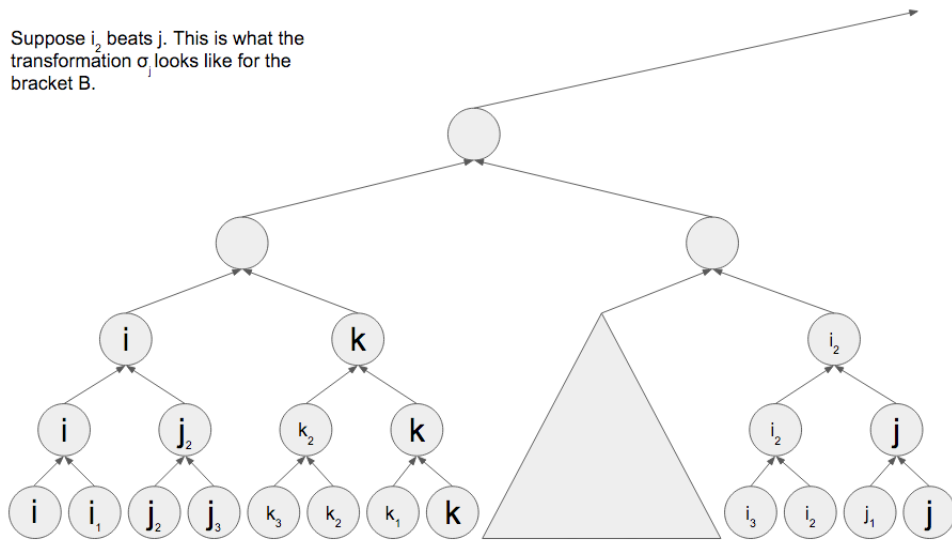same bracket.

Figure 8 Two partial brackets B_3, B_4.

Both brackets on the previous slide map to this sub-bracket. This
shows that the naive mapping is not injective.

Figure 9 Swapping the subtrees corresponding to j, k in both brackets above yields this bracket.

Suppose i_2 beats j.

Figure 10 A partial bracket B_2.

Suppose i_2 beats j. This is what the transformation σ_j looks like for the bracket B.

Figure 11 $\sigma_j(B_2)$.